

Elke Maria Deubzer

Die Ordnung im Kopf -

**Begriffliche Wissensstrukturen
zur Entwicklung
benutzerorientierter Anordnungen
von Funktionen im Raum**

München 2002

Die Ordnung im Kopf –
Begriffliche Wissensstrukturen zur Entwicklung
benutzerorientierter Anordnungen
von Funktionen im Raum

Elke Maria Deubzer

Vollständiger Abdruck der von der Fakultät für *Wirtschaftswissenschaften*
der Technischen Universität München zur Erlangung des akademischen Grades eines
Doktors der Philosophie (Dr. phil.)
genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. A. Büssing

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Heiner Bubb
2. Priv.-Doz. Dr. Gisa Aschersleben, Ludwig-Maximilians-Universität München

Die Dissertation wurde am 17.10. 2002 bei der Technischen Universität München eingereicht
und durch die Fakultät Wirtschaftswissenschaften am 09. 04. 2003 angenommen.

Vorwort

Die Arbeit ist für praktisch *und* wissenschaftlich arbeitende Psychologen, Ergonomen und Nutzerforscher geschrieben. Sie zeigt einen neuen Lösungsweg für das Problem „Wie muss man Funktionen im Raum anordnen, damit NutzerInnen sie schnell und einfach finden?“. Und sie liefert einen pragmatisch formulierten Leitfaden zur Entwicklung benutzerorientierter, erwartungsgerechter Anordnungen von Funktionen im Raum. Die Arbeit zeigt, wie durch die Anwendung von Grundlagentheorien und empirischen Untersuchungen diese praktische Problemstellung gelöst werden kann. Sie zeigt, dass Theorien und Gesetze nicht automatisch Rezepte für die Lösung eines Problems liefern, sondern dass es notwendig ist, Theorien umfassend zu verstehen und spezifisch auf das Problem anzuwenden. Das methodische Vorgehen unterscheidet sich somit von einem sonst üblichen, eher induktionistisch-orientierten Vorgehen in der Nutzerforschung und in den entsprechenden ISO-Normen zur benutzerorientierten Gestaltung.

Praktisch orientierten Köpfen wird die Arbeit an manchen Stellen zu detailliert sein; z.B. wenn zehn Theorien oder sieben Erhebungsmethoden daraufhin geprüft werden, welche sich zur Lösung des Anordnungsproblems eignen. Theoretisch und methodologisch orientierten Köpfen wird die Arbeit wahrscheinlich an manchen Stellen zu knapp sein und sie werden an manchen Stellen eine intensivere Auseinandersetzung mit der aktuellen Literatur vermissen. Ich habe versucht eine Ebene zu finden, die das Vorgehen und die vielen Entscheidungen aus den Bereichen der Nutzerforschung, Begriffsforschung, Wissenschaftstheorie, Methodenlehre und Statistik nachvollziehbar und überprüfbar macht und dennoch den roten Faden erkennen lässt.

Die Fragestellung dieser Arbeit habe ich in meiner Beratungstätigkeit gewonnen, in Aufträgen für Unternehmen aus dem Einzelhandel, der Automobilindustrie, für Hersteller von Konsumprodukten, Investoren aus dem Immobilienbereich, Instituten des Finanzwesens und vielen anderen. Und ich danke meinen Auftraggebern, insbesondere den Unternehmern Herrn Hans Feneberg, Herrn Peter Feneberg, Herrn Bernhard Kupsch und Herrn Hans Heiland von BMW, dass sie in neuartige, psychologische Projekte investierten und Vertrauen und Risikobereitschaft bei der Umsetzung zeigten.

Den Anspruch praktische Probleme wissenschaftlich und methodologisch reflektiert zu lösen, verdanke ich einmal meinem Vater und seinen Gesprächen über Wissenschaftstheorie und Philosophie am Mittagstisch und bei Spaziergängen. Aber auf diesen Weg haben mich auch viele Bekannte gebracht, die mich schief anschauten, wenn ich ihnen von meinen Projekten erzählte und die mich dann skeptisch fragten, was ich da als Psychologin Handfestes beitragen könne. So musste ich für mein berufliches Selbstverständnis die Frage klären: Wie muss ich in der Problemlösung vorgehen, damit die von mir vorgeschlagenen Maßnahmen auch funktionieren werden – vor allem wenn die Umsetzung hohe Kosten mit sich bringt!

Die Realisierung dieser Arbeit verdanke ich zum einen meinen Praktikantinnen Sonja Pedell, Barbara Köstler, Petra Kinateder, Bettina Wecker, Christiane Berker und Katja Kosubek, die meine Themen mit großem Engagement und Fleiß in ihren Diplom-Arbeiten und Projekten bearbeiteten.

Zum anderen verdanke ich die Realisierung meinen Betreuern. Sie ermutigten mich zur Integration meiner berufsbegleitenden, wissenschaftlichen Tätigkeiten in dieser Doktorarbeit und gaben mir eine Unterstützung, wie ich sie nicht erwartet hatte. Ich danke Frau Privatdozentin Dr. Gisa Aschersleben vom Max Planck Institut für psychologische Forschung für Ihre liebevolle, prägnante und fachlich kompetente Auseinandersetzung mit meinem Ideen und Texten. Ich danke Herrn Professor Dr. Heiner Bubb, Leiter des Fachbereiches Ergonomie an der Technischen Universität München für seine Unterstützung und seine umsichtige, fachliche Kritik meiner Ideen und Ausführungen. Und ich danke Herrn Professor Dr. Wolfgang Prinz, Leiter des Max Planck Instituts für psychologische Forschung, für seine Offenheit und Bereitschaft mir als angewandte arbeitende Psychologin die Teilnahme an den Doktoranten-Veranstaltungen und die Benützung der Institutsbibliothek zu ermöglichen, was mir eine große Unterstützung war.

Zuletzt danke ich Frau Nanni Deubzer, Frau Ingola Lammers, Frau Christina Salaba und Herrn Rigobert Böck, die diese Arbeit auf Punkt und Komma überarbeiteten bzw. mir halfen sie in eine professionelle Fassung zu bringen.

Gliederung

Vorwort	III
0. Zusammenfassung der Arbeit	1
1. Einführung in die Problemstellungen dieser Arbeit	5
2. Entwicklung des methodischen Vorgehens zur Lösung des praktischen Problems	11
2.1 Aktuelles methodisches Vorgehen in der Literatur	11
2.1.1 Ursprünge, Ziele und Kontext nutzerorientierter Maßnahmen	11
2.1.2 Methodologische Ansätze in der Nutzerforschung	14
2.1.3 Analyse von ISO-Normen und aktueller Literatur	16
2.1.4 Analyse von zwei ergonomischen Lehrbüchern	24
2.1.5 Zusammenfassung und Konsequenzen	26
2.2 Wissenschaftstheoretische Überlegungen zum Vorgehen	29
2.2.1 Zum Verhältnis von Theorie und Anwendung	29
2.2.2 Erste methodologische Konsequenzen	32
2.3 Entwicklung eines methodischen Vorgehens	33
2.3.1 Methodologie der Anwendungsentwicklung in Analogie zur Methodologie der Theorieentwicklung	34
2.3.2 Darstellung des methodischen Vorgehens	35
2.4 Zum Vergleich: Methodisches Vorgehen im Maschinenbau	42
2.5 Zusammenfassung	46
3. Problemanalyse und Bestimmung möglicher Theoriebereiche	49
3.1 Aktuelle Strategien, um Suchprozesse zu erleichtern	49
3.2 Formale Analyse von Anordnungen	51
3.3 Ergebnisse einer Explorationsstudie zum Suchverhalten	53
3.4 Erste Zuordnung zu möglichen Theoriebereichen	57
3.4.1 Erwartungen in der Inferenzforschung	57
3.4.2 Prüfung verschiedener Formen der Wissensrepräsentation	61
3.4 Zusammenfassung und weiteres Vorgehen	65

4. Auswahl einer geeigneten Theorie für das Anordnungsproblem	67
4.1 Einführung in die Welt der Begriffe	67
4.2 Kriterien für die Auswahl einer geeignete Theorie	69
4.3 Darstellung und Prüfung von Begriffstheorien	71
4.3.1 Exemplartheorien: Begriffe als Mengen erlebter Exemplare	71
4.3.2 Prototypentheorien: Begriffe als Prototypen	73
4.3.3 Gegenstandsorientierte Merkmalsrepräsentationstheorien	77
4.3.4 Verhaltensorientierte Merkmalsrepräsentationstheorien	81
4.3.5 Begriffe in der Theorie der Antizipativen Verhaltenssteuerung	86
4.3.6 Überblick über das Ergebnis der Prüfung	89
4.4 Darstellung und Prüfung von Begriffsstrukturtheorien	91
4.4.1 Semantische Netzwerke	93
4.4.2 Propositionale Netzwerke und Andersons ATC-Modelle	98
4.4.3 Schemata-Modelle	104
4.4.4 Hierarchische und ereignisbezogene Begriffsstrukturen	109
4.4.5 Begriffsstrukturen im Modell der Antizipativen Verhaltenssteuerung	116
4.4.6 Überblick über das Ergebnis der Prüfung	118
4.5 Zusammenfassung	120
5. Anwendung der Theorie auf das Anordnungsproblem	121
5.1 Anwendung von drei Gesetzen auf das Anordnungsproblem	121
5.2 Zusammenfassung: Problemlösungsweg und Hypothesen	130
6. Entwicklung des Untersuchungsdesigns	133
6.1 Untersuchungsvariablen und Stichprobenanforderungen	133
6.2 Bestimmung der Erhebungsmethoden	135
6.2.1 Anforderungen an die Methoden	135
6.2.2 Übersicht über Methoden der Begriffsforschung	137
6.2.3 Nicht geeignete Methoden	138
6.2.4 Aufgaben mit Ähnlichkeitsurteilen als geeignete Methoden	145
6.3 Bestimmung der Auswertungsmethoden	149
6.3.1 Metrische und nonmetrische Multidimensionale Skalierung	150
6.3.2 Hierarchische Clusteranalyse	154
6.4. Übersicht über Methoden für den Vorversuch	157
6.5 Zusammenfassung	159

7. Darstellung von vier empirischen Studien	161
7.1. Ermittlung einer nutzerorientierten Anordnung von Einzelhandelsgeschäften für Einkaufszentren	162
7.1.1 Einleitung	162
7.1.2 Vorversuch zur Auswahl typischer Einzelhandelsgeschäfte	164
7.1.3 Hauptversuch	166
7.1.4 Interpretation der Ergebnisse	170
7.1.5 Beurteilung der verwendeten Methoden	173
7.1.6 Zusammenfassung	174
7.2. Ermittlung einer nutzerorientierten Anordnung von Serviceeinrichtungen für Flughäfen	175
7.2.1 Einleitung	175
7.2.2 Hauptversuch:	177
7.2.3 Interpretation der Ergebnisse	187
7.2.4 Beurteilung der verwendeten Methoden	191
7.2.5 Zusammenfassung	191
7.3. Ermittlung einer käuferorientierten Anordnung von Zimmerpflanzen für Verkaufsflächen	193
7.3.1 Einleitung	193
7.3.2 Vorversuch: Auswahl der Pflanzen für den Hauptversuch	197
7.3.3 Hauptversuch	200
7.3.4 Nachversuch: Property Fitting Analyse	211
7.3.5 Befragung und Literaturrecherche für die Interpretation	216
7.3.6 Gesamtschau und Interpretation der Ergebnisse	216
7.3.7 Beurteilung der verwendeten Methoden	224
7.3.8 Zusammenfassung	225
7.4. Ermittlung einer käuferorientierten Anordnung von Süßgebäck und Überprüfung der Ergebnisse in einem Feldexperiment	227
7.4.1 Einleitung	227
7.4.2 Input-Studie: Erhebung der Begriffsstruktur von Süßgebäck	230
7.4.3 Evaluations-Studie: Überprüfung der Begriffsstruktur-orientierten Anordnung von Süßgebäck auf das Kundenverhalten	242
7.4.4 Gesamtschau und Interpretation der Ergebnisse	253
7.4.5 Beurteilung der verwendeten Methoden	256
7.4.6 Zusammenfassung	257

8. Interpretation der Ergebnisse der vier Studien in Hinblick auf die Hypothesen	259
8.1 Gesamtinterpretation der Ergebnisse zu Hypothese 1	259
8.2 Gesamtinterpretation der Ergebnisse zu Hypothese 2	261
8.3 Gesamtinterpretation der Ergebnisse zu Hypothese 3	267
8.4 Gesamtinterpretation der Ergebnisse zu Hypothese 4	268
8.5 Zusammenfassung	269
9. Auswertung der Ergebnisse in Hinblick auf die verwendete Theorie	271
9.1 Ein Untersuchungsthema für die Forschung	271
9.2 Weiterentwicklung der Begriffstheorie von Klix et al.	273
9.3 Differenzierung der Prognose zur semantischen Ähnlichkeit von Objekten	283
9.4 Zusammenfassung	284
10. Leitfaden für die Entwicklung nutzerspezifischer und erwartungsgerechter Anordnungen von Objekten und Funktionen im Raum	285
10.1 Leitfaden zur Entwicklung empirisch fundierter Vorgaben	285
10.2 Beschreibung eines spekulativen Vorgehens	298
10.3 Zusammenfassung	301
11. Ausblick	303
<i>Literaturverzeichnis</i>	307

0. Zusammenfassung der Arbeit

Wie muss man Objekte und Funktionen im Raum anordnen, damit NutzerInnen sie einfach und intuitiv finden? Die vorliegende Arbeit startet mit dieser praktischen Problemstellung und spannt einen Bogen über die Wissenschaftstheorie, Methodologie bis hin zur Kognitionsforschung. Und es wird gezeigt, dass durch dieses Gespann von Theorie und Praxis ein fruchtbarer Lösungsweg entsteht: Wenn die Anordnung der Objekte im Raum der Ordnung der Begriffsstrukturen der NutzerInnen entspricht, finden diese einfacher, was sie suchen. Die Arbeit mündet in einen theoretisch und empirisch fundierten Leitfaden für die Entwicklung nutzerspezifischer, erwartungsgerechter Anordnungen.

Das *erste Kapitel* führt in die Gestaltungsprobleme der Praxis ein: Wie müssen Geschäfte in einem Einkaufszentrum oder Serviceeinrichtungen in einem Flughafen angeordnet sein, damit die NutzerInnen sie schnell und einfach finden? Wie müssen Artikel in einem Supermarkt oder Pflanzen in einem Gartencenter nutzerspezifisch und erwartungsgerecht angeordnet sein? Die Literatur zeigt nämlich, dass Leitsysteme und Übersichtspläne üblich sind, aber das Suchproblem nicht zufrieden stellend lösen, weil sie selbst zu viele Probleme mit sich bringen. Um eine neue Problemlösung zu finden, muss deshalb anders an das Problem heran gegangen werden.

Im *zweiten Kapitel* wird daher zunächst das methodologische Vorgehen geklärt. Das Vorgehen soll wissenschaftlichen Standards entsprechen. Zunächst wird gezeigt, dass in ISO-Normen und Literatur der Nutzerforschung ein eher induktionistisch-orientiertes – und damit nicht zufrieden stellendes – Vorgehen empfohlen wird. Dann wird in Anlehnung an die Wissenschaftstheorie und Methodologie des Forschens ein deduktionistisch-orientiertes, theoretisch und empirisch fundiertes Vorgehen zur Problemlösung entwickelt: Es sieht zunächst eine theoretische Fundierung, d.h. die Suche nach einer – zum Problem – passenden Theorie vor, dann eine spezifische Anwendung der Theorie auf das praktische Problem. Daraus ergeben sich zwei hypothesengeleitete, empirische Untersuchungen an entsprechenden NutzerInnen. Die erste Studie ist eine Input-Studie, um nutzerspezifische, „erwartungsgerechte“

Anordnungsvorgaben zu entwickeln. Die zweite Studie ist eine Evaluations-Studie, um die Anordnungslösung dann am Nutzerverhalten zu überprüfen.

Im *dritten Kapitel* startet der konkrete Problemlösungsprozess mit einer Explorationsstudie und ersten theoretischen Einordnungen. Die Ergebnisse der Explorationsstudie zeigen, dass Nutzer eine Erwartung darüber haben, wo etwas zu finden ist. Die Sichtung der Kognitionsliteratur zeigt: Erwartungen über den Findeort können als spontane, automatische Schlussfolgerungen (Inferenzen) betrachtet werden. Sie werden auf der Basis von Begriffsstrukturen gebildet. Der Inhalt der Erwartungen wird somit durch den Inhalt und den – durch Gesetzmäßigkeiten bedingten – Aufbau der Begriffsstrukturen bestimmt. Die Begriffsstrukturen scheinen der Schlüssel zu einer erwartungsgerechten Anordnung zu sein: Die Objekte bzw. Funktionen sollten im Raum entsprechend ihrer begrifflichen Beziehungen in der Begriffsstruktur angeordnet sein. Somit wird eine Theorie gesucht, die die Frage beantwortet: Nach welchen Kriterien sind Begriffe im Gedächtnis „angeordnet“?

Im *vierten Kapitel* werden zehn Theorien über Begriffe und Begriffsstrukturen gesichtet und eine zum Problem passende ausgewählt. Die Theorie von Klix (1971, 1992, 1998), Hoffmann (1980, 1983, 1986) et. al. zeigt sich geeignet.

Im *fünften Kapitel* wird diese Theorie auf das Problem angewendet. Der Problemlösungsweg lautet: es ist zunächst die Begriffsstruktur - definiert als begriffliche Ähnlichkeit der Objekte bzw. Funktionen - zu erheben; sie ist dann in eine räumliche Anordnung zu übersetzen: was begrifflich nahe (semantisch ähnlich) ist, sollte räumlich nahe sein; was begrifflich fern (semantisch unähnlich) ist, sollte räumlich fern sein. Die Theorie lässt drei Gesetzmäßigkeiten ableiten, die Prognosen bzw. Hypothesen für die begrifflich-semantische Nähe der Objekte und Funktionen formulieren lassen: a) Die Begriffsstruktur ist abhängig von der Handlungsvorerfahrung – Personen einer Nutzergruppe stimmen in der Begriffsstruktur überein. b) Der Begriffsstruktur liegen handlungsbezogene und nicht kategoriale Kriterien zugrunde. c) Die Begriffsstruktur ist stabil und nicht abhängig vom aktuellen Handlungsziel. d) Anordnungen, die mit der Begriffsstruktur der NutzerInnen korrespondieren, unterstützen deren Such- und Findeprozesse.

Im *sechsten Kapitel* wird basierend auf den Hypothesen ein prinzipielles Untersuchungsdesign entwickelt. Für die Erhebung der semantischer Ähnlichkeit von Objekten und Funktionen werden Sortiertechniken gewählt, für die Auswertung die Nonmetrische Multidimensionale Skalierung und die Hierarchische Clusteranalyse.

Im *siebten Kapitel* werden vier praktische Untersuchungen (Einkaufszentrum, Flughafen, Süßgebäckregal und Zimmerpflanzen) vorgestellt, in der jeweils zwischen 40 und 60 entsprechende NutzerInnen bzw. KäuferInnen ca. 25-30 anzuordnende Objekte bzw. Funktionen auf semantische Ähnlichkeit beurteilen. Die begriffsstrukturorientierte Anordnung wird in einem Feldexperiment evaluiert.

Im *achten Kapitel* werden die Ergebnisse aller vier Studien übergreifend in Hinblick auf die Hypothesen interpretiert. Es zeigt sich, dass a) die Begriffsstrukturen personenübergreifend stabil und – wenn auch nur in einem geringen Maße – von den spezifischen Handlungserfahrungen der Personen (FachverkäuferInnen versus KundInnen, Vielflieger versus Wenigflieger) abhängig sind: Personen mit ähnlichen Handlungserfahrungen haben Übereinstimmungen in den semantischen Ähnlichkeitsbeurteilungen; Personen mit unterschiedlichen Handlungserfahrungen unterscheiden sich in den Ähnlichkeitsbeurteilungen; b) Der Ähnlichkeitsbeurteilung liegen handlungsbezogene Kriterien zugrunde. Die Begriffsstruktur ist kein Abbild der objektiven Realität. Eine Analyse verweist sogar auf handlungs“ziel“bezogene Kriterien. c) Die Instruktion hat keinen Einfluss auf die Ähnlichkeitsbeurteilung. d) Bei einer Anordnung von Objekten im Raum, die der Begriffsstruktur der NutzerInnen entspricht, benötigen die Nutzer weniger Zeit und haben kürzere Wege.

Im *neunten Kapitel* werden die Untersuchungsergebnisse zu einer Weiterentwicklung der Begriffstheorie von Klix, Hoffmann et al. benützt. Ausgangspunkt der neuen Theorie sind die Annahmen, dass Begriffsbildung ein Ergebnis von Handlungs-, Lern- und Merkmalsabstraktionsprozessen ist und dass Handlungsziele eine stark ordnende Funktion haben. Es zeigt sich, dass hiermit schlüssige Annahmen über partonomische (ereignisbezogene) und taxonomisch-hierarchische Begriffsstrukturen gebildet werden können. Die neuen Annahmen korrespondieren mit den Untersuchungsergebnissen und lassen semantische Ähnlichkeitsbeurteilung differenzierter prognostizieren: semantisch

ähnlich ist, was in einem Handlungskontext und/oder für das gleiche Handlungsziel verwendet wird, d.h. funktional äquivalent ist.

Im *zehnten Kapitel* wird abschließend ein allgemeiner Leitfaden für Praktiker zur Entwicklung nutzerspezifischer, erwartungsgerechter Anordnungen zusammengefasst: Es werden kurz die theoretische Fundierung, die drei Hypothesen für die semantische Ähnlichkeit dargestellt und für zwei hypothesengeleitete, empirische Untersuchungen das Untersuchungsdesign beschrieben: für die Input-Studie mit NutzerInnen zur Ermittlung der begrifflich-semantischen Ähnlichkeit der anzuordnenden Objekten bzw. Funktionen, deren Ergebnisse dann zu Vorgaben für die Gestalter (Architekten, Designer, Konstrukteure) formuliert werden. Und das Untersuchungsdesign für die Evaluationsstudie, mittels der dann Gestaltungsvarianten an NutzerInnen daraufhin überprüft werden, welche Lösung das Such- und Findeverhalten tatsächlich am Besten unterstützt.

Die Arbeit endet im *elften Kapitel* mit einem Ausblick auf neue Fragestellungen für die Nutzerforschung und die Entwicklung nutzerorientierter Problemlösungen.

1. Einführung in die Problemstellungen dieser Arbeit

Diese Arbeit behandelt zwei Probleme: ein praktisches und ein methodologisches Problem. Das praktische Problem lautet: Wie ordnet man Objekte oder Funktionen so an, dass die NutzerInnen sie schnell und einfach finden können? Das methodologische Problem lautet: Wie muss man aus wissenschaftlicher Sicht vorgehen, um nutzerorientierte, erwartungsgerechte Anordnungen zu entwickeln?

Das praktische Problem

Ein Lebensbereich mit Anordnungsproblemen ist z.B. der *Handel*. Wer in Lebensmittelmärkten einkauft, der kennt die verzweifelte Suche nach Tomatenmark oder Pfefferkörnern, nach Mozzarella oder Brühwürfeln. Kein Wunder, wenn sich hier Such- und Findeprobleme einstellen. Schließlich hat schon ein durchschnittlicher Supermarkt 4.000 bis 6.000 Lebensmittelartikel und ein Verbrauchermarkt hat zwischen 25.000 und 40.000 Artikel. Verkaufsflächengestalter und Regalbelegungsexperten aus der Zentrale einer Lebensmittel-Einzelhandelskette machen sich in der Regel zwar sehr viel Mühe bei der Regalbelegung (so nennt man die Anordnung der Artikel im Regal), doch nicht immer steht das Kriterium: „einfach für die Kundin zu finden“ an oberster Stelle. Vielmehr wird die Regalbelegung von Kriterien wie Spanne, Absatz, Herstellerkonditionen und von ausgefeilten Regalbelegungsprogrammen bestimmt, die mit den Daten der Scannerkassen gefüttert werden (vgl. Kröber-Riehl 1999).

Die Wünsche der Kundinnen und Kunden klingen dagegen meist sehr einfach: „Ach, würden die nicht immer alles neu umräumen, dann würde ich mich zurechtfinden!“. Implizit liegt hier die Meinung zugrunde, dass eine konstante Anordnung und damit eine Gewöhnung bzw. ein Lernen der Schlüssel zur Lösung ist. Doch es wird angezweifelt, dass jede beliebige Anordnung – Hauptsache der Nutzer kennt sie seit Jahren – zu gleich guten Such- und Findeleistungen führt.

In zwei Studien wird das Problem der „kundenorientierten“ Anordnung von Artikeln im Handel untersucht. Als „kundenorientiert“ oder besser „kundInnenorientiert“ wird eine Anordnung definiert, die den Erwartungen der KundInnen entspricht. Mit der

Schreibweise KundInnen soll daran erinnert werden, dass nicht nur Männer sondern auch Frauen einkaufen. In den meisten Geschäften sind die Kunden sogar überwiegend Frauen. Die Kurzform soll ein lästiges Ausschreiben von Kundinnen und Kunden ersparen. Doch da sie künstlich ist, wird sie nicht durchgängig angewendet. Eine Studie beschäftigt sich beispielhaft mit der „kundInnenorientierten“ Anordnung von Süßgebäck und Keksen im Supermarkt. Die andere Studie beschäftigt sich mit der „kundInnenorientierten“ Anordnung von Zimmerpflanzen im Gartencenter.

Der *zweite Lebensbereich* mit „Anordnungsproblemen“ sind *Gebäude*, vor allem öffentliche. Wer schon einmal in einem Flughafen war, kennt ganz sicher die hilflose Suche nach dem richtigen Ticketschalter, das sich Durchfragen zum Meeting-Point, die Suche nach einer Toilette oder den Ärger beim Nicht-Finden des Informationsschalters. Ein anderes Beispiel sind Einkaufszentren. Hier gibt es ähnliche Probleme mit der Suche nach dem richtigen Geschäft: Welche Geschäfte erwartet der Kunde am Eingang des Einkaufszentrums und welche in der Tiefe des Gebäudes? Welche Läden braucht ein Bäcker in der Nachbarschaft und welche ein Schuhgeschäft? Wo ist die KundIn bereit, um sich mit den Angeboten eines Schmuckgeschäftes zu beschäftigen und wo mit denen eines Sportgeschäftes? Bei beiden Gebäudetypen stellt sich die Frage, wie die vielen verschiedenen Einrichtungen und Geschäfte angeordnet werden müssen, damit der Nutzer sie einfach und spontan findet und die Suche nicht zu einem orientierungslosen Zickzack-Lauf wird.

Eine Voraussetzung für florierende Geschäfte ist die richtige Lage. Richtige Nachbarn beleben das Geschäft. Sie ziehen gemeinsam KundInnen an. Entsprechend den Ergebnissen aus eigenen, unveröffentlichten Untersuchungen sind dann auch die KundInnenströme harmonisch. Die KundInnen müssen nicht lange suchen und im Zickzack laufen. Sie finden die Geschäfte, die sie suchen, in einer harmonischen Abfolge. Für die KundInnen bedeutet ein schnelles Finden bessere Orientierung (vgl. Purcell 1986), ein größeres Sicherheitsgefühl (vgl. Flade 1990) und weniger schmerzende Füße aufgrund von Irrwegen. „Falsche“ Nachbarn behindern sich m.E. gegenseitig. KundInnen können ausbleiben und in der Folge die Umsätze. Dieses Wissen über eine „kun-

den- und verkaufsorientierte“ Funktionsstruktur bräuchten meiner Ansicht nach Architekten, Bauherren und Centermanager schon zu Beginn der Planungsphase:

- *Architekten und Gebäudeplaner* bräuchten Vorgaben für eine Anordnung der Einzelhandelsflächen, bevor sie mit den ersten Entwürfen starten. Auf dieser Basis könnten sie dann die Gebäudestruktur und Funktionsplanung für das konkrete Gebäude entwickeln. Schließlich haben sie auch noch ganz andere Probleme bei der Grundrissplanung zu lösen, wie die Statik, die Haustechnik (Elektrik, Energie, Zu- und Abwasser etc.), den Brandschutz, ästhetische Anmutung und die Kosten.
- *Bauherren und Investitionsentscheider* bräuchten objektive, empirische Kriterien für die Beurteilung der verschiedenen Grundrissentwürfe und Planungsvarianten. Fehlentscheidungen schlagen sich in Euro und Cent nieder.
- *Centermanager* bräuchten objektive Leitlinien für eine ertragsorientierte Branchenbelegung der Flächen. Sie sind in hohem Maße an langjährigen Mietern interessiert, denen es wirtschaftlich gut geht. Häufiger Wechsel von Geschäften bedeutet nicht nur Umsatzausfall für die Einzelfläche, sondern schädigt auch die anderen Geschäfte und so das Image des ganzen Einkaufszentrums.¹

In zwei Studien wird das Problem der nutzerorientierten, erwartungsgerechten Anordnung von Funktionsflächen in der Architektur untersucht. Eine Studie beschäftigt sich mit der Anordnung von Flughafeneinrichtungen, die andere mit der Anordnung von Geschäften in einem Einkaufszentrum. Der Einfachheit halber wird meistens von NutzerInnen als der Kurzform von Nutzerinnen und Nutzer gemeint. Doch da diese Kurzform nur daran erinnern soll, dass Frauen und Männer sich unterscheiden ansonsten aber künstlich ist, wird sie auch nicht durchgängig angewendet.

Das methodologische Problem

Um nutzerorientierte Anordnungen zu entwickeln, muss vorher das methodologische Vorgehen geklärt werden. Lösungsansatz und Lösungsweg sollen nachvollziehbar, überprüfbar, zuverlässig, gültig sein, damit sie für andere Anwendungsfälle wiederhol-

¹ U.a. Aussagen von Frau S. Perzl, Geschäftsführerin der Gewerbepark Vermietungs GmbH Regensburg

bar und übertragbar sind (vgl. Herrmann 1993); z.B. auf andere Geschäfte, Sortimente, Gebäude etc. und sogar auf andere Bereiche (z.B. Bedienoberflächen, Steuerungspulte oder Cockpits).

Unterschiedliche Annahmen über die Ursachen und Wirkungszusammenhänge haben unterschiedliche Auswirkungen auf die Wahl der Methoden und damit auf die Maßnahmen. Folgende Überlegungen am Beispiel „Artikelanordnung im Supermarkt“ sollen dies belegen: Man könnte z.B. der Überzeugung sein, dass ein gutes Zurechtfinden eine Frage der Gewöhnung und des Lernens ist, und dass die KundInnen die Anordnung nur gut genug kennen müssen. So eine Überzeugung würde dazu führen, zuerst einmal zu erheben, was die KundInnen bereits kennen. Man könnte dann Lehrmaßnahmen entwickeln und die KundInnen entsprechend multimedial (Infozettel, Videofilme, Führungen etc.) belehren. Man könnte versuchen sicherzustellen, dass sich möglichst an der Anordnung nichts mehr ändert, damit sie sich auch wirklich daran gewöhnen können. Vielleicht würde man auch Normierungen anstreben, damit Mitbewerber sich auch an diese Anordnung halten müssen.

Man könnte aber auch der Überzeugung sein, dass die Anordnung Bewegungs- und Handlungsgesichtspunkten folgen sollte: Große, sperrige Ware am Beginn, damit sie als Erstes in den Einkaufskorb kann. Alle leichten Artikel in die Reckzone (d.h. auf Regalböden höher als 150 cm) und die schweren in die Bückzone (d.h. auf Regalböden niedriger als 80 cm). Oder aber man setzt auf motivationale Faktoren: Die Ware, die Spaß macht (wie Kosmetik, Geschirr, Süßwaren) an den Anfang und die Muss-Ware (Mehl, Zucker, Toilettenpapier) möglichst in den hinteren Raumbereich und vor die Kasse Artikel für den spontanen Kauf (Kaugummi, kleine Süßigkeiten etc.). Zu denken ist auch an ein Leitsystem mit großen Hinweistafeln über jeder Gruppe. Möglich ist auch eine Kombination all dieser Annahmen und Maßnahmen. Oder aber man verzichtet auf solche theoriegeleiteten Annahmen und setzt auf Befragungen von KundInnen und NutzerInnen. Man führt Tiefeninterviews oder Focus-Gruppen durch um per Befragung zu erfahren, welche Anordnung sich die NutzerInnen wünschen! Empfehlungen, wie man die Such- und Findeleistung auf einer Verkaufsfläche, in einem Flughafen o.Ä. optimieren kann, haben Konsequenzen. Der Auftraggeber muss nicht nur für die Arbeit des Psychologen zahlen. Er riskiert auch Kosten und Vertrauensverluste bei seinen MitarbeiterInnen und KundInnen. Die KundInnen finden sich schlechtestenfalls noch weniger

zurecht als vorher. Es entsteht ein Schaden für das Unternehmen und die betroffenen Personen.

Welcher Ansatzpunkt, welche Theorie, welche Methoden sind zu wählen? Wodurch lassen sich Verbesserungen herstellen? Als praktisch arbeitende PsychologIn oder ErgonomIn wird man laufend vor Entscheidungen gestellt. Er/sie geht dann entweder naiv vor und erhofft sich die Lösung aus dem Mund des Nutzers. Oder er „...berücksichtigt allgemeine Theorien, deren Auswahl und Kombination ihm selber überlassen bleiben. Er findet methodische Lösungen.... Er wählt aus den allgemein akzeptieren Routinen diejenigen aus, die nach seiner Überzeugung wirksam sind...“ (Witte 1991, S. 47). Bei all diesen Alternativen sollte der angewandte Psychologe entscheiden, welche Lösung die zuverlässigste und effektivste ist und welche Kosten sie produziert. Diese Entscheidung kann er aber nur rational lösen, „wenn er zurückgreift auf die wissenschaftlichen Grundlagen seines Faches.“ (Witte 1991 S. 55). Insbesondere der schnelle Rückgriff auf Routinen „... bedeutet zwar eine erhebliche Entlastung, bei dem häufig vorhandenen Zeitdruck aber auch eine Stagnation der Praxis. Man entschließt sich, ein Repertoire von Routinen zu erlernen und anzuwenden, ohne weiter die Wirksamkeit zu kontrollieren. ... Spätestens an dieser Stelle muss der wissenschaftlich ausgebildete Praktiker aus ethischen Gesichtspunkten ein schlechtes Gewissen bekommen, weil unklar bleibt, ob nicht andere Routinen größeren Nutzen (weniger Schaden), größere Zufriedenheit und größere methodische Anerkennung bringen.“ (Witte 1991, S. 56).

Als PraktikerIn hat man nicht die Wahl methodologisch zu arbeiten oder nicht. Man wird nolens volens Methodologie-Entscheidungen treffen und gegebenenfalls *methodologicien malgre lui* sein. Wenn man sich aber mit einer hausgemachten Methodologie begnügt, dann wird man in dieser Hinsicht nicht die bestmögliche Ressource gewählt haben. Methodologie-Regeln können, solange sie implizit bleiben, nicht kritisiert werden. Und damit besteht auch keine Chance, dass sie verbessert werden. Eine gute Methodologie ist für den Praktiker bzw. die Praktikerin eine gute Ressource. Sie ist unentbehrlich, ob man es will oder nicht, ob man es weiß oder nicht (vgl. Radnitzky 1992, Herrmann 1979, 1993, Greif 1993). Das folgende Kapitel 2 beschäftigt sich somit zunächst mit methodologischen Fragen.

2. Entwicklung des methodischen Vorgehens zur Lösung des praktischen Problems

Wie muss der Problemlösungsprozess zur Entwicklung nutzerorientierter Anordnungen gestaltet sein, damit er auch wissenschaftlichen Standards entspricht? Dies soll im Folgenden geklärt werden. *Im ersten* Abschnitt werden zunächst die Ursprünge, Ziele und Kontextbedingungen benutzerorientierter Forschung dargestellt. Dann folgt ein Überblick über methodologische Grundpositionen und es wird in der Literatur nach einem methodischen Vorgehen gesucht. Da hier keine zufrieden stellende, explizit formulierte Methodologie zu finden ist, wird im *zweiten* Abschnitt ein Ausflug in die Wissenschaftstheorie und die Forschungsmethodologie (Radnitzky 1994) unternommen, um dort Empfehlungen für eine Anwendungsmethodologie zu erhalten. Im *dritten* Abschnitt wird dann ein methodisches Vorgehen für den Problemlösungsprozess entwickelt. Im *vierten* Abschnitt erfolgt ein Ausflug in die Konstruktionsmethodik, einem Fachbereich des Maschinenbaus, um zu zeigen, dass hier ein ähnlich methodisches Vorgehen zu finden ist. Der *fünfte* Abschnitt gibt abschließend eine Zusammenfassung.

2.1 Aktuelles methodisches Vorgehen in der Nutzerforschung

Wie ist das heute übliche, methodische Vorgehen zur Gestaltung nutzerorientierter Lösungen? Um hierauf eine Antwort zu geben, werden ISO-Normen, aktuelle, ausgewählte Literatur der Nutzerforschung sowie Lehrbücher der Ergonomie daraufhin analysiert, welches Vorgehen sie nahe legen, um nutzerorientierte Gestaltungslösungen zu entwickeln. Zur Einführung wird kurz dargestellt, welche Ursprünge die Nutzerforschung hat und in welchem Kontext sie arbeitet.

2.1.1 Ursprünge, Ziele und Kontext nutzerorientierter Maßnahmen

Ursprünge Der Aspekt der Nutzer- bzw. Bedienerfreundlichkeit fand seinen Ausgangspunkt in Bedienungsproblemen der Nutzer mit computergestützten, interaktiven Systemen (Shakel 1957, 1991). In der Literatur ist er unter verschiedenen Bezeichnungen bekannt geworden, wie Human-Computer-Interaction (HCI), Man-Machine-Interaction

(MMI), Software-Ergonomie, User-Interface-Design (UID), Human Factor (HF), User Centered Design (UCD), Usability Engineering, User Research. Die ersten Ansätze beziehen sich ausschließlich auf Nutzerprobleme mit computergestützten Systemen, die Letzteren zeigen bereits in der Bezeichnung den breiteren Ansatz, sodass unter Nutzerforschung auch allgemein Aktivitäten für die Gestaltung von Konsum- und Investitionsgütern, Arbeitsplätzen, Gebäuden o.a. zu verstehen sind. Inzwischen hat sich Nutzerforschung als ein eigener Fachbereich innerhalb der Ergonomie – als der übergeordneten Wissenschaft zur Anpassung der Umwelt an den Menschen – etabliert. Bedienerfreundlichkeit erweist sich für Unternehmen nicht nur als Schlagwort, sondern als monetärer Marketing- und Vertriebsvorteil (vgl. Nielsen 1993, Green & Jordan 1999, Mayhew 1999, Shneiderman 2002). Heute sind an Universitäten eigene Lehrinstitute für Nutzerforschung zu finden. Konzerne wie Nokia, Siemens, Renault, Daimler/Chrysler etc. betreiben eigene Usability Laboratories (User-Labs). Und es gibt einige, wenige selbständige Dienstleistungsinstitute.

Ziele der Nutzerforschung Heutzutage werden drei Ziele verfolgt: eine nutzerorientierte Auswahl der Funktionen (User Centered Functionality), eine bedienerfreundliche Gestaltung der Funktionen (Usability) und ein nutzerorientierter Bedienspaß (vgl. Abb. 2.2) (Carroll & Thomas 1988, Carroll 1991, Shakel & Richardson 1991, Nielsen 1993, Green & Jordan 1999).

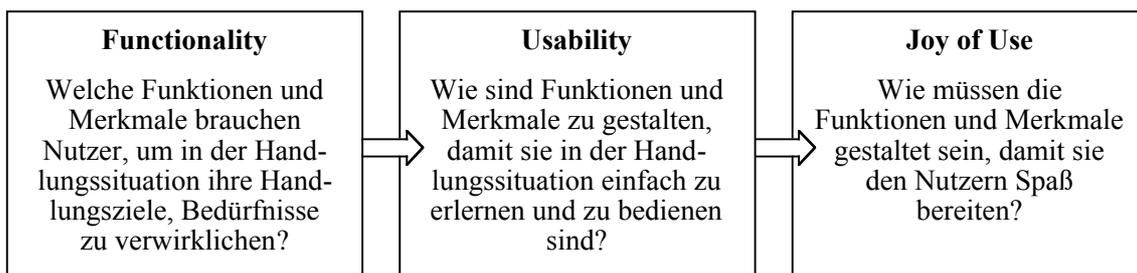


Abbildung 2.2. Ziele der aktuellen Nutzerforschung für die Gestaltung von Produkten

Was heißt dies? In puncto *Usability* sollen die Gestalter (z. B. Designer, Konstrukteure, Architekten) darin unterstützt werden, die Funktion bedienerfreundlich zu realisieren im Sinne von selbsterklärend, erwartungskonform, flexibel, einfach zu erlernen und zu bedienen, etc. (vgl. Shakel 1991, Nielsen 1993, DIN EN ISO 9241-11 1998, DIN EN ISO

13407, 1999). Dies Ziel stand am Anfang. Dann rückte der *Joy of Use* in den Fokus. Bereits 1988 wiesen Carroll und Thomas darauf hin, dass es nicht ausreicht bei der Gestaltung nur auf Bedienerfreundlichkeit zu achten, sondern dass die NutzerInnen auch Produkte brauchen, die z.B. angenehme Wahrnehmungseindrücke (z.B. sehen, tasten, riechen, hören) und Gefühle auslösen, die Spaß oder Anerkennung in sozialer Interaktion (z.B. durch Symbole zur Rollen- und Gruppenidentifikation) bewirken oder die Sinn vermitteln, weil sie Werte verkörpern (z.B. Ökologie, Technikbegeisterung, Minimalismus, Ästhetik etc.) (vgl. Carroll & Thomas 1988, Jordan 1999, Burmester, Hassenzahl & Koller 2002). Ein m. E. in der Nutzerforschung noch vernachlässigter Punkt ist die *User Centered Functionality*. Hier soll sichergestellt werden, dass das zu gestaltende Produkt (egal ob Konsumgut, Soft- oder Hardware-System, Gebäude etc.) auch wirklich diejenigen Funktionen hat, die die Nutzer zur Erreichung ihrer Handlungsziele in der konkreten Handlungssituation benötigen bzw. wünschen. Er ist sehr wichtig, vor allem, wenn es sich um die Entwicklung von Funktionen handelt, die die NutzerInnen aufgrund einer technischen Neuerung noch gar nicht kennen können.

Kontext Seit Beginn steht die Frage im Vordergrund, wie nutzerorientierte Ziele im Gestaltungsprozess umgesetzt werden können (vgl. Carroll 1991, Eason & Harker 1991). Daher ist es wichtig, diesen Prozess näher zu betrachten. Als Produktentwicklungs- oder Gestaltungsprozess wird der Ablauf verstanden, in dem die Teilnehmer des Entwicklungsprozesses (z.B. aus den Bereichen Marketing, Entwicklung, Design und Vertrieb) das neue Produkt entwickeln. Jedes Unternehmen, jede Fachdisziplin (Maschinenbau, Informatik, Design, Architektur etc.) verfügt über eine eigene Terminologie und Systematik. Doch um den Ablauf zu beschreiben, eignen sich folgende vier Phasen: die Orientierungs-, die Konzept-, die Realisierungs- und die Marktphase (vgl. Abb. 2.3). In der *Orientierungsphase* wird der Anforderungskatalog bzw. Briefing formuliert. Hier wird erarbeitet, für welche Nutzergruppe und welche Nutzungsziele das Produkt sein soll, welche Funktionen es haben soll und es werden allgemeine Anforderungen aus dem Bereich der Bedienung, Technik, Material etc. formuliert. In der *Konzeptphase* werden erste Vorentwürfe und Konzepte entwickelt. In der *Realisierungsphase* entstehen dann Konkretisierungen (z.B. CAD-Visualisierung, Mock-ups, Prototypen etc.). In der *Marktphase* finden dann schließlich die Markteinführung und Marktbeobachtung statt. Abhängig von der Produktentwicklungsphase unterscheiden sich die Beiträge, die Nut-

zerforscher leisten (sollten): In der Orientierungsphase benötigen die Gestalter Vorgaben (Input) darüber, welche Funktionen das zu gestaltende Produkt haben soll, wie diese nutzerorientiert zu realisieren sind und was schließlich zu tun ist, damit das Produkt Freude bereitet. Die Notwendigkeit solche Vorgaben bereits zu Beginn des Entwicklungsprozesses bereitzustellen, wurde allerdings erst in den letzten Jahren erkannt. Vorher konzentrierte sich die praktische Arbeit von Nutzerforschern fast ausschließlich auf die Bewertung von Gestaltungsvarianten in der Konzept- und Realisierungsphase (Evaluations-Studien).

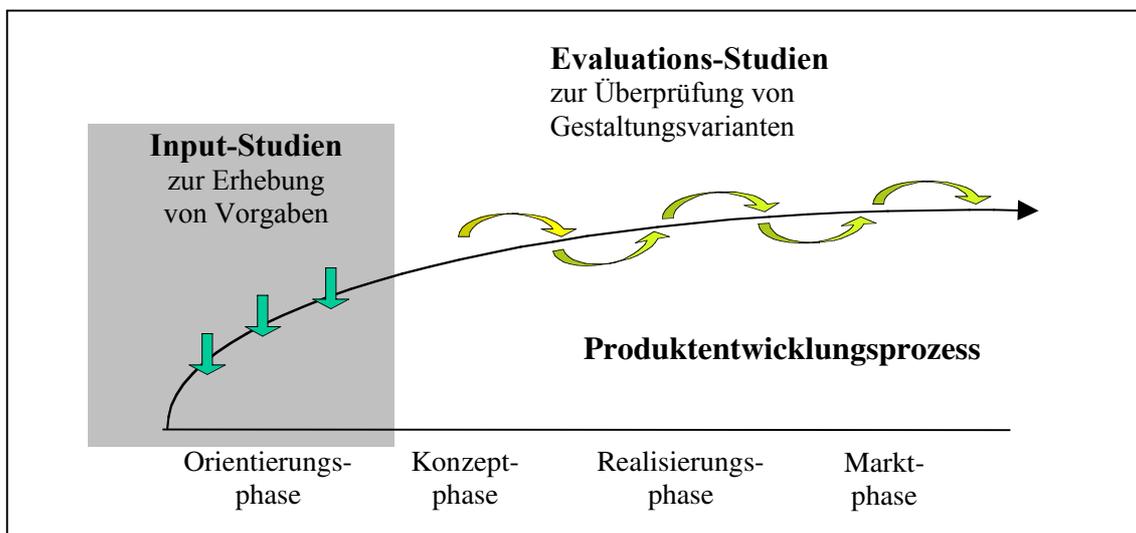


Abbildung 2.3. Input- und Evaluationsstudien mit NutzerInnen im Produktentwicklungsprozess

2.1.2 Methodologische Ansätze in der Nutzerforschung

Schon frühzeitig wurde diskutiert, wie methodisch vorzugehen sei, um im Entwicklungsprozess zur Entwicklung nutzerorientierter Gestaltungen beizutragen. Die Positionen spannen sich – vereinfacht zusammengefasst – zwischen einem deduktionistisch- und einem induktionistisch-orientiertem methodologischen Vorgehen auf (vgl. Kittle-House-Manifesto dargestellt in Carroll 1991).

Das *deduktionistisch-orientierte Vorgehen* geht von der klassischen Überlegung aus, experimentalpsychologische Methoden *und* Theorien auf das Anwendungsproblem anzuwenden und zu lösen. Dieses Vorgehen wurde 1920 von Münsterberg, einem Schüler von Wundt, formuliert, der als Pionier der angewandten Psychologie gilt (Münsterberg

nannte es Psychotechnik). Münsterberg proklamierte die Nutzung der Theorien und Methoden der experimentellen Psychologie für die Erreichung praktischer Ziele und Entdeckung. Kritisiert wird an diesem Ansatz, dass mit einer Anwendung von Theorien und Methoden noch lange keine gute Problemlösung entsteht. Experimente würden unter kontrollierten Bedingung ablaufen. Die Zahl der Variablen sei stark eingeschränkt. Praktische Probleme hingegen sind in der Regel hoch komplex. In Experimenten falsifizierte Gesetze würden der Komplexität der Problemstellung nicht gerecht. Die Anwendung von Theorien und Gesetzen auf Probleme bedürfe eines eigenen Vorgehens, einer Anwendungsmethodologie. So wurde dieses Vorgehen auch schnell als naiver Deduktionismus abgeurteilt (vgl. Herrmann 1993).

Die radikal gegenteilige Position stellt ein *induktionistisch-orientiertes Vorgehen* dar. Sie wird z.B. von Jack Carroll (1991) vertreten: nach seiner Auffassung sei das klassische deduktive Vorgehen beim tatsächlichen Design von Softwarewerkzeugen praktisch erfolglos geblieben. Es gäbe keine innovative Designlösung, die eindeutig als Leistung einer theoretischen Perspektive zurückgeführt werden kann. Römische Baumeister hätten auch bereits zu einer Zeit stabile Brücken gebaut, in der die entsprechenden physikalischen Gesetze noch nicht explizit vorgelegen hätten (Carroll 1991, S. 2). Er legt vielmehr nahe, vorhandene bedienfreundliche Designlösungen daraufhin zu untersuchen, warum sie funktionieren, nachträglich die theoretischen Erkenntnisse zu extrahieren, um diese dann mit wissenschaftlichen Methoden auf ihre Gültigkeit zu überprüfen. Carroll bestreitet nicht, dass Software-EntwicklerInnen bereits nach theoretischen Überlegungen – sei es Alltagswissen oder wissenschaftlichen Erkenntnissen – arbeiten. Aber er schlägt eine nachträgliche Extraktion der theoretischen Erkenntnisse aus dem praktisch erfolgreichen Artefakt und nicht direkt aus den zugrundeliegenden Theorien vor. Siegfried Greif (1993 S. 152 ff) kritisiert dieses Vorgehen mit dem Argument des Zirkelschlusses und verweist auf das Induktionsproblem: Aus einer erfolgreichen Gestaltungslösung lasse sich nicht schlüssig eine Theorie ableiten. Bereits die Beobachtungssprache, mit der die erfolgsbestimmenden Variablen beschrieben würden, enthalte theoretische Voraussetzungen, die expliziert werden müssten. „Wenn er nur auf einen methodologisch angreifbaren induktiven Weg zur Entwicklung einer zukünftigen Theorie verweist, ohne selbst eine überprüfbare Theorie vorzulegen, können wir ihm nicht folgen.“ (Greiff 1993, S. 155).

Zwischen diesen beiden gegensätzlichen Positionen finden sich Vorgehensweisen, die beides zu verbinden versuchen. So gibt es für Greif kein einfaches deduktives oder induktives Verhältnis zwischen Theorie und Praxis. „Das Spannungsverhältnis zwischen Theorie und Praxis wird als dialektisches Wechselverhältnis in komplexen Problemlöseprozessen interpretiert.“ (1991 S. 159.) Landauer (1985) vertritt mit seinem Anspruch, die Psychologie sei die „Mother of Invention“, eine eher deduktionistische Position. Herrmann (1979, 1993) fordert neben einer Grundlagenforschung eine deutlich unterschiedene Anwendungsforschung. Theorien und Methoden der Grundlagenforschung seien nicht direkt auf komplexe Praxisprobleme anwendbar. Er weist Grundlagentheorien den Charakter von „Mittelbausteinen“ zu (Herrmann 1979 S. 133 ff). Bei einer Problemlösung gingen diese sozusagen vermittelt in die Problemlösung ein. Es bedürfe einer eigenen Anwendungsforschung (Psychotechnologie).

2.1.3 Analyse von ISO-Normen und aktueller Literatur

In den vergangenen Jahren wurden internationale ISO-Normen zur benutzerorientierten Gestaltung entwickelt. Zwei dieser Normen werden im Folgenden daraufhin untersucht, welches methodische Vorgehen sie nahe legen. Sie werden kritisiert und es wird dazu neue Literatur dargestellt. Als Erstes wird die ISO 13407 (1999) zur „benutzerorientierten Gestaltung interaktiver Systeme“ dargestellt und kritisiert. Sie zielt auf den Prozess und beschreibt, in welchen Schritten benutzerorientierte Lösungen zu entwickeln sind. Als Zweites wird die ISO 9241-10 (1998) zur „Gebrauchstauglichkeit“ analysiert. Sie beschreibt, welche inhaltlichen Gestaltungsprinzipien einzuhalten sind, um benutzerorientierte Lösungen zu entwickeln. Diese Normen werden aus folgenden Gründen gewählt: 1) ISO-Normen sind Übereinkünfte von internationalen Experten (Nutzerforscher, Ergonomen etc.). Sie haben eine hohe Verbindlichkeit für Praktiker und zeigen, wie man aktuell vorgeht (bzw. vorgehen sollte); 2) sie sind relativ aktuell; 3) sie beziehen sich auf die Entwicklung benutzerorientierter Hardware *und* Software interaktiver Systeme; 4) computergestützte, interaktive Systeme sind Vorreiter in der Erforschung der Benutzerorientiertheit (Shakel 1957, 1991), auch für nicht computergestützte Systeme.

Analyse der Empfehlungen zum Gestaltungsprozess Im Folgenden wird die ISO 13407 dargestellt und kritisiert. Dann wird sie mit der aktuellen Literatur abgeglichen.

Darstellung der DIN EN ISO 13407 Entsprechend dieser Norm zeichnet sich ein benutzerorientierter Gestaltungsprozess aus durch a) die aktive Beteiligung der Benutzer sowie durch ein klares Verständnis der Benutzerbelange und der Erfordernisse der Aufgaben; b) eine geeignete Aufteilung der Funktionen zwischen Benutzer und angewandter Technik; c) eine iterative Bearbeitung, d.h. durch einen Wechsel von Gestaltungsarbeit und einer Einbeziehung von Benutzern; d) die Arbeit in multidisziplinären Teams. Die ISO-Norm schreibt vier Phasen und Gestaltungsmaßnahmen vor (siehe Abb. 2.1).

1. *Phase: Verstehen und Spezifizieren des Nutzungskontextes* In dieser Phase soll eine Nutzeranalyse (z.B. Kenntnisse, Fertigkeiten, Übungsgrad, körperliche Eigenschaften, Neigungen etc.), eine Aufgabenanalyse (z.B. Art, Häufigkeit, Dauer) und eine Situationsanalyse durchgeführt werden.
2. *Phase: Spezifizieren der benutzerorientierten Anforderungskriterien* Dann sollen die Anforderungen, die sich aus den Nutzerbelangen, den Aufgaben und den Erfordernissen der Situation ergeben, spezifiziert, bewertet und in messbare Kriterien übersetzt werden.
3. *Phase: Entwerfen von Gestaltungslösungen und Überprüfungen an Benutzern* Hier sollen Gestaltungslösungen – unter Ausnutzung der Ergebnisse der ersten Phase, des vorhandenen Standes der Technik und der Verwendung der vorhandenen wissenschaftlichen Erkenntnisse – in Form von Skizzen, Simulationen, Modellen bzw. Prototypen entwickelt werden. Diese Gestaltungsentwürfe sind zukünftigen Nutzern vorzustellen und die Nutzer sollen daran (simulierte) Aufgaben ausführen.
4. *Phase: Bewerten der Lösungen nach benutzerorientierten Kriterien* Schließlich sollen die Gestaltungslösungen anhand der Anforderungskriterien und der Erkenntnisse über das Benutzerverhalten bewertet werden. Die Erkenntnisse über die Reaktionen der Benutzer sowie die Bewertungen sind dem Gestaltungsteam rückzukoppeln.

Eventuell bedarf es einer Neubearbeitung der Gestaltungslösung (Phase 3) mit einer erneuten Bewertung (Phase 4), so lange, bis das Produkt die Anforderung erfüllt. Der ma-

ximale Nutzeffekt sei dann zu erreichen, wenn mit den Benutzern mehrere Wiederholungsmessungen durchgeführt werden.

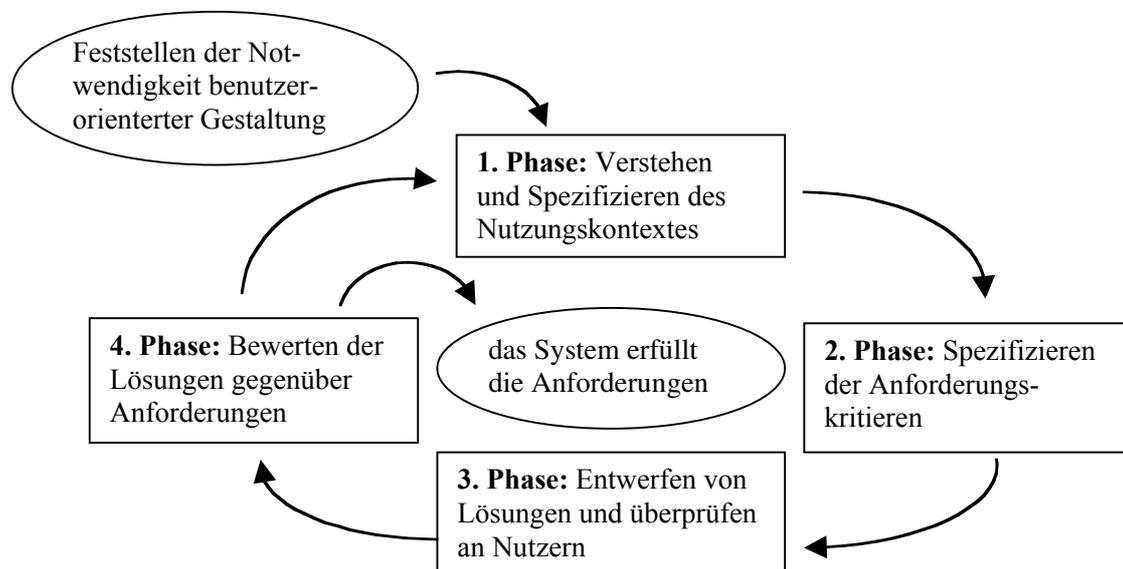


Abbildung 2.1. Phasen benutzerorientierter Gestaltungsmaßnahmen (zitiert nach ISO 13407, 1999; der Text wurde aus Verständlichkeitsgründen leicht modifiziert)

Kritik der DIN EN ISO-Norm 13407 Diese Norm ist wichtig, weil sie sich pragmatisch am Entwicklungsprozess orientiert. Sie fordert zu Beginn Untersuchungen für nutzerfundierte Vorgaben sowie Evaluations-Studien mit NutzerInnen zur Überprüfung der Gestaltungslösungen. Untersucht man die Norm hinsichtlich einer wissenschaftlichen Methodologie, so lassen sich jedoch drei wesentliche Kritikpunkte formulieren:

Erstens: Das Vorgehen zur Entwicklung von Vorgaben ist nicht theoriefundiert! In Phase 1 wird eine Nutzer-, Aufgaben- und Kontextanalyse nahe gelegt, mit dem Ziel hieraus Anforderungskriterien für eine nutzerorientierte Gestaltung zu entwickeln. Dieses Vorgehen ist pragmatisch, aber auch naiv induktionistisch! Allein aufgrund detaillierter, phänomenologischer Beschreibungen ist es noch nicht möglich Anforderungen zu entwickeln, die darlegen, wie etwas zu gestalten ist, damit die Nutzer das gewünschte Verhalten zeigen. Dies kommt einem Zirkelschluss nahe. Wäre dies möglich, bräuchte es keine Forschung und Suche nach Gesetzen und Wirkfaktoren.

Außerdem ist eigentlich allgemein bekannt, dass bereits jede Untersuchung und Beobachtung theoretische Inhalte impliziert, schlechtestenfalls unreflektierte und nicht auf

Eignung überprüfte Annahmen. Diese Annahmen müssten daher geklärt werden und durch wissenschaftliche, geeignete Theorien ersetzt werden. Wird z.B. die bisherige Erfahrung der Nutzer analysiert, dann impliziert dies ja die Annahme, dass die Lernerfahrung ein Merkmal ist, das Einfluss auf die Benutzerfreundlichkeit hat. Ohne Kenntnis von Theorien und Gesetzen z.B. über Wissenserwerb und Gedächtnis bleiben Fragen offen: z.B. Welche Art des Vorwissens (episodisches, begrifflich-abstraktes, handlungsbezogenes, implizites, explizites Vorwissen?) sollte erhoben werden? Welche Methoden (z.B. Fragebogen, nondirektives Interview, Experimente mit recall- oder recognition-Verfahren) sind dafür geeignet? Und es bleibt schließlich unklar, wie das Verhalten zu interpretieren ist und wie die Erkenntnisse im Gestaltungsprozess zu berücksichtigen sind.

In der aktuellen Literatur finden sich für die Entwicklung von Anforderungen und Vorgaben z.B. bei Deborah Mayhew (1999) und Ben Shneiderman (2002) die gleichen Vorgehensweisen, wie sie in der ISO-Norm 13407 zu finden sind: Es ist eine Nutzer-, eine Aufgaben- und eine Kontextanalyse durchzuführen. Sie liefern einen umfangreichen Methodenkatalog aus standardisierten und nicht-standardisierten Befragungen mittels Interview, Fragebogen oder Gruppendiskussionen (Focus-Gruppen), teilnehmenden oder verdeckten Beobachtungen, stakeholder meetings, Aufgabenszenarios, Tätigkeitsanalysen usw. (vgl. auch Green & Jordan 1999, Mayhew 1999, Shneiderman 2002).

Mayhew (1999) und Shneiderman (2002) schlagen in dieser Phase zusätzlich einen Schritt vor: der Problemstellung sollen passende Design-Prinzipien zugeordnet werden. Bei der Beschreibung der Prinzipienauswahl beschränkt sich Mayhew auf praktische Literaturhinweise mit Sammlungen für Designrichtlinien, welche Arten von Design-Guidelines es gibt, ihre Wichtigkeit und die Frage, wie man sie mit der Nutzer-, Aufgaben- und Kontextanalyse verbindet, um Anforderungen zu entwickeln. Shneiderman nennt acht goldene Gestaltungsregeln. Wie die Auswahl inhaltlich zu gestalten ist, thematisieren weder Mayhew noch Shneiderman. Ein spezifiziertes Vorgehen, wie Design-Richtlinien anzuwenden sind, wird auch nicht gegeben.

Zweitens: Die Wahl der Methoden für die Evaluation ist nicht theoriegeleitet. Die Darstellungen in Phase 3 behandeln zwar ausführlich praktische Aspekte der Überprüfung: z.B. in welcher Form die Gestaltungslösungen den Nutzern dargeboten werden sollten

(Skizzen, Prototypen etc.), dass festgelegt werden muss, welche Teile des Systems mit welchen Verfahren zu bewerten sind etc. Es wird aber nicht erwähnt, dass Erhebungs- und Auswertungsmethoden selbst Theorieinhalte implizieren. Die Wahl der Methoden sollte – bei einem wissenschaftlichen Vorgehen – auch durch die Untersuchungsfrage sowie die zugrunde liegende Theorie definiert werden, um ein in sich stimmiges Untersuchungsdesign zu entwickeln (vgl. Bortz 1985, Sarris 1990). Dies wird aber nicht thematisiert.

Auch Nielsen (1993), Mayhew (1999) und Shneiderman (2002) geben praktische Tipps, wie solche Usability Tests durchzuführen sind, dass z.B. die Nutzer für das Programm repräsentative Testaufgaben erhalten müssen. Sie liefern eine lange Liste an quantitativen Verfahren und thematisieren die Notwendigkeit reliabler und valider Methoden. Und sie nennen möglich messbare Verhaltensmaße, wie z.B. die Bediendauer, Anzahl der Fehler, die Häufigkeit, mit der der Nutzer seine Frustration äußert, die Dauer der „toten“ Zeit, bei der der Nutzer inaktiv ist, etc. Die Wahl der Methoden orientiert sich m.E. mehr an der Frage der Reliabilität und Validität als an methodologischen Kriterien (vgl. Eason & Harker 1991, Tetzlaff & Mack 1991, Kanis 1999). Auffallend ist z.B., dass für eine Interpretation der Fehler Techniken, wie „lautes Denken“ empfohlen werden. Der Bedeutung quantitativer Methoden innerhalb der Nutzerforschung liegen wahrscheinlich ähnliche Motive zugrunde, wie sie Bungard (1993 S. 378) für die gesamte Arbeits- und Organisationspsychologie unterstellt: eine zentrale Rolle nimmt hier die Frage nach der Adäquatheit (Objektivität, Reliabilität, Validität, Eichung, etc.) der Methoden ein, weil dies anscheinend einen wissenschaftlichen Status garantiert.

Drittens: Phänomenologische Fehlerbeschreibungen liefern noch keine Veränderungsempfehlungen! Es wird in Phase 3 eine nicht näher spezifizierte Aufgabensimulation mit Fehleranalyse empfohlen. Den Nutzern werden z.B. Aufgaben mit der Gestaltungslösung gegeben und es wird erhoben, welche Fehler sie dabei machen. Quantitative Methoden liefern aber nur quantitative Daten, z.B. über Reaktionszeiten und Fehleranzahl. Es wird nicht erläutert, in welcher Form diese Daten zu interpretieren sind.

Kanis (1999) kritisiert an diesem in der Nutzerforschung üblichen Vorgehen, dass Fehler auf summativen Effekten basieren. Quantitative Daten, wie Reaktionszeiten und Fehleranzahl gäben keine Empfehlungen für Designveränderungen. Er schlägt daher qualitative Methoden, wie Tiefeninterviews, Handlungsanalysen vor, weil sie brauchba-

rere Erkenntnisse für Designverbesserung liefern (vgl. Kanis 1999, 2000). Doch dies reicht m.E. nicht aus. Wenn man Fehler als Ausgangspunkt für eine Verbesserung der Gestaltung benützen will, müssten der Fehlerbeschreibung kausale bzw. aktionale Fehlermodelle (vgl. Bubb 1992) zugrunde gelegt werden. Eine rein phänomenologische bzw. auftretensorientierte Fehlerbeschreibung gibt es a) wohl kaum (vgl. Frese & Zapf 1991) und würde b) auch nicht ausreichen. Denn um Gestaltungsveränderungen zu empfehlen, bedarf es ursachenorientierter Klassifizierungsansätze. Und diese verlangen eindeutig die Kenntnis zugrunde liegender Wirkzusammenhänge. Es wäre also notwendig zu wissen, welche Gesetzmäßigkeiten des Verhaltens den Fehlern zugrunde liegen bzw. welche Theorie sich hinsichtlich der Aufgabenstellung am besten eignet. Die DIN EN ISO 13407 thematisiert dies nicht einmal.

Analyse der Empfehlungen zu Design-Richtlinien Im Folgenden wird die ISO 9241-10 analysiert, die Design-Richtlinien zur benutzerorientierten Gestaltung liefert. Nach der Darstellung erfolgt die Kritik mit dem Abgleich zur aktuellen Literatur.

Darstellung und Kritik der DIN EN ISO-Norm 9241-10 Die DIN EN ISO-Norm 9241 „Ergonomic requirements for office work with visual display terminals“ (1998) liefert inhaltliche Vorgaben „für Gebrauchstauglichkeit“ von Software und interaktiven Geräten. Sie fordert z.B. für die Gestaltung von „benutzerfreundlichen“ Dialog-Systemen die Realisierung von folgenden sieben Prinzipien: *Aufgabenangemessenheit, Selbsterklärbarkeit, Erwartungskonformität, Steuerbarkeit, Fehlerrobustheit, Individualisierbarkeit und Lernförderlichkeit*. Das Prinzip *Erwartungskonformität* wird im Folgenden untersucht:

Unter dem Prinzip der *Erwartungskonformität* werden Lösungen bezeichnet, die hinsichtlich der verwendeten Begriffe und Bedienweisen dem Wissen und den Erwartungen der Benutzer entsprechen und konsistent sind. Es werden sieben Kriterien genannt, wie: „Es sollte ein Vokabular verwendet werden, das dem Nutzer im Bezug zur Aufgabenbewältigung geläufig ist.“ „Die Rückmeldungen sollten dem Wissen des Nutzers entsprechen.“ „Das Dialogverhalten und die Darstellungen sollten konsistent sein.“ Aber neben diesen Kriterien wird kein Erklärungswissen geben, wie diese angemessen auf ein spezifisches Problem anzuwenden sind oder wie ein Gestalter selbständig prüfen

kann, ob er sie eingehalten hat. In der Norm findet sich weder eine Darstellung noch ein Verweis auf Denk-, Gedächtnis- oder Begriffstheorien. Doch gerade das Verständnis solcher Theorien ist m.E. notwendig, um die Bedeutung von Erwartungen bzw. von Wissensinhalten zu verstehen und um z.B. entsprechend erwartungsgerechte Funktionen, Bedienschritte oder Bezeichnungen zu realisieren. Es wird nicht nahe gelegt, die NutzerInnen zu untersuchen, um z.B. zu erfahren, auf welchem Wissensstand sie sich befinden, welches Vokabular sie kennen etc. Wie aber soll dann der Entwickler verständliche, erwartungsgerechte Lösungen entwickeln, wenn er nicht weiß, wie Erwartungen gebildet werden, welches Vorwissen die NutzerInnen haben und welchen Einfluss das Vorwissen auf die Erwartungsbildung hat?

Innerhalb der Literatur wurden bereits frühzeitig eine Vielzahl von Design-Richtlinien und Prinzipien formuliert (vgl. Tetzlaff & Mack 1991). Di Sessa (1991) nannte abstrakt-formulierte Design-Prinzipien wie: „Eine Isomorphie zwischen dem Mentalen Modell und der Bedienoberfläche erleichtert das Verständnis von komplexen Problemen.“ oder „System-Einfachheit und strukturelle Integrität sind wichtige Aspekte für die Erlernbarkeit des Systems.“ Auch die Designprinzipien von Carroll und Kollegen (1991) sind noch abstrakt gehalten, wie „Prominentes Positionieren von zentralen Funktionen hebt angemessene Unterziele besser hervor.“ Nielsen (1993) nennt 10 „Usability Heuristics“, wie „Verwende einfache, natürliche Darstellungen“, „Verwende die Sprache des Nutzers“, „Halte die Gedächtnisbelastung des Nutzers so gering wie möglich“, „Achte auf Konsistenz“, „Gebe eindeutige Rückmeldungen“, „Ermögliche Abkürzungen“, „Gib klare Fehlermeldungen“, „Helfe Fehler zu verhindern“ usw. Er erläutert sie an einigen Beispielen, gibt aber keine eindeutigen, konkreten Anhaltspunkte, woran der Gestalter z.B. feststellen kann, dass er die Gedächtnisleistung gering gehalten hat oder dass er auf Konsistenz geachtet hat. Bezüge bzw. Verweise zu Grundlagentheorien, die ein Hintergrundwissen geben, macht er nicht.

Shneiderman (2002, S. 100-101) nennt acht goldene Regeln, die sich größtenteils mit denen von Nielsen decken. Er stellt sie auf zwei Seiten vor. Darstellungen, wie man diese Regeln auf ein Gestaltungsproblem anwendet, finden sich nicht. Bezüge bzw. Verweise zu psychologischen Grundlagentheorien und somit erklärendes Hintergrundwissen für die Gestalter finden sich nicht. Shneiderman schreibt noch einige Seiten zuvor (2002, S. 73): „Glücklicherweise bekommen Designer Anleitungen in dreierlei Form:

1. Anspruchsvolle Theorien und Modell auf einem hohen, übergeordneten Level; 2. Prinzipien auf einem mittleren Level und 3. spezifische und praktische Richtlinien“. Doch es finden sich in seinem Buch weder Theorien, die die acht Regeln begründen, noch Verweise auf sie.

Die Beispiele zeigen bereits, dass sich die Design-Prinzipien in ihrem Abstraktionsgrad, in der Art der Formulierung, der Tiefe der Elaboriertheit und dem Erklärungsanteil unterscheiden (vgl. Tetzlaff & Mack 1991, Mayhew 1999). Gemeinsam ist ihnen ihre pragmatische Stoßrichtung: die Anwendung im Designprozess. Doch für die konkrete Gestaltungsaufgabe bieten sie allenfalls ein Rezept – aber kein wirkliches Erklärungs- und Hintergrundwissen. Sie sind in der Regel nicht abgeleitet bzw. mit Theorien hinterlegt. Es gibt keine Hinweise wie diese Regeln konkret auf das Problem anzuwenden sind. Somit lassen diese Designprinzipien den Verdacht eines naiven Deduktionismus aufkommen. Doch auch die Gestalter bräuchten ein problemlösungsorientiertes Wissen. „DesignerInnen können nicht aus ihrer Verantwortung als ExpertInnen für ihr Anwendungsfeld entlassen werden.... So gesehen sind sie auch immer zugleich TheoretikerInnen, die mit ihrem Design in dialektischen Problemlöseprozessen zwischen praktischen Gestaltungszielen, Theorien, Gestaltungsprinzipien und Designlösungen vermitteln.“ (Greif 1993, S. 158).

Der einzige mir bekannte, systematische Versuch Design-Guidelines und Grundlagentheorien aus dem Bereich der Kognitiven Psychologie zu verbinden, stellt das Buch „Applying Cognitive Psychology to User-Interface-Design“ von Margaret Gardiner und Bruce Christie (1990) dar. Sie stellen kognitive Grundlagentheorien (z.B. zum Arbeitsgedächtnis, zum episodischen und semantischen Gedächtnis) vor und leiten daraus insgesamt 162 Design-Guidelines ab. So formulieren sie auf der Basis der Theorien zum Semantischen Gedächtnis folgende Empfehlungen „Verwende bei der Einführung neuer Konzepte Bekanntes, damit bestehende Wissenskörper verwendet werden.“ „Verwende bekanntes Material, Situationen, Arbeitsmethoden und bedeutsame Analogien, um ein gutes Nutzerverhalten zu bewirken.“ An einigen Beispielen wird gezeigt, wie einige der 162 Richtlinien konkret umgesetzt werden können. Gardiner und Christie liefern somit zwar eine theoretische Herleitung und Begründung für 162 Design-Richtlinien, aber sie liefern kein methodisches Vorgehen, wie eine praktische Problemstellung zu lösen ist:

wie die Problemstellung zu analysieren ist, wie die passende Richtlinie auszusuchen und dann anzuwenden ist. Und es wird gar nicht thematisiert, was zu tun ist, wenn eine Problemstellung durch die 162 aufgeführten Richtlinien nicht abgedeckt wird. So gibt es z.B. keine Richtlinie für die Entwicklung nutzerorientierter, erwartungsgerechter Anordnungen. Dieser Aspekt ist nicht vorgesehen.

2.1.4 Analyse von zwei ergonomischen Lehrbüchern

Bleibt noch der Blick in die Ergonomie, als der übergreifenden Wissenschaft zur Anpassung der Umwelt an den Menschen. Im Folgenden werden zwei Lehrbücher der (deutschen) Ergonomie daraufhin untersucht, welches methodische Vorgehen hier empfohlen wird, um praktische Probleme wissenschaftlich zu lösen. Es werden die beiden Standardwerke: „Ergonomie“ (Hrsg. Schmidtke 1993) und „Arbeitswissenschaften“ (Luczak 1993) analysiert.

Analyse des Lehrbuches „Ergonomie“ von Schmidtke (Hrsg. 1993). In den ersten vier Kapiteln werden ausgewählte, arbeitsgestaltungsbezogene Grundlagen der Medizin (Anatomie, Biomechanik, Neurophysiologie), der Psychologie (Wahrnehmung, Intelligenz, Lernen, Motivation), der Belastung und Beanspruchung (Physische und psychische Belastung, Vigilanz, Stress) und der physikalischen, chemischen Umweltbeeinflussung (Strahlung, Lärm, Schwingungen, Klima, Druck, chemische Stoffe) zusammenfassend dargestellt. Im fünften Kapitel wird das meta-theoretische Modell der Systemtheorie (Kybernetik) vorgestellt und es wird es an verschiedenen Beispielen gezeigt, dass es sich zur Analyse menschlicher (Arbeits-)Tätigkeiten eignet. In Kapitel sechs und sieben werden Grundlagen der Arbeitsgestaltung und Arbeitsstrukturierung an angewandten Beispielen dargestellt. Das Buch schließt mit einem Kapitel über normative Aspekte und eine Verankerung der Arbeitswissenschaften in Normen und Verordnungen.

Den vielen dargestellten, praktischen Beispielen liegt zwar ein eindeutiges methodisches Vorgehen zugrunde: Theorien und empirische Erkenntnisse werden auf das Problem angewendet. Aber es fehlt eine explizit dargestellte Methodologie, d.h. eine all-

gemeine Vorgehensweise zur Lösung von Problemen in der Ergonomie. Es wird nicht explizit dargestellt, wie im konkreten Fall vorzugehen ist, welche Schwierigkeiten sich dabei ergeben können, worauf zu achten ist usw. Somit wird dieses Vorgehen nicht offen zur Disposition gestellt und es wird nicht zur Anwendung des Vorgehens und zur Auseinandersetzung damit aufgefordert.

Analyse des Lehrbuches „Arbeitswissenschaften“ (Luczak 1993). Luczak startet sein Lehrbuch mit grundlegenden Ansätzen der Arbeitswissenschaften (wie die Systemergonomie, das Belastungs-Beanspruchungs-Konzept oder die Handlungsregulationstheorie). Er stellt Erhebungsmethoden der Arbeitsanalyse und grundlagenbezogene Theorien dar (wie zur Arbeitsperson, zur menschlichen Informationsverarbeitung, zur menschlichen Kraft- und Energieerzeugung etc.). Erst im fünften Kapitel thematisiert Luczak das Theorie-Praxis-Verhältnis der Arbeitswissenschaften. Er beschreibt folgenden Problemlösezyklus für arbeitswissenschaftliche Gestaltungsfragen: Ausgehend von einer „globalen“ Zielsetzung erfolgt eine Zerlegung (Analyse) des Gesamtproblems in Teilprobleme, „...bis die Komplexität der Einzelprobleme soweit reduziert ist, dass verfügbare Lösung herangezogen oder neue Lösungen gefunden werden können.“ (Luczak 1993, S. 350) „Die Teilprobleme finden dabei ihre Lösungen in der entsprechenden Aspektwissenschaft¹ wie der Soziologie, der Pädagogik, der Medizin oder der Psychologie. Die Einzellösungen werden dann zur einer Gesamtlösung zusammengefasst (Synthese).“ (vgl. Abb. 2.5).

Luczak schreibt, diesem zerlegenden Problemlösungsprozess stehe das Bedürfnis der Praktiker gegenüber, gefiltertes, handlungsbereites Problemlösungswissen und Routinen an die Hand zu bekommen, um praktische Aufgabenstellungen zu lösen und „...nicht mehr nach Begründungszusammenhängen zu fragen. ... Auf den Problemlöseprozess übertragen bedeutet dies, dass Probleme im Einzelfall nicht mehr auf eine Ebene herunter gebrochen werden, die eine wissenschaftlich begründete Lösung der Teilprobleme erlaubt, sondern durch Anwendung von Regeln dieser Prozess abgekürzt wird.“ (ebd. S.

¹ Jede Aspektwissenschaft (wie z.B. die Soziologie, die Pädagogik, die Medizin oder die Psychologie) betrachtet das Arbeitsgestaltungsproblem aus ihrem fachlich definierten Blickwinkel und löst es auf der Basis ihrer Theorien und Methoden.

351). Eine Zerlegung in Teilprobleme erfolgt nur in einem solchen Grade, dass bekannte Regeln und Verfahren angemessen angewandt werden können.

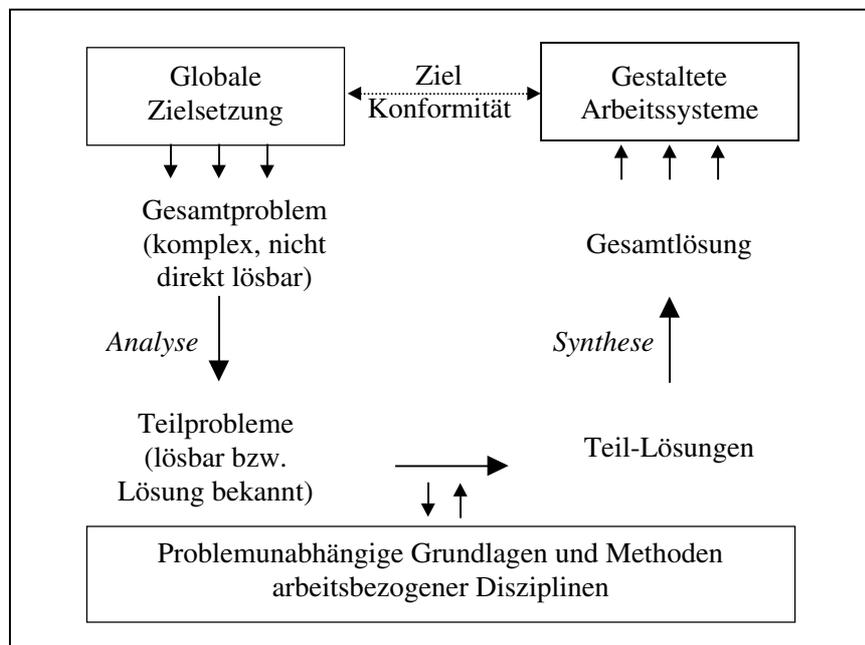


Abbildung 2.5. Problemlösezyklus in arbeitswissenschaftlichen Gestaltungsfragen (aus Luczak 1993 S. 350)

Luczak thematisiert hiermit das Verhältnis von Problemstellung, Theorie, Methoden und Problemlösung explizit. Er vertritt ein deduktives methodologisches Vorgehen. Aber die Darstellung bleibt nur allgemein. Er liefert kein methodisch differenziertes Vorgehen, er zeigt nicht explizit am Beispiel, wie eine Gesamtaufgabenstellung zu zerlegen ist, wie Teilprobleme herauszuarbeiten sind, wie man in der „richtigen“ Aspektwissenschaft nach „richtigen“ Theorien und Erkenntnissen sucht, die das Problem lösen können, und wie man dann die Lösungen eventuell mehrerer Teilprobleme wieder zusammenführt und das Gesamtproblem löst. Er stellt nicht dar, welche Schwierigkeiten sich ergeben können, worauf zu achten ist, welche Standards einzuhalten sind, usw.

2.1.5 Zusammenfassung und Konsequenzen

In den Ergonomie-Lehrbüchern findet sich eine nicht näher spezifizierte deduktionsorientierte Methodologie. In den ISO-Normen und der Literatur zur Nutzerorien-

tierten Gestaltungen wird ein iterativer Gestaltungsprozess mit einer Einbeziehung von Nutzern und die Verwendung von Design-Richtlinien empfohlen: in der Orientierungsphase werden empirische Studien zur Entwicklung von Designvorgaben empfohlen und in der Konzept- bzw. Realisierungsphase empirische Evaluations-Studien zur Überprüfung von Gestaltungsvarianten. Dies lässt zunächst eine Mischung von deduktionistischer und induktionistischer Vorgehensweise vermuten. Die Ausführungen für diese Studien sind sehr praxisbezogen und orientieren sich merklich an den praktischen Bedingungen des Produktentwicklungs- bzw. Gestaltungsprozess. Es fehlen allerdings eindeutig formulierte methodologische Vorgaben für die Entwicklung der Untersuchungsdesigns.

Es taucht der Verdacht eines naiven induktions-orientierten Empirismus auf, wenn z.B. vorgeschlagen wird, allein aus einer Nutzer-, Aufgaben- und Kontextanalyse nutzerorientierte Designvorgaben zu formulieren, wenn Beobachtungen, Befragungen nahe gelegt werden, ohne darauf hinzuweisen, dass hier vorher geklärt werden muss, welche theoretischen Annahmen, Hypothesen, Beobachtungskategorien etc. vorliegen.

Naiv empiristisch erscheinen die vielfältigen, empfohlenen Methodenkataloge, die zwar eine Einbeziehung von Nutzern garantieren und sagen, wann welche Methode im Produktentwicklungsprozess an wie vielen Personen anzuwenden ist. Aber Erhebungs- und Auswertungsmethoden sind nicht bloße Werkzeuge. Ein Verweis auf die Reliabilität und Validität reicht da nicht aus. Es ist Traxel (nach Gigerenzer 1981, S. 29) zu zitieren: „Es muss eindringlich betont werden, dass die Methoden den Charakter von Werkzeugen haben. Die Methodik ist ein Arsenal von Werkzeugen; die Methodologie ist eine wissenschaftliche Werkzeuglehre. Dies möge stets im Blick behalten werden. Die Methoden lehren uns nicht, wie wir neue Probleme entdecken können, welche Fragestellungen als wichtig oder weniger wichtig zu gelten haben und welcher weitere Gebrauch von den Ergebnissen gemacht werden kann.“

Naiv induktionistisch erscheint auch die Hoffnung aus phänomenologischen Fehleranalysen mit quantitativen Daten, z.B. zu Reaktionszeiten und Fehleranzahl, Hinweise für die Gestaltungsoptimierung zu erhalten (vgl. Kanis 2000). Fehlern liegen konfundierende Effekte zugrunde, die nur durch kausale Fehlermodelle und Wirkprinzipien er-

klärt und behoben werden können. Die empfohlenen Design-Richtlinien bzw. Prinzipien geben auf den ersten Blick den Anschein eines theoriegeleiteten, deduktiven Vorgehens. Doch sie haben eher Rezeptcharakter. Sie stellen kein problemlösungsorientiertes Wissen bereit, das den Gestaltern a) die Notwendigkeit und die Hintergründe der Prinzipien erklärt und das sie b) befähigt dieses Wissen selbständig auf praktische Probleme anzuwenden. Doch dies brauchen sie, sodass hier ein naiv deduktionistisches Vorgehen nahe gelegt wird. „Wissenschaftlich fundierte Praxis ist ein methodisches Vorgehen.... die Übernahme tradierten Know-hows, daumenregelgestützte Routinen usf. machen eine Tätigkeit noch nicht zu einer wissenschaftlichen Tätigkeit.“ (Herrmann 1993 S. 168)

Die folgenden Zitate sind zwar älter, aber sie geben meines Erachtens nach den methodologischen Stand der Nutzerforschung immer noch treffend wieder. Paul J. Barber beklagte bereits 1988 (S. 7): „*It is often said, that what is applicable is not so much the theoretical stuff of psychology as its elaborate repertoire of methods of investigation and discovery.*“ und forderte eine Verbindung von Grundlagenforschung und Anwendung. Doch bis heute hat sich aber m.E. die von ihm beklagte Situation nicht grundlegend geändert. Vielmehr scheint das Vorgehen so zu sein, wie es Hans-Jörg Bullinger Leiter des Fraunhofer Institutes für Arbeitswissenschaften und Organisation beschreibt (1994 S. 272): „*Eine erfolgreiche Produktentwicklung wird in zunehmendem Maße von der Methodenintegration in den unterschiedlichen Entwicklungs-(Gestaltungs)phasen geprägt.*“

Es fehlt ein methodisches Vorgehen für Input- und Evaluations-Studien, das Aussagen über das Verhältnis von Problemstellung (=Aufgabenstellung), Theorien, Methoden und Ergebnisinterpretation macht. Luczak (1993, S. 351 ff) begründet dieses Fehlen mit einem Bedürfnis der Praktiker nach schnell verfügbaren Routinen und Regeln. Meine Erfahrungen als Praktikerin sind gegenteilig: erst wenn man den Problemlösungsweg prinzipiell versteht und praktizieren kann, hat man die nötigen Erfahrungen, Fertigkeiten und Beurteilungskriterien, um Routinen und Regeln sicher und gezielt anwenden zu können. Außerdem sind viele Praxisprobleme meiner Erfahrung nach nicht mit Routinen zu lösen, weil es hierfür noch keine Routinen gibt oder weil das Problem viele ideosynkratische Anteile hat. Dann ermöglicht erst eine theoretische Fundierung ein tieferes Verständnis der Wirkzusammenhänge und gewährleistet nicht nur Sicherheit und Ori-

entierung im Problemlösungsprozess, sondern liefert auch einen geeigneten Ansatz. Im Folgenden wird daher ein methodisches Vorgehen für die Lösung des Anordnungsproblems entwickelt.

2.2 Wissenschaftstheoretische Überlegungen zum Vorgehen

Wie muss ein Problemlösungsprozess aussehen, der wissenschaftlichen Standards entspricht, wie Nachvollziehbarkeit, Überprüfbarkeit, Wiederholbarkeit, Gültigkeit und Übertragbarkeit für andere Anwendungsfälle? Im folgenden Abschnitt wird diese Frage aus der Sicht der Wissenschaftstheorie untersucht.

2.2.1 Zum Verhältnis von Theorie und Anwendung

Die Aufgabe der Wissenschaft ist – aus Sicht der Wissenschaftstheorie – teils theoretisch und teils praktisch. Theoretisch ist sie dort, wo sie Theorien und Gesetze entwickeln soll, um Erklärungen zu geben. Praktisch ist sie dort, wo sie Voraussagen treffen bzw. helfen soll, praktische Probleme zu lösen und praktische Anwendungen zu entwickeln (vgl. Popper 1973, S. 377).

Die Wissenschaftstheoretiker Hempel und Oppenheim formulierten folgende Aussagen über das Verhältnis von Erklärungen, Vorhersagen und praktischen Anwendungen, in dem nach ihnen benannten *Hempel-Oppenheim-Schema* (siehe Popper 1973): Sie behaupten, dass theoretischen Erklärungen, Vorhersagen und praktischen Anwendungen die gleiche logische Argumentationsstruktur zugrunde liegt! In allen drei Fällen braucht es zwei Prämissen: 1.) die Kenntnis des betreffenden allgemeinen Gesetzes sowie 2.) die Kenntnis der spezifischen Anfangs-, Rand- und Situationsbedingungen, unter denen dieses Gesetz gilt (siehe Abb. 2.6).

Ein Beispiel soll das Gesagte verdeutlichen: für die Erklärung, warum eine Brücke hält, als auch für die Vorhersage, ob sie halten wird, braucht es die Kenntnis von Ge-

setzmäßigkeiten, wie des Hebelgesetzes, sowie die Kenntnis der spezifischen Situationsbedingungen, wie Materialbeschaffenheit, Wind- und Strömungsverhältnisse etc. Für den Bau einer Brücke braucht es dieses Wissen u.U. nicht unbedingt explizit. Es kann also sein, dass der Brückenbauer intuitiv diese Gesetzmäßigkeiten verwirklicht. Aber – nach dem Stand der Erkenntnis – müssen sie verwirklicht sein, damit die Brücke auch hält und der Brückenbauer wird sich leichter tun, wenn er sie kennt und anwendet.

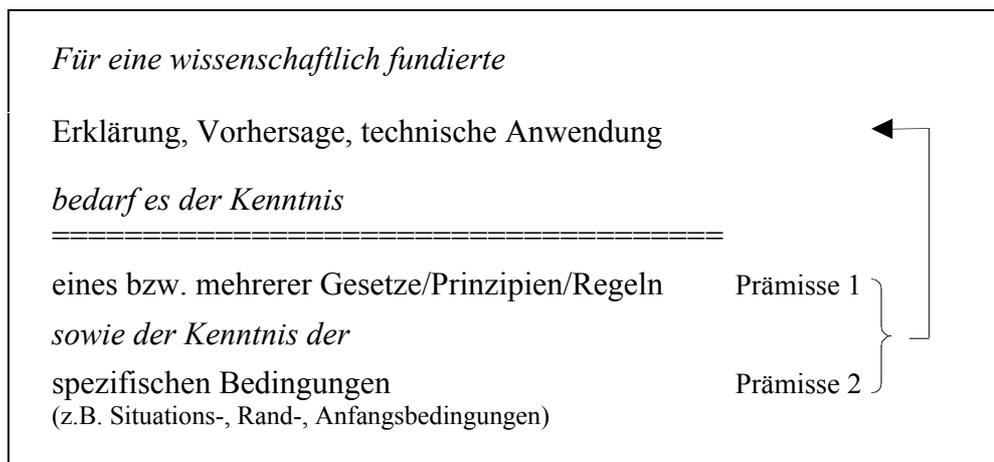


Abbildung 2.6. Theorie-Praxis-Verhältnis nach Hempel/Oppenheim (visualisiert nach Popper 1973)

Exkurs: Begriffsdefinitionen von Gesetz, Theorie, Modell Im Folgenden wird *Gesetz* synonym verwendet für die Begriffe Regel, Prinzip, Wirkprinzip, Gesetzmäßigkeit. Alle Versuche, verbindliche Kriterien für die Gesetzesartigkeit von Aussagen sowie eindeutige Abgrenzungskriterien für die oben genannten Begriffe anzugeben, sind bisher gescheitert (vgl. Popper 1976, Jammer 1994, Mehrrens 1999). Ein Gesetz – als eine Allaussage – erhebt Anspruch auf die Gültigkeit an allen Orten und zu allen Zeiten. Die empirische Überprüfbarkeit des Vorliegens der Situationsbedingungen und des Explanandums hingegen bezieht sich auf eine konkrete, ideosynkratische Situation, d.h. auf singuläre, historische Ereignisse.

Eine *Theorie* ist ein wissenschaftliches Lehrgebäude, das nach definierten, in der Methodologie der entsprechenden Wissenschaft beschriebenen Methoden entsteht (Seiffert 1994). Im strengeren Sinn ist eine Theorie ein System von Aussagen, das mehrere Hypothesen und Gesetze umfasst. Die oben genannten Begriffe werden in der folgenden Arbeit im Sinne von Popper (1934, 1976) als von Menschen kreierte und verfertigte Aussagen verstanden, die die Regeln und Wirkzusammenhänge, denen die Wirklichkeit folgt, abbilden sollen, die aber grundsätzlich Vermutungswissen darstellen und mittels des Falsifikationsprinzips verworfen werden können, und – falls vorhanden – durch geeignetere (im Sinne von Occam's Razor mit weniger Annahmen und Einschränkungen) ersetzt werden können. Unter *Modell* wird ein Abbild, eine Repräsentation verstanden, die das Original für bestimmte (Modell-)Benutzer, in einer bestimmten

Zeitspanne, hinsichtlich eines definierten Zweckes repräsentiert (Stachowiak 1994). In der Psychologie wird der Modellbegriff teils synonym mit Theorie verwendet.

Erklärung, Vorhersage und Anwendung unterscheiden sich nur hinsichtlich des Zieles: Während bei der Erklärung das zu erklärende Ereignis (Explanandum) bereits bekannt ist und nach dem zugrunde liegenden Gesetz und den Situationsbedingungen gesucht wird, handelt es sich bei der Vorhersage bzw. bei einer Anwendung um ein vorherzusa-gendes bzw. in der Zukunft noch unbekanntes, zu realisierendes Ereignis (Abb. 2.7.).



Abbildung 2.7. Zum Verhältnis von Erklärung, Vorhersage und Anwendung: die Grundlagenforschung sucht nach Gesetzen und Situationsbedingungen, die Anwendungsforschung wendet sie an, um technische Lösungen oder Vorhersagen zu entwickeln

Der Grundlagenforscher versucht demnach für ein Ereignis das zugrunde liegende Gesetz und die spezifischen Situationsbedingungen aufzudecken und in einer Theorie zu formulieren. Gegebenenfalls wendet er die Theorie an, um sie zu überprüfen (Robotik, Reengineering). Der Praktiker wendet in der Regel die Theorie bzw. das Gesetz und die Kenntnis der Situationsbedingungen an, um eine Vorhersage über ein Ereignis zu machen oder um eine technische Anwendung zu realisieren. Diese Gleichsetzung von Erklärung, Vorhersage und praktischer Anwendung wird in der Wissenschaftstheorie unter dem Begriff der „Strukturellen Gleichheitsthese“ erörtert (vgl. Popper 1973, Küttner 1994).

Dem *Hempel/Oppenheim-Schema* liegen folgende Annahmen zugrunde:

1. Jedem Ereignis liegt ein Gesetz (ein oder mehrere Gesetze oder Gesetzesketten) zugrunde, das das Ereignis „verursacht“. Dieses *Prinzip der Nomologisierung* (principle of lawfulness) hat sich in der Geschichte der Wissenschaft bewährt und ist von großem heuristischen Wert für die Suche nach neuen Gesetzen. Das Gesetz kann bereits entdeckt und formuliert sein. Dann steht es Praktikern für Vorhersagen oder Anwendungen zur Verfügung. Es kann aber auch sein, dass das entsprechende Gesetz von Grundlagenforschern noch zu entdecken und zu formulieren ist.
2. Die Kenntnis des zutreffenden Gesetzes ist eine notwendige, aber noch keine hinreichende Bedingung zur Erklärung oder Vorhersage eines Ereignisses bzw. zur Entwicklung einer praktischen Anwendung. Jedes Ereignis ereignet sich in einer spezifischen Situation mit spezifischen Bedingungen. Es bedarf also darüber hinaus auch noch der Kenntnis der spezifischen bzw. typischen Anfangs-, Rand- und Situationsbedingungen, unter denen dieses Gesetz gilt.

Wichtigste Kritik (Küttner 1994) fand das Hempel-Oppenheim-Schema bei Stegmüller (1973), der es nur für nomologisch-deduktive (kausal-bedingte) Ereignisse bejaht, aber nicht für statistisch-induktive (korrelativ-bedingte) Ereignisse (vgl. Küttner 1994, Sarris 1990). Als Beispiele für eine statistisch-induktive Prognose werden z.B. Wahlergebnisse, Wettervorhersagen oder Wirtschaftsprognosen genannt (vgl. Sarris 1990). Im Folgenden wird jedoch davon ausgegangen, dass es sich bei dem Problem der nutzerorientierten Anordnung nicht um ein statistisch-induktives Ereignis handelt. Es wird vielmehr ein kausales Verhältnis unterstellt. Und es wird – im Sinne des Prinzips der Nomologisierung – angenommen, dass eine oder mehrere Gesetzmäßigkeiten im Rahmen definierter Situationsbedingungen dafür verantwortlich sind, dass Nutzer leicht oder schwer das finden, was sie suchen.

2.2.2 Erste methodologische Konsequenzen

Wendet man das Hempel-Oppenheim-Schema auf die Problemstellung dieser Arbeit an, ergeben sich daraus folgende Konsequenzen: Für einen wissenschaftlichen Problemlösungsprozess braucht es ein explizites, theoriegeleitetes und an Nutzern empirisch fun-

diertes Vorgehen! Das bedeutet, dass als Erstes Theorien und Gesetzmäßigkeiten zu bestimmen sind, die die ursächlichen Bedingungsfaktoren für Such- und Findeprozesse in Anordnungen benennen, erklären und vorhersagen können und natürlich auch die Situationsbedingungen, unter denen sie gelten. Dann ist ein entsprechendes Untersuchungsdesign zu entwickeln und durchzuführen, um an den entsprechenden NutzerInnen die spezifischen Vorgaben für eine nutzerorientierte Anordnung zu ermitteln.

In dieser Arbeit wird somit eine Methodologie verfolgt, die man als ein *deduktiv-theoriegeleitetes und empirisch-fundiertes Vorgehen* beschreiben kann. Es zwingt zum einen die theoretischen Grundlagen zu klären und auf das Problem passende Theorien und Gesetze anzuwenden. Und es fordert zum anderen eine empirische Untersuchung, die spezifische, nutzerorientierte Lösungen liefert.

Dieses hier skizzierte Vorgehen unterscheidet sich im Ansatz grundlegend von dem induktionistischen Vorgehen, wie es Carroll (1991; siehe oben) vertritt. Es unterscheidet sich auch von dem aktuellen Vorgehen in der Nutzerforschung oder vom Vorgehen wie es die ISO-Normen 13407 oder 9241-10 empfehlen. Die gängige Praxis wäre dort zum einen a) z.B. NutzerInnen zu befragen, wie sie sich eine such- und findefreundliche Anordnung wünschen, was man als induktiv-phänomenologisch-empirisches Vorgehen beschreiben kann; sowie zum anderen b) Design-Richtlinien und Regeln vorzugeben, die die Gestalter einzuhalten haben, was man als simplifiziertes, deduktionistisches Vorgehen beschreiben kann. Nachdem nun das prinzipielle Vorgehen im Ansatz geklärt ist, eröffnet sich die Frage nach dem konkreten, methodischen Vorgehen.

2.3 Entwicklung eines methodisches Vorgehens

Im Folgenden wird ein methodisches Vorgehen für Lösung praktisch angewandter Problemstellungen – wie die Entwicklung nutzerorientierter, erwartungsgerechter Anordnungen – entwickelt. Hierfür wird zunächst ein Ausflug in die Methodologie der Theorieentwicklung unternommen, um von dort die allgemeine Grundlagen für eine Anwendungsmethodologie zu klären.

2.3.1 Methodologie der Anwendungsentwicklung in Analogie zur Methodologie der Theorieentwicklung

Um eine Anwendungsmethodologie zu formulieren, wird folgende Überlegung angestellt: Die „Strukturelle Gleichheitsthese“ von Hempel und Oppenheim geht davon aus, dass Erklärung (Theorien) und Anwendung (Problemlösungen) die gleiche logische Struktur zugrunde liegt: für beides braucht es die Kenntnis der entsprechenden Gesetzmäßigkeiten und Situationsbedingungen, die das Ereignis bedingen bzw. herstellen lassen. Wenn Erklärung und Anwendung die gleiche logische Struktur zugrunde liegt, dann liegt auch der Methodologie der Theorieentwicklung und der Anwendungsentwicklung die gleiche logische Struktur zugrunde. Es wird daher im Weiteren davon ausgegangen, dass sich eine Methodologie der Anwendungsentwicklung analog zur Methodologie der Theorieentwicklung verhält.

In der Literatur finden sich sehr elaborierte Darstellungen zur Methodologie der Theorieentwicklung. Unterschiedliche Autoren weisen dabei nur geringfügige Unterschiede auf (vgl. Bortz 1985, Sarris 1990). Im Folgenden werden diese Phasen skizziert sowie die wesentlichen Anforderungen, die im Prozess der Theorieentwicklung erfüllt sein müssen. Ziel ist es, diese dann auf eine Methodologie der Anwendungsentwicklung zu übertragen. Auf eine ausführliche Darstellung der Methodologie der Theorieentwicklung wird verständlicherweise verzichtet, da sie in jedem Standardwerk der Methodenlehre nachzulesen ist.

So unterscheiden z.B. sowohl Bortz (1985 S. 3) als auch Sarris (1990 S. 113) sieben Phasen im Theorieentwicklungsprozess: In der ersten Phase findet die Problemerkundung und -formulierung (Literatursichtung, Explorationsstudien etc.) statt mit dem Ziel, den wissenschaftlichen Status des Problems kennen zu lernen und einen ersten theoretischen Rahmen zu formulieren. In der zweiten Phase wird die Theorie formuliert und logisch und begrifflich geprüft. In der dritten Phase werden von dieser Theorie Hypothesen abgeleitet, um die Theorie an der Realität zu überprüfen. In der vierten Phase wird ein Untersuchungsplan zur Hypothesentestung entworfen. Die Variablen werden operationalisiert, die Untersuchungsstichproben, die Erhebungs- und Auswertungsmethoden werden festgelegt. In der fünften Phase findet die Versuchsdurchführung statt. In der

sechsten Phase werden die Daten analysiert und ausgewertet. In der siebten Phase werden die Ergebnisse interpretiert und die Theorie wird im Hinblick auf die Ergebnisse bewertet (siehe auch Abb. 2.8.).

Für die inhaltliche Realisierung dieser Phasen werden eindeutige Anforderungen formuliert, damit Untersuchungsgegenstand (Problem), Theorie und Methoden theoretisch wie empirisch aufeinander abgestimmt sind: Das theoretische Modell und der Untersuchungsgegenstand stehen in einer Interaktion und müssen daher zusammenpassen (vgl. Gigerenzer 1981, Bortz 1985, Sarris 1990). Erhebungs- und Auswertungsmethoden sind nicht bloße Werkzeuge. Sie implizieren selbst theoretische Inhalte, wie ein messtheoretisches Skalierungsmodell (Traxel, zitiert nach Gigerenzer 1981). Diese dürfen nicht ausgeblendet werden, weil sie Einfluss auf die Daten haben. Eine Trennung von Methode und Inhalt sowie von Messung und Theorie geht sonst auf Kosten der Interpretation der Untersuchungsergebnisse. Die Interpretation der Ergebnisse kann nur in Bezug auf die Hypothesen bzw. die zugrunde liegende Theorie erfolgen. Insgesamt müssen die sieben Phasen ein aufeinander abgestimmtes Bezugssystem ergeben.

2.3.2 Darstellung des methodischen Vorgehens

Im Folgenden wird ein methodisches Vorgehen zur Entwicklung nutzerorientierte Anwendungen formuliert. Es besteht aus zehn Phasen, wobei die ersten sieben Phasen analog der Theorieentwicklung sind – natürlich mit dem großen Unterschied, dass hier nicht eine Theorie entwickelt, sondern gesucht und empirisch auf das Problem angewendet werden muss. Die Phasen acht, neun und zehn müssen zusätzlich eingeführt werden, um die gewonnene Problemlösung in Gestaltungsvarianten umzusetzen und diese dann an Nutzern zu überprüfen, ob sie tatsächlich nutzerorientiert ist und die gewünschten Verhaltensweisen hervorruft. Abbildung 2.8 gibt einen ersten Überblick über die zehn Phasen im Vergleich zu den Phasen der Theorieforschung.

Zum Vergleich:
Phasen der Theorieentwicklung
nach Bortz (1985) und Sarris (1990)

1. Problemanalyse
2. Theorieentwicklung
3. Hypothesenformulierung
4. Untersuchungsdesign
5. Durchführung des Versuchs
6. Datenanalyse und Auswertung
7. Interpretation und Entscheidung

Entwurf für diese Arbeit:
**Phasen einer theoretisch und empirisch fundierten
Anwendungsentwicklung**

1. Problemanalyse, Eingrenzung auf erste Theoriefelder
2. Auswahl einer zum Problem passenden Theorie
3. Anwendung der Theorie auf das Problem, Hypothesen
4. Entwicklung des Untersuchungsdesigns
5. Durchführung des Versuchs
6. Datenanalyse und Auswertung
7. Interpretation und Entscheidung
8. Formulieren der Vorgaben für die Gestalter



Gestalter entwickeln Gestaltungsvarianten


9. Überprüfen der Gestaltungsvarianten an Nutzern
10. Bestimmen der besten Gestaltungsvariante

Abbildung 2.8. Phasen der Anwendungsentwicklung im Vergleich zu den Phasen der Theorieentwicklung

Dieses Vorgehen legt zwei empirische Untersuchungen an NutzerInnen nahe: eine Input-Studie (Phasen 1–8) zur Ermittlung von Gestaltungsvorgaben, wie Objekte nutzerorientiert anzuordnen sind. Und eine Evaluations-Studie (Phase 9–10) zur Überprüfung von Gestaltungsvarianten, d.h. von Anordnungsvarianten. Abbildung 2.9 visualisiert die Studien. Im Folgenden wird das methodische Vorgehen für die Input- und Evaluations-Studie jeweils vorgestellt und erläutert.

Vorgehen und Anforderung für die Theoretische Fundierung

1. Phase: Problemanalyse und Eingrenzung auf erste Theoriebereiche In der ersten Phase müssen zunächst das Problem und die Aufgabenstellung analysiert werden, mit

dem Ziel eine erste, grobe, theoretische Einordnung vorzunehmen: Liegen dem Problem eher wahrnehmungspsychologische, kognitions-, bewegungs-, problemlösungs- oder gar motivationspsychologische Prozesse zugrunde und welche genau?

Um dieses Ziel zu erreichen, stehen unterschiedliche Instrumente zur Verfügung; z.B. a) qualitative Erkundungsstudien, bei denen Nutzer befragt werden, wie sie sich in Suchsituationen verhalten und wovon für sie der Finde-Erfolg abhängt. Dabei dienen die Ergebnisse der Befragung der Problemerkundung und Theorieeingrenzung, nicht aber der Formulierung von Gestaltungsvorgaben! b) formale Analysen des Untersuchungsgegenstandes, z. B. mit der Frage: Wie lässt sich eigentlich eine Anordnung von Objekten im Raum beschreiben? Was ist mit Anordnung gemeint? c) Sichtung der einschlägigen Literatur, um zu klären, wie Anwendungs- oder Grundlagenforscher das Anordnungs-, Such- und Findeproblem theoretisch einordnen bzw. bisher praktisch lösen. d) Informationsgespräche mit Praktikern und Forschern, die am gleichen Problem arbeiten. Diese Phase muss sorgfältig und kritisch durchgeführt werden, um sich eventuell nicht zu früh auf nicht lösungsgerechte oder bisher gängige Theoriebereiche festzulegen, was dann zu routinierten, nicht neuen und ineffektiven Problemlösungen führen kann (vgl. Witte 1991).

2. Phase: Auswahl einer geeigneten Theorie Diese Phase hat das Ziel eine geeignete Theorie auszuwählen, die in der Lage ist, die Wirkstruktur, die Bedingungsfaktoren, o.Ä., die dem Problem zugrunde liegen, zu erklären und Vorhersagen zu machen, was zu tun ist, um das Problem zu lösen. Diese Phase unterscheidet sich somit grundlegend von der zweiten Phase der Grundlagenforschung, bei der eine Theorie zu formulieren ist!

Für die Auswahl einer geeigneten Theorie sind zunächst *Auswahlkriterien* zu definieren. Nach der Definition der Auswahlkriterien sind Theorien aus der Literatur zu sichten, zu bewerten und die geeignete auszuwählen. Eventuell entwickeln sich auch die Auswahlkriterien parallel zur Sichtung der Theorien. Ergebnis dieser Phase sollte (möglichst) eine Theorie sein, die den Auswahlkriterien entspricht und sich für die Anwendung auf das Problem eignet.

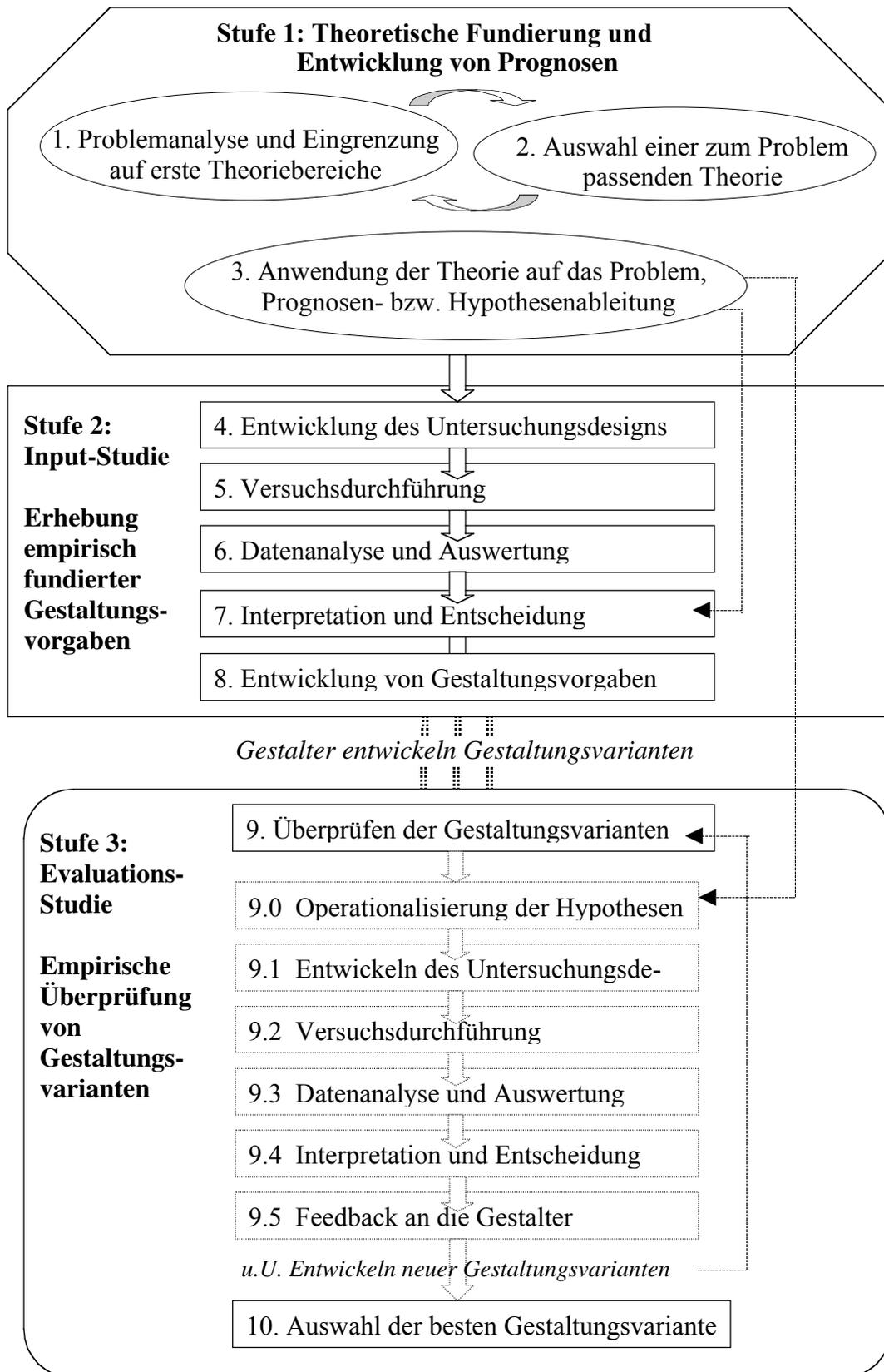


Abbildung 2.9. Methodisches Vorgehen für die Entwicklung nutzerorientierter Lösungen

Es reicht nicht, dem Problem Gestaltungsrichtlinien oder -prinzipien, wie sie in ISO-Normen oder Handbüchern formuliert sind, zuzuordnen. Für eine „gute“, nutzerorientierte Problemlösung bedarf es einer Theorie, deren Gesetzmäßigkeiten auf das Problem spezifisch – unter Berücksichtigung der Situationsbedingungen – anzuwenden sind, um eine gute Lösung zu erarbeiten. Das leisten Richtlinien und Prinzipien von Normen alleine nicht, weil sie allenfalls Rezepte liefern.

Bei der Theoriesichtung ist zu bedenken: Viele „Theorien“ der Psychologie sind erst in einem frühen Entwicklungsstatus und liefern allenfalls Beschreibungskategorien. In einem solchen Fall fehlen Gesetzmäßigkeiten, die auch Vorhersagen ermöglichen (vgl. Sarris 1990). Die Psychologie ist eine relativ junge Wissenschaft und hat im Allgemeinen den Status einer Protowissenschaft (vgl. Sarris 1990, Witte 1991): Einzelne psychologische Fachbereiche (wie die Wahrnehmungspsychologie oder die Psychophysik) haben bereits den Status einer „harten“, andere (wie die Emotions-, Handlungspsychologie) haben hingegen erst den Status einer „weichen“ Wissenschaft. Somit sind die Theorien – je nach Fachbereich – in einem sehr unterschiedlich elaborierten und empirisch fundierten Zustand. Für den Praktiker hat das die Konsequenz, dass er für sein Problem im schlimmsten Falle nur „ungenügende“ (d.h. wenig aussagestarke, ungenügend geprüfte etc.) oder aber gar keine Theorie findet. Häufig stehen aber auch mehrere Theorien unterschiedlicher Güte nebeneinander und der Praktiker muss entscheiden, welche der Theorien die seinem Problem zugrunde liegenden Wirkprinzipien am besten abbildet.

3. Phase: Anwendung der Theorie auf das Problem Da die Theorie wahrscheinlich mehrere gesetzesartige Aussagen beinhaltet, sind zunächst diejenigen Gesetzmäßigkeiten auszuwählen, die zu dem Problem passen und die Wirkstruktur des Problems beschreiben lassen. Aus diesen Gesetzmäßigkeiten sind Prognosen im Sinne von Hypothesen zu formulieren, die eine Vorhersage erlauben, unter welchen Bedingungen das Such- und Findeverhalten der Nutzer erleichtert wird. Im weiteren Prozedere müssen diese Hypothesen dann empirisch überprüft bzw. inhaltlich konkretisiert werden. Es wird daran erinnert, dass nach dem Hempel-Oppenheim-Schema nicht nur die Gesetzmäßigkeit notwendig ist, sondern auch die Kenntnis und Berücksichtigung der spezifischen Bedingungen. Es kann also sein, dass sich das Gesetz bei unterschiedlichen Objekten, NutzerInnen oder Suchsituationen anders realisiert.

Vorgehen und Anforderung für die Input-Studie

Ziel der Input-Studie ist es, die Gestaltungsvorgaben empirisch zu erheben.

4. Phase: Entwicklung des Untersuchungsdesigns In der vierten Phase wird das Untersuchungsdesign entwickelt. Eine empirische Untersuchung braucht es aus folgenden Gründen: Zum einen ist eine empirische Untersuchung notwendig, um die Vorgaben empirisch an den Nutzern zu erheben. Es kann z.B. sein, dass das Such- und Findeverhalten von den Personenvariablen der NutzerInnen abhängt. Dies kann nur mit unterschiedlichen Nutzer-Stichproben herausgefunden werden. Oder aber das Such- und Findeverhalten ist abhängig von Situationsvariablen. Dann müssten die Situationsvariablen im Versuch variiert werden.

Zum anderen soll eine empirische Untersuchung überprüfen, ob die ausgewählte Theorie passt. Denn stimmt das tatsächliche Verhalten der NutzerInnen mit dem prognostizierten Verhalten überein, dann ist dies ein Hinweis darauf, dass die Theorie geeignet ist, die zugrunde liegenden Wirkfaktoren zu benennen und damit das Verhalten zu prognostizieren. Dann ließe sich die Theorie bzw. die Gesetzmäßigkeit auf ähnliche Anordnungsprobleme übertragen! (Was aber nicht bedeutet, dass es nicht noch geeignetere Theorien geben könnte!)

Im Untersuchungsdesign sind die abhängigen, unabhängigen und die zu kontrollierenden Variablen zu definieren, die Stichprobenkriterien sind festzulegen sowie die Erhebungs- und Auswertungsmethoden. Die Anforderungskriterien an das Untersuchungsdesign sind die gleichen, wie sie an ein Untersuchungsdesign im Kontext einer Theorieentwicklung gestellt werden. Sie sind ausführlich in Lehrbüchern der Methodologie des Forschens nachzulesen (vgl. Bortz 1985, Sarris 1990). So müssen z.B. Erhebungsmethoden, Auswertungsmethoden, Stichproben und Theorie aufeinander abgestimmt sein.

5. Phase: Versuchsdurchführung Jetzt findet die Durchführung der Untersuchung statt. Wird die Untersuchung gründlich geplant und das Untersuchungsdesign sorgfältig entwickelt, dürfte es in dieser Phase keine Schwierigkeiten geben.

6. Phase: Datenanalyse und Auswertung Hier findet die statistische Analyse und Auswertung der erhobenen Daten statt, mittels der im Untersuchungsdesign bestimmten

statistischen Auswertungsverfahren. Die Auswertungsverfahren werden angewendet und die Ergebnisse beschrieben.

7. Phase: Interpretation und Entscheidung In der siebten Phase findet die Interpretation der Ergebnisse in Bezug auf die Hypothesen statt. Stimmt das tatsächliche Verhalten mit den Prognosen (= Hypothesen) überein, dann lässt dies zwei Schlussfolgerungen zu: a) die Theorie ist hinsichtlich des Problems geeignet, b) die Ergebnisse sind in die Problemlösung zu übersetzen. Die Ergebnisse werden somit nicht „frei“ interpretiert, sondern im Hinblick auf die Hypothesen. Und die Ergebnisse zeigen klar, was getan werden muss. Für diese Arbeit heißt das, dass die Ergebnisse vorgeben, wie die Objekte anzuordnen sind, damit die NutzerInnen schnell und einfach das finden, was sie suchen. Um zu überprüfen, ob die Ergebnisse tatsächlich ein schnelleres Finden ermöglichen, braucht es noch weitere Phasen:

8. Phase: Entwicklung der Gestaltungsvorgaben In der achten Phase sind die Ergebnisse für ein Gestaltungsbriefing aufzubereiten, damit die Gestalter die Lösung umsetzen können. Sie brauchen Vorgaben, wie die Objekte bzw. Funktionen im Raum, im Regal oder auf der Bedieneroberfläche „nutzerorientiert“ anzuordnen sind. Dabei ist zu beachten, dass das Briefing in einer für den Gestalter angemessenen Form aufbereitet wird, damit es auch eindeutig und unmissverständlich verstanden wird. Dann können diese Vorgaben von Gestaltern in konkrete Gestaltungsvorschläge umgesetzt werden.

Vorgehen und Anforderung für die Evaluations-Studie

Ziel der Evaluations-Studie ist es, Gestaltungslösungen an NutzerInnen zu überprüfen.

9. Phase: Überprüfen der Gestaltungsvarianten In der neunten Phase findet dann schließlich eine empirische Überprüfung der Gestaltungsrealisierungen statt. Gestaltungsrealisierungen sind z.B. Regalbelegungen, Modelle, Prototypen. Die Gestalter entwickeln – wie bereits erwähnt – in der Regel mehr als eine Gestaltungsvariante. Die Überprüfung dieser Varianten muss dabei in *einer eigenen Evaluations-Studie* durchgeführt werden. Hierfür gilt ein identisches Vorgehen: Formulierung der Hypothesen (9.1), Entwicklung eines Untersuchungsdesigns (9.2), Versuchsdurchführung (9.3),

Datenanalyse und Auswertung (9.4), Interpretation und Entscheidung (9.5) und schließlich das Feedback an die Gestalter (9.6). Dabei muss die Operationalisierung der Hypothesen, die Definition der Stichproben, der Erhebungs- und Auswertungsmethoden, die Datenanalyse, die Interpretation etc. wiederum den wissenschaftlichen Anforderungen genügen, wie sie die Methodologie des Forschens fordert. Die Evaluations-Studie ist durchzuführen und die Ergebnisse sind in Hinblick auf die Hypothesen zu interpretieren. Ergebnis so einer Evaluations-Studie ist a) festzustellen, welche Gestaltungsrealisierung die beste ist; und möglicherweise b) Verbesserungsvorschläge zu formulieren. Zeigen sich starke Probleme, so müssen die Gestalter u.U. unter der Berücksichtigung der Erkenntnisse der Evaluations-Studie die Gestaltungsvarianten überarbeiten. Eventuell ist es notwendig die neuen Varianten wieder in einer Evaluations-Studie zu überprüfen.

Phase 10: Auswahl der besten Gestaltungsvariante Abschluss des Prozesses ist die Auswahl und Bestimmung der besten Gestaltungsvariante.

So weit zur allgemeinen Darstellung. Das Ziel, ein allgemeines, methodisches Vorgehen für die Lösung des Anordnungsproblems zu entwickeln, ist somit erreicht. Bevor es konkret umgesetzt wird, erfolgt noch ein Blick über die Fachgrenzen hinüber in den Fachbereich Maschinenbau zur Konstruktionsmethodik. Beim Konstruieren werden auch Probleme gelöst und technische Anwendungen realisiert. Daher soll untersucht werden, welches methodische Vorgehen sich dort etabliert hat.

2.4 Zum Vergleich: Methodisches Vorgehen im Maschinenbau

Erst seit Anfang der 1960er Jahre gibt es innerhalb der Maschinenbaulehre den eigenen Fachbereich Konstruktionslehre. Ziel ist es, sowohl eine allgemein gültige, fach- und produktneutrale als auch eine produktspezifische Methodologie des Konstruierens und Problemlösens zu entwickeln (Koller, 1994). Aus der Geschichte sind zwar die Ergebnisse genialer Konstrukteure bekannt, nicht aber ihre Methoden. Weder zu den Arbeiten von Archimedes, Leonardo da Vinci, James Watt oder Konrad Zuse gibt es Informatio-

nen über die Vorgehensweisen, die zu solch großen Ingenieursleistungen geführt haben (Koller 1994).

Pahl & Beitz (1977, 1997) geben einen kurzen historischen Abriss und zitieren Reuleaux von 1854, der angesichts der Komplexität des Konstruierens noch resignierend die Leistung allein der Kunst des Ingenieurs zuschrieb. „Allein die Inbetrachtziehung aller dieser Umstände und ihre richtige Würdigung können nicht in einer absoluten Form gesehen und daher weder allgemein behandelt noch eigentlich gelehrt werden. Sie sind vielmehr einzig Sache der Intelligenz und des Scharfblicks des entwerfenden Ingenieurs.“ Als eigentlicher Ausgangspunkt methodischen Konstruierens gelten die Arbeiten von Wörgerbauer (1943) (vgl. Koller 1994, Pahl & Beitz 1997). Und in der Folge wurden Phasen des Konstruierens herausgearbeitet, die inzwischen in der VDI Richtlinie 2221 als allgemein verbindliche Empfehlung niedergeschrieben wurden.

Im Folgenden wird die Vorgehensweise einer Produktentwicklung, wie sie die VDI-Richtlinie 2221 (1996) empfiehlt, vorgestellt (Abb. 2.10). Ziel ist es, zu zeigen, dass innerhalb des Konstruktionsprozesses der Bezug zwischen dem praktischen Problem einerseits und Theorien und Gesetzmäßigkeiten andererseits sehr bewusst reflektiert wird. Die Anwendung von Gesetzen und Theorien im Problemlösungsprozess ist hier Standard. Ingenieure lernen sehr explizit z.B. physikalische, chemische, biologische Wirkprinzipien/Gesetze (leider nicht humanwissenschaftliche) und die typischen Situationsbedingungen, unter denen sie gelten, im kreativen Konstruktionsprozess anzuwenden.

Pahl und Beitz (1997) unterscheiden im Konstruktionsprozess vier Ebenen: Der Spezifikation der Funktionen (1. Funktionsebene) folgt die Funktionsdarstellung auf der Basis der entsprechenden physikalischen, chemischen etc. Wirkprinzipien (2. Prinzip-ebene), um im Zuge der weiteren Konkretisierung die Bauelemente festzulegen (3. Gestaltungsebene) sowie die herstellungstechnische Dokumentation (4. Ausarbeitungsebene). Prinzipiell wird der Konstruktionsprozess in sieben Arbeitsschritten vorgeschlagen. Sie sind in Abbildung 2.10 dargestellt. Es handelt sich nur um „elementare“ Arbeitsschritte für eine strategische Handlungsanweisung. Detaillierte Schritte werden nicht aufgeführt, da sie problemspezifisch sind und von der Person des Problemlösers abhängen (Pahl & Beehlich 1981, Pahl & Beitz 1997).

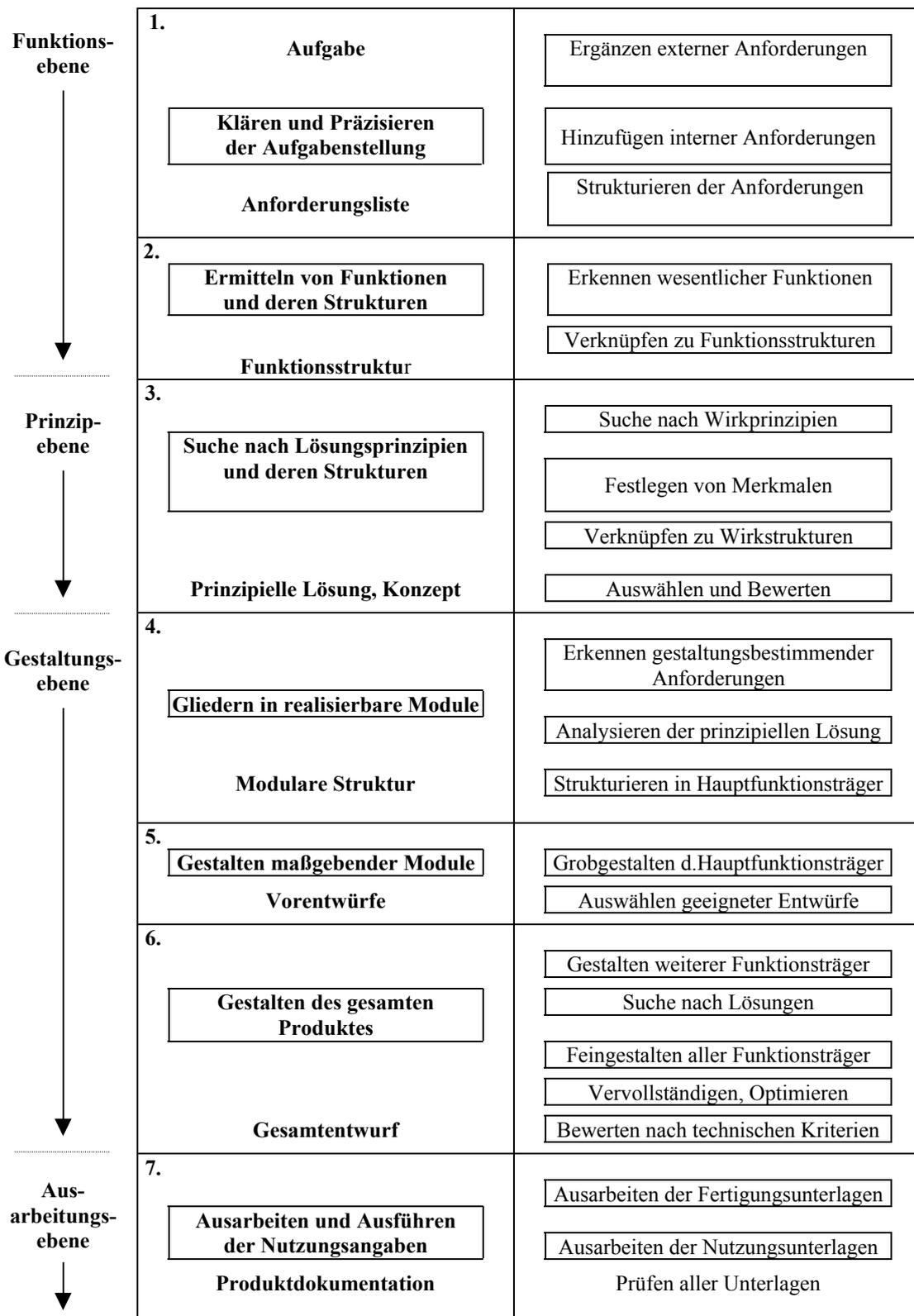


Abbildung 2.10. Phasen des Konstruktionsprozesses für eine Produktentwicklung nach VDI 2221 (vgl. Hütte 1991, Koller 1994, Pahl & Beitz 1997)

Im *ersten* Schritt der Produktentwicklung wird eine Idee für das Produkt entwickelt, die Aufgabenstellung geklärt und präzisiert. Die Anforderungen an das Produkt und deren Bedeutung werden in einer Anforderungsliste festgelegt.

Im *zweiten* Schritt werden die wesentlichen Funktionen herausgearbeitet, d.h. die Tätigkeiten, die das technische System leisten soll (z.B. Heben, Pumpen, Leiten, Steuern). Diese Funktionen werden in einer Funktionsstruktur zusammengeführt.

Im *dritten* Schritt erfolgt die Entwicklung einer Prinziplösung. Eine Prinziplösung braucht einmal die Festlegung auf die geeigneten *Wirkprinzipien/Gesetzmäßigkeiten* der Physik, Chemie, Biologie etc., die den Funktionen zugrunde liegen (z.B. das Impuls- bzw. Reibungswiderstandsgesetz, wenn mittels eines festen Körpers Zug- oder Druckkräfte übertragen werden sollen). Zum anderen braucht sie die Festlegung auf den *Effektträger*, d.h. den Werkstoff (wie Metall, Kunststoff, Holz, Flüssigkeiten, Gase etc.) mittels dessen die Funktion realisiert werden soll. Es lassen sich verschiedene Prinziplösungen entwickeln durch eine Variation von Wirkprinzip und Effektträger. Dies entspricht den Aussagen des Hempel-Oppenheimer-Schemas (siehe Abschnitt 2.1.1), das besagt, dass es für eine technische Anwendung zwei Prämissen bedarf: einer „Gesetzmäßigkeit“ sowie der spezifischen (Anfangs-, Rand-, Situations-)Bedingungen.

Im *vierten* und im *fünften* Schritt wird die Baustruktur entwickelt. Es werden Module entworfen und ausgewählt. Im *sechsten* Schritt wird der Gesamtentwurf festgelegt. Im *siebten* Schritt erfolgt schließlich die Festlegung auf die herstellungstechnischen Lösungen: die Form, die Bemessungen, die Oberflächenbeschaffenheiten aller Einzelteile, alle Werkstoffe, alle Herstellungsverfahren etc. Alle Schritte benötigen Analysephasen, vorläufige Entwürfe, Bewertungskriterien, Bewertung und Auswahl der Entwürfe. Ergebnis dieses Schrittes ist die Erstellung aller Fertigungsunterlagen und der gesamten Produktdokumentation.

Betrachtung: Besonders der dritte Schrittes ist im Kontext dieser Arbeit wichtig: Hier werden Theorien und Gesetzmäßigkeiten ausgewählt, die im Problemlösungsprozess eingesetzt werden. Lösungen werden in einem rationalen Vorgehen erarbeitet (vgl. Hütte 1991, Koller 1994 Pahl & Beitz 1997). „Es gäbe kein Getriebe, wenn es den Hebeleffekt nicht gäbe, und es gäbe keine Schraube, wenn es den Keileffekt nicht gäbe,

und es gäbe keine Federn, wenn es keine Elastizität gäbe u.a.m. ...Systematische Anwendung geeigneter physikalischer, chemischer und/oder biologischer Effekte zur Realisierung technischen Tuns ist ein wesentliches Hilfsmittel des methodischen Konstruierens von Prinziplösungen.“ (Koller 1994 S. 99). Zu ergänzen wäre dieser Satz, dass ebenso psychologische Gesetzmäßigkeiten angewendet werden müssten, damit die technischen Geräte auch die Bedürfnisse und Anforderungen der Menschen, die sie bedienen, erfüllen.

Mit diesem Ausflug in die Ingenieurwissenschaften zeichnet sich das Problem ab, in dem sich Psychologen und Ergonomen befinden: Im Gegensatz zu Ingenieuren tun sich diese Disziplinen schwerer, wenn es um Anwendung von Wirkprinzipien, Gesetzen und Theorien geht. Im Verhältnis zur Physik und den Ingenieurwissenschaften sind Psychologie und Ergonomie junge (Proto-)Wissenschaften (vgl. Sarris 1990). Die Vielzahl von experimentellen Erkenntnissen und Theorien, die z.B. in der Psychologie entwickelt wurden, wurden m. E. noch nicht auf „Brauchbarkeit und Effizienz“ für die Anwendung geprüft, bewertet und dementsprechend „bereitgestellt“. Die Ingenieurwissenschaften, insbesondere der Maschinenbau als angewandte Wissenschaft der Physik, haben eine sehr viel ältere wissenschaftliche Tradition als die angewandte Psychologie. Ihnen liegen physikalische Gesetzmäßigkeiten gekämmt, geschliffen und überprüft in Theorie und Praxiseinsatz vor, wie die Gesetze der Mechanik (z.B. Hebel-, Impuls-, Trägheits-, Fliehkraftgesetze), der Elektrizitätslehre (z.B. magnetische Induktions-, elektrostatische Feldgesetze), der Optik (z.B. Strahlensatz) oder der Wärmelehre (z.B. die thermodynamischen Gesetze) (vgl. Koller 1994, Pahl & Beitz 1997). Und sie haben bereits eine gewisse Tradition in einer Methodologie des Problemlösens.

2.5 Zusammenfassung

Dieses Kapitel klärte, wie das Anordnungsproblem methodisch zu lösen ist: in welchen Schritten die Problemlösung zu erarbeiten ist und welchen Kriterien sie folgen sollte.

Auf der Suche nach einer explizit dargestellten Methodologie wurden im *ersten* Abschnitt a) methodische Grundpositionen der Nutzerforschung, b) wichtige ISO-Normen

(ISO 13407, ISO 9241-10) und Literatur der Nutzerforschung (Shakel & Richardson 1991, Caroll 1991, Nielsen 1993, Green & Jordan 1999) sowie c) zwei wichtige Lehrbücher der Ergonomie (Schmidtke 1993, Luczak 1993) untersucht. Das Ergebnis der Suche war unbefriedigend: es fanden sich keine expliziten Darstellungen eines methodisches Vorgehen für einen wissenschaftlichen Problemlösungsprozess. Der Leser wird mit Routinen versorgt (in Form von Empfehlungen zum Gestaltungsprozess, Design-Guidelines und Methodenkatalogen), aber nicht mit einer Methodologie, in der die Lösungsschritte nach definierten Kriterien aufeinander abgestimmt sind.

Im *zweiten* Abschnitt wurde damit begonnen, ein explizites methodisches Vorgehen für den Problemlösungsprozess zu formulieren. Es erfolgte ein Ausflug in die Wissenschaftstheorie, um das Verhältnis von Theorie und Praxis zu klären. Es zeigte sich die Bedeutung von Gesetzmäßigkeiten: Ziel der Grundlagenforschung ist es – vereinfacht gesagt – nach Gesetzen zu suchen und Theorien zu formulieren, um Ereignisse zu erklären. Ziel der Anwendungsforschung ist es, bereits vorhandene Theorien und Gesetzmäßigkeiten anzuwenden, um Ereignisse entsprechend herzustellen. Aufgrund der strukturellen Gleichheit von Theorieentwicklung und Anwendungsentwicklung wurde eine Methodologie der Anwendungsentwicklung in Analogie zur Methodologie zur Theorieentwicklung entwickelt. Vorteil hierbei ist, dass in der Literatur bereits eine explizite Methodologie der Theorieentwicklung vorliegt (vgl. Bortz 1985, Sarris 1990).

Im *dritten* Abschnitt wurde dann ein methodisches Vorgehen für den Problemlösungsprozess entwickelt. Dieses Vorgehen fordert eine theoretische und empirische Fundierung des Anordnungsproblems: Es muss zuerst nach einer Theorie gesucht werden, die zum Anordnungsproblem passt und Bedingungen für erwartungsgerechte Anordnungen nennt. Die Gesetze der Theorie sind auf das Problem anzuwenden und es sind Prognosen im Sinne von Hypothesen zu bilden. Es ist ein entsprechendes Untersuchungsdesign zu entwickeln, um dann anhand einer (Input-)Studie empirisch fundierte Daten für eine nutzerorientierte Anordnung zu erhalten. In einer Evaluations-Studie können dann Gestaltungsvarianten überprüft und bewertet werden.

Im *vierten* Abschnitt erfolgte ein Ausflug in die Konstruktionsmethodik (ein Fachbereich des Maschinenbaus), um zu sehen wie dort praktische Problem gelöst werden. Es wurde gezeigt, dass dort ein theoriegeleitetes, wissenschaftliches Vorgehen unter dem

Begriff „Methodisches Konstruieren“ bereits Standard ist (vgl. Hütte 1991, Koller, 1994, Pahl & Beitz 1997, VDI Richtlinie 2221 1996).

Die Phasen des hier formulierten methodischen Vorgehens werden in den folgenden Kapiteln abgearbeitet. Das nächste Kapitel startet mit der ersten Phase, einer intensiven Analyse von Anordnungen und des Such- und Findeverhaltens von NutzerInnen, um Bedingungsfaktoren und erste Theoriebereiche für die Problemlösung zu identifizieren.

3. Problemanalyse und Bestimmung möglicher Theoriebereiche

Was tun Personen, wenn sie etwas suchen? Was ist eine Anordnung? Und was sind mögliche Bedingungsfaktoren, damit Anordnung und Suchverhalten zusammenpassen? Im folgenden Kapitel findet die Problemanalyse statt. Ziel ist es, erste Theorien für eine Problemlösung zu identifizieren. Das Kapitel gliedert sich in fünf Abschnitte: Im *ersten* Abschnitt werden zunächst die aktuellen Strategien von Gestaltern dargestellt, um Nutzer in ihren Suchprozessen zu unterstützen. Und es wird anhand von Ergebnissen in der Literatur gezeigt, welche Probleme diese Strategien mit sich bringen. Im *zweiten* Abschnitt erfolgt eine systematische Auseinandersetzung mit dem Problemgegenstand. Es wird untersucht, was Anordnungen definiert, wie sich Anordnungen formal beschreiben lassen und welche Form der Beschreibung sich für die Anordnungsprobleme dieser Arbeit eignet. Im *dritten* Abschnitt werden die Ergebnisse einer empirischen Explorationsstudie vorgestellt: Personen wurden gefragt, was sie „tun“, wenn sie nach einem Gegenstand suchen. Die Explorationsstudie hat das Ziel, das Suchverhalten von NutzerInnen nach psychologischen Bedingungsfaktoren zu analysieren. Im *vierten* Abschnitt wird die Analyse fortgesetzt mit der Suche nach einer Problemlösung versprechenden Theoriebereich. Das Kapitel endet im *fünften* Abschnitt mit einer Zusammenfassung der wesentlichen Ergebnisse der Problemanalyse in Form von Schlussfolgerungen und Annahmen, und es erfolgt das erste Skizzieren eines inhaltlichen Lösungsweges. Die Abschnitte bauen systematisch aufeinander auf.

3.1 Aktuelle Strategien, um Suchprozesse zu erleichtern

Um Orientierungs- und Suchprobleme zu minimieren, verwenden Ladenplaner und Architekten üblicherweise Übersichtspläne und Leitsysteme. So finden sich in jedem Flughafen, in jedem Einkaufszentrum und in größeren Lebensmittelmärkten Übersichtspläne und Hinweisschilder. Eigentlich sollen sie NutzerInnen helfen, sich gut zurechtzufinden. Doch Untersuchungen zeigten, dass sie selbst Probleme erzeugen und das Suchproblem nicht befriedigend lösen helfen. So gaben in einer Studie über den Flughafen Fort-Worth Dallas in Texas 30% der Befragten an, dass es zu viele Schilder gäbe, und 76% hatten Schwierigkeiten, die Wegweiser zu verstehen. (Seidel 1983).

Die Probleme, die Übersichtspläne und Leitsysteme mit sich bringen, veranlassten zahlreiche Untersuchungen (vgl. Appleyard 1977; Levine et al. 1984; Warren 1994). Es wurden z.B. folgende Faktoren für eine bessere Verständlichkeit von Übersichtsplänen und Leitsystemen erarbeitet: Die Pläne und Schilder müssen prominent und sichtbar im Raum angebracht werden. Übersichtspläne sind entsprechend dem Blickfeld des Betrachters zur direkten Übertragung der Karte auf die realen Gegebenheiten anzubringen. Die Symbole und Hinweiszeichen müssen hoch verständlich sein. Die Übersichtspläne und Schilder müssen in ausreichender Häufigkeit und Rhythmik vorhanden sein. Die Lesbarkeit im visuellen Blickfeld, wie Kontrast, Größe, Höhe, Ausleuchtung, muss sehr gut sein. Mit Leitsystemen können also viele „Fehler“ gemacht werden.

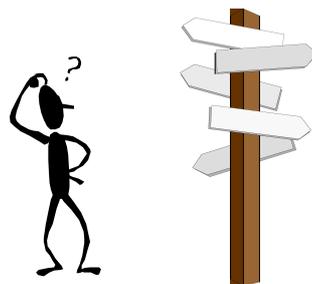


Abbildung 3.1. Der aktuelle Lösungsweg, um Suchprobleme zu erleichtern

Weisman (1981) konnte darüber hinaus in einer Studie zeigen, dass sowieso nur 18% der Versuchspersonen Übersichtspläne nutzten. Und die Untersuchungen von Butler, Acquino, Hisson, Scott (1993) weisen darauf hin, dass man bei komplizierten, architektonischen Grundrissen lieber ganz auf Übersichtskarten verzichten sollte, weil sie zu viel Zeit und zu hohe Gedächtnisanstrengung bedeuten. Die Ergebnisse legen die Vermutung nahe, dass die Suchprobleme durch optische Hinweisreize alleine nicht zu lösen sind (s. Abb. 3.1). Und es wird angenommen, dass das Problem tiefer liegt. Dies bestätigt den Ansatz, die Anordnung der Objekte im Raum näher zu betrachten. Mit diesen Aussagen wird nicht behauptet, dass man auf Übersichtspläne und Leitsysteme gänzlich verzichten sollte. Nein, sie sind wichtig, aber ihnen wird eine andere Funktion zugeschrieben: Übersichtspläne und Leitsysteme bestärken meines Erachtens den suchenden Nutzer im Sinne eines Reinforcements (Verstärkung) in seinen Orientierungserwartungen, dass er auf dem richtigen Weg ist.

3.2 Formale Analyse von Anordnungen

Was definiert eine Anordnung von Objekten im Raum, wie von Süßgebäck im Regal oder von Geschäften in einem Einkaufszentrum? Und welche Beschreibungsform eignet sich, um nutzerorientierte Anordnungen von solchen Objekten zu beschreiben? Diese Fragen werden im Folgenden untersucht.

Definitionen Spricht man davon, dass Objekte angeordnet sind, beinhaltet das, dass die Objekte nicht zufällig, chaotisch im Raum verteilt sind, sondern dass die Verteilung einer Ordnung folgt. Diese Ordnung definiert sich durch Regeln, die die Beziehungen zwischen den Objekten bestimmen. Es wird angenommen, dass eine Anordnung dann „nutzerorientiert“ ist, wenn die Objekte nach Regeln angeordnet werden, die den Nutzern entsprechen. In dieser Arbeit soll der Fokus ausschließlich auf kognitive Aspekte gelegt werden. Körperliche, anthropometrische Aspekte sind bereits ausführlich von der klassischen Ergonomie untersucht (vgl. Grandjean, 1979, Kirchner 1990). (Im Handbuch für „Räumlich-ergonomische Gestaltung“ (Schriftenreihe der Bundesanstalt für Arbeitsschutz Kirchner 1990) wird z.B. eine Optimierung hinsichtlich kognitiver Aspekte gar nicht erwähnt.) Ob eine Anordnung körperlichen, anthropometrischen Anforderungen entspricht, müsste also gesondert untersucht werden.

Beschreibungsformen von Anordnungen Nun zur Frage, wie man Anordnungen formal beschreiben kann. Die Topographie und die Topologie beschäftigen sich mit den formal-strukturellen Kriterien von Anordnungen. Die Topographie ist die allgemeine Lehre der Beschreibung von Orten, Geländen, Landschaften. Die Topologie ist eine sehr junge mathematische Disziplin, die sich mit Gruppen beschäftigt. Entsprechend diesen Disziplinen gibt es vier Beschreibungsformen: die *metrische*, *perspektivische*, *topologische* und die *relationale* Beschreibung. Es soll nun untersucht werden, welche dieser vier Beschreibungsformen sich am besten für das Problem dieser Arbeit eignet.

Bei einer *metrischen* Beschreibung werden Objekte im Raum durch exakte Distanzangaben zwischen den Objekten (z.B. in Meter oder Zentimeter) definiert. Dies ist eine äußerst präzise, aber auch starre Beschreibungsform. Angenommen, es lässt sich eine nutzerorientierte Regalanordnung für Süßgebäck finden, dann erscheint es nicht not-

wendig, diese in metrischen Distanzangaben zu definieren. Eine Fachkraft, die die Regalbelegungspläne entwickelt, könnte damit auch nur bedingt etwas anfangen. Denn die Länge des Süßgebäckregals ist abhängig von der Supermarktgröße. Der Regalbeleger bräuchte eine Vorgabe für eine nutzerorientierte Süßgebäckanordnung, die er auf unterschiedlichste Regalgrößen umsetzen kann. Die Frage nach den metrischen Distanzen zwischen den Objekten berührt anthropometrische und biomechanische Probleme, also Probleme der körperlichen Erreichbarkeit. Das ist eine wichtige Frage, doch sie soll nicht Gegenstand dieser Arbeit sein. Hier soll es nicht um anthropometrische Aspekte gehen! Eine metrische Beschreibung ist zunächst nicht die richtige Beschreibungsform für Anordnungen von Objekten im Raum.

Eine *perspektivische* Beschreibung hat einen fixen Blick-/Standpunkt. Die Objekte sind dadurch eindeutig in ihrem räumlichen Bezug definiert als: vor/hinter, links/rechts, oben/unten definiert. Perspektivische Beschreibungen werden meist in Untersuchungen über das räumliche Gedächtnis zugrunde gelegt (vgl. Knauff 1997, Herrmann & Schweizer 1998). Sie werden z.B. in der Cockpit-Forschung verwendet. In der Tat hat hier der Nutzer in Bezug zu den Cockpit-Instrumenten immer die gleiche Perspektive. Für die Anwendungsfälle dieser Arbeit ist eine perspektivische Beschreibung jedoch ungeeignet! Die Nutzer von Supermärkten, Einkaufszentren, Flughäfen bewegen sich im Raum. Und das, was vorher links vor einer Person liegt, kann beim Laufen dann später rechts hinter ihr liegen. Eine perspektivische Beschreibungsform macht also keinen Sinn!

Bei einer *topologischen* Beschreibung wird die Anordnung der Objekte nur noch durch Nachbarschaften oder Gruppen definiert. Sie zielt also nur auf den Gruppierungs- bzw. Nachbarschaftserhalt. Die absoluten Distanzen sind nicht relevant und sie berücksichtigen auch nicht die relative Lage von Objekten zueinander. Eine topologische Anordnung von Flughafeneinrichtungen beschreibt, in welcher Gruppe von Einrichtungen z.B. die Ticketschalter bzw. der Meeting-Point liegen, und ob die Gruppen z.B. benachbart sind und welche anderen Gruppen benachbart angrenzen. In Bezug auf die Problemstellungen dieser Arbeit ist eine topologische Beschreibungsform gut geeignet.

Die *relationale* Beschreibung von Objektanordnungen zielt auf die relative Lage der Objekte zueinander. Relationalen Anordnungen liegen Ähnlichkeits- bzw. Klassifizierungskriterien und Distanzkriterien zugrunde, die die relative Lage zueinander bestimmen. Dabei muss die Anordnung nicht durchgängig von einem Kriterium, sondern kann durch unterschiedliche Kriterien bestimmt sein. Eine relationale Anordnung von Flughafeneinrichtungen beschreibt, welche Ähnlichkeit z.B. die Ticketschalter in Bezug zum Meeting-Point oder zum Abflug-Gate haben. Relationale Anordnungen eignen sich auch als Beschreibungsform für das Anordnungsproblem.

Die Auseinandersetzung mit formalen Aspekten von Anordnungen erbringt folgende *Schlussfolgerungen* in Hinblick auf eine Problemlösung: Um Objekte nutzerorientiert anzuordnen, erscheint es notwendig, die Objekte nach den Kriterien der Nutzer anzuordnen. Im Rahmen dieser Arbeit wird untersucht, welche Objekte bzw. Objektgruppen aus Nutzersicht nebeneinander oder gruppiert liegen sollten und/oder in welchen Relationen die Objekte zueinander liegen sollten. Auf eine metrische oder perspektivische Beschreibungsform für eine nutzerorientierte Anordnung wird in dieser Arbeit verzichtet und ist in einer gesonderten anthropometrischen/biomechanischen Studie zu überprüfen.

3.3 Ergebnisse einer Explorationsstudie zum Suchverhalten

Wie gehen Personen vor, wenn sie etwas suchen? Diese Frage wurde in einer empirischen Explorationsstudie untersucht. Die Ergebnisse werden hier dargestellt. Ziel dieser Explorationsstudie ist es, Bedingungsfaktoren herauszufinden, die dafür verantwortlich sind, dass Nutzer schnell das finden, was sie suchen. Die Nutzer wurden nicht befragt, wie sie sich eine Anordnung wünschen, da Nutzer m.E. dann das beschreiben, was sie bereits kennen. Wie bei Erkundungsstudien üblich (vgl. Bortz 1984), wurde nur eine kleine Versuchspersonengruppe mit qualitativen Methoden untersucht. Die Ergebnisse haben keinen Anspruch auf Repräsentativität, was bei einer Studie dieser Art auch nicht notwendig ist (vgl. Bortz 1984, Virzi 1990, 1992 und Kanis & Ariz 2000). Acht Nutzer (vier Frauen und vier Männer im Alter von 20-70 Jahren) wurden gebeten zu beschrei-

ben, was sie tun, wenn sie nach einem Objekt suchen. Die Befragungen waren nondirektiv: es gab also eine Einstiegsfrage, die weiteren Fragen wurden aus den Äußerungen der befragten Personen gebildet, durch Rückformulierungen, Zusammenfassungen etc. Die Einstiegsfragen lauteten: „Wo suchen Sie in einem Supermarkt Spaghetti?“ oder „Wo suchen Sie in einem Flughafen eine Toilette?“ Die Auswertung erfolgte durch eine qualitative Inhaltsanalyse, mit dem Ziel, ein möglichst breites Spektrum an Aussagen über Suchverhaltensweisen zu erhalten. Die Interviews dauerten ca. 15 min.

Die Ergebnisse zeigen fünf Aspekte, die das Suchverhalten bestimmen. Sie werden im Folgenden dargestellt (siehe Abb. 3.2) und durch Originalzitate belegt:



Abbildung 3.2. Ergebnis der Explorationsstudie: fünf Aspekte, die das Suchverhalten bestimmen

1. *Planungsverhalten*: Die Suche wird durch die verfügbare Zeit und das Ziel bestimmt. Die befragten Personen berichten, dass eine klare Vorstellung, was sie wollen und Zeitdruck zu einem überlegtem Vorgehen führen. Eine vage Zielvorstellung und viel verfügbare Zeit verführt sie dazu, sich treiben zu lassen.

- „...entweder ich will was ganz Bestimmtes, oder ich schlendere durch und schau, was mich anspricht!“

2. *Denkverhalten*: Das Suchverhalten wird durch Erwartungen über den Ort bestimmt. Die befragten Personen äußern, dass eine „erste“ Erwartung, wo das gesuchte Objekt zu finden ist, quasi automatisch parat steht.

- „...Es ist mir nicht bewusst, aber ich weiß, wo ich hingehen muss ...“

- „...ich hab' ein klassisches Bild von einem Supermarkt im Kopf, aber es ist nicht der, wo ich immer einkaufe ... das ist gar kein echter, ich weiß, das klingt unlogisch ...“
- „Zuerst geh ich dorthin, wo ich es vermutet habe, wenn es da nicht ist, werde ich richtig in meinem Lauf gebremst ... dann geht der bewusste Denkprozess los.“

3. *Problemlöseverhalten*: Findet die Person das gesuchte Objekt an dem erwarteten Ort nicht, muss sie bewusst überlegen oder fragt eine andere Person.

- „Wenn ich beim ersten Mal das dann nicht finde, wird's schwierig!“
- „Ich überlege mir, wo das sein könnte; ich grabe in meinem Gehirn ...“
- „Wenn es da nicht ist, wo ich es vermute, bin ich sauer, dann renn ich und frag mürrisch `ne Verkäuferin!“

4. *Wahrnehmungsverhalten*: Die Personen berichten, dass sie das, was sie sehen, sofort auswerten, dass aber auch die Erwartung darüber, wie das gesuchte Objekte aussieht, ihr Verhalten steuert.

- „... ich stelle mich hin und mache mir einen Überblick, wo es sein könnte!“
- „... man erkennt es an den Farben, an der Packungsform, an den blau-gelben Pappkartons mit dem Sichtfenster ...“
- „Ich schaue in das Regal und sehe Marmeladen, da weiß ich, da sind Spaghetti nicht benachbart... wo Reis ist, da such ich dann...“

5. *Bewegungsverhalten*: Das Suchverhalten wird durch zielgerichtetes oder planloses Umhergehen bestimmt. Man hofft, das Gesuchte schon „irgendwie“ anzutreffen, oder bewegt sich zielgerichtet dorthin, wo man das Objekt erwartet.

- „... ich laufe auf dem vorgegebenen Weg und schaue links und rechts ...“
- „... Spaghetti sind in der Mitte, das ist immer so!“

Um NutzerInnen das Finden von Objekten zu erleichtern, ergeben sich nun drei Ansatzpunkte: 1) Eine den Erwartungen der NutzerInnen entsprechende Gruppen- und Nachbarschaftsbildung und relative Anordnung der Objekte im Raum bilden. 2) Visuelle Hinweisreize liefern, z.B. durch identifizierungsrelevante Positionierung der Objekte, Ausleuchtung, durch Leitsysteme etc. 3) Eine nutzerorientierte Wegeführung gestalten, z.B. durch die Lage der zentralen Wege, der Nebenwege, Stichwege, Rundwege etc. (vgl. Abb. 3.3).

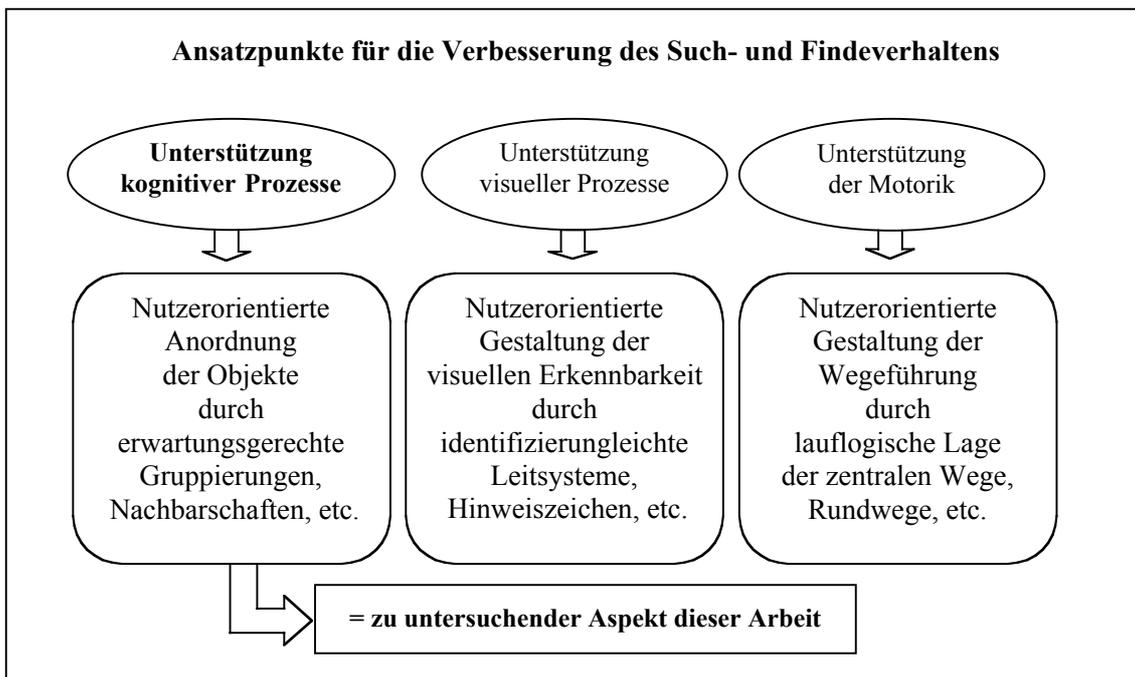


Abbildung 3.3. Ansatzpunkte für die Verbesserung des Such- und Findeverhaltens von Nutzern

Schlussfolgerung aus den Ergebnissen der Explorationsstudie: In dieser Arbeit soll der erste Punkt weiter untersucht werden: die Anordnung der Objekte im Raum, die mit den Erwartungen der Nutzer korrespondiert. Dabei werden unter Nutzern jene Personen verstanden, die die Objekte normalerweise nutzen und nun innerhalb der Objektmenge etwas suchen. Die beiden Optimierungsmöglichkeiten „visuelle Gestaltung“ und „Wegeführung“ werden hier nicht weiter verfolgt, auch wenn sie hinsichtlich einer nutzerorientierten Gestaltung für wichtig gehalten werden.

Die Aussagen der befragten Personen zeigen, dass es anscheinend einen Unterschied gibt zwischen Erwartungen, die spontan entstehen und solchen, die erst durch Überlegungen gebildet werden müssen. *In dieser Arbeit wird davon ausgegangen, dass eine Anordnung von Objekten im Raum dann nutzerorientiert ist, wenn sie den spontanen, intuitiven Erwartungen der Nutzer entspricht.* Die Befragung zeigte, dass es Nutzer als weniger angenehm erleben, wenn sie „gezwungen“ sind, bewusst zu überlegen, wo was sein könnte. Im Weiteren wird untersucht, was aus kognitionspsychologischer Sicht unter „spontanen“ Erwartungen zu verstehen ist und was den Inhalt solcher Erwartungen beeinflusst.

3.4 Erste Zuordnung zu möglichen Theoriebereichen

Die Ergebnisse weisen auf die Bedeutung der Erwartungen der Suchenden hin, wo sie was finden. Erwartungen sind Schlussfolgerungen. Deshalb wird zuerst untersucht, was die Inferenzforschung (Lehre von den Schlussfolgerungen, einem Teilgebiet der kognitiven Psychologie) über Erwartungen beitragen kann. Die Antwort auf die Frage: Wovon hängen Erwartungen ab? könnte der Schlüssel dazu sein, die Objekte für die Suchenden erwartungsgerecht anzuordnen.

3.4.1 Erwartungen in der Inferenzforschung

Innerhalb der aktuellen Kognitionspsychologie beschäftigt sich die Inferenzforschung und die Problemlösungsforschung (vgl. Johnson-Laird 1983, Engelkamp 1990, Nisbett 1993, van der Meer 1998) mit den Regeln des menschlichen Schlussfolgerns. Van der Meer (1998) definiert Inferenzen (inferre lat.: hineinragen, folgern) als Ergebnis kognitiver Schlussfolgerungsprozesse, die auf der Grundlage von bestehendem Wissensbesitz bzw. aktuell gegebener Wahrnehmungsinhalte neue Informationen erzeugen. Diese neu erschlossenen Informationen können dabei als neues Wissen, Erwartungen, Einstellungen, Urteile, Hypothesen, Problemlösungen etc. wirken. Das entstehende Resultat wird durch die Auswahl und Manipulation von Informationen sowie durch die Qualität der verarbeiteten Informationen bestimmt.

In der Forschung gibt es unterschiedliche Ansätze, Inferenzen in Abhängigkeit von ihrer Spezifik zu ordnen. Die folgende Darstellung ist angelehnt an die Inferenztaxonomie von Elke van der Meer (1998). Sie unterscheidet klassische, prozessorientierte und inhaltsabhängige Taxonomien. Diese Taxonomien haben beschreibenden, aber keinen erklärenden bzw. prognostischen Charakter. Sie sind auch nicht trennscharf, sondern überlappend. Ziel ist es, jenen Inferenztypus zu identifizieren, der zu Erwartungen bezüglich des Ortes von Objekten passt.

1. *Die klassische Taxonomie* fußt auf formal-logischen Kriterien und unterscheidet z.B. *induktive Inferenzen*, d.h. Schlussfolgerungen, bei denen – im Sinne eines Wahr-

scheinlichkeitsschlusses – von einer Teilmenge von Beobachtungen auf eine Gesamtmenge geschlossen wird (z.B. von einer beobachteten Menge an weißen Schwänen wird geschlossen, dass alle Schwäne weiß sind.). Beim induktiven Schließen wirken Generalisierungs- und Diskriminationsstrategien.

Davon unterschieden werden *deduktive Inferenzen*, d.h. Schlussfolgerungen, bei denen aus allgemeinen Gesetzen oder Prämissen Aussagen abgeleitet werden. Als Beispiel dient hier der Syllogismus: *Alle Menschen sind sterblich. Sokrates ist ein Mensch. Daher gilt: Sokrates ist sterblich.*“ Ein Syllogismus wird formal durch zwei Prämissen und einen Schluss dargestellt, der formal-logischen Regeln folgt, um einen gültigen, „wahren“ Schluss zu ziehen. Menschen neigen jedoch – wie allgemein bekannt ist – dazu, Schlüsse nicht automatisch nach formal-logischen Regeln zu ziehen. Vorwissen, Voreinstellungen, Kontexteffekte etc. (vgl. Johnson-Laird 1989, van der Meer 1998) beeinflussen Schlussfolgerungen.

Eine weitere Auseinandersetzung mit Erkenntnissen über induktive bzw. deduktive Schlussfolgerungen wird hinsichtlich der Fragestellung, wo sich ein Objekt befindet, als nicht zielführend betrachtet, denn sie thematisieren ausschließlich formal-logische Kriterien des Schließens und lassen z.B. das Vorwissen, auf dessen Basis die Schlüsse gezogen werden, außer Acht. Dies erscheint aber bei den vorliegenden Schlussprozessen als unabdingbar. Denn Schlüsse, wo ein Objekt zu finden ist, werden aller Voraussicht nach auch auf der Basis von Wissensinhalten gezogen.

2. *Inhaltsabhängige Taxonomien* betonen die Bedeutung des Vorwissens für die Schlussfolgerungen. Beispiele für inhaltsabhängige Inferenzen sind z.B. Analogien, wie *„Füller verhält sich zu Schulheft wie Kreide zu Tafel.“* *„Bäcker verhält sich zu Lebensmittelgeschäften wie Boutique zu Bekleidungsgeschäften.“* Solche Analogien kann nur jemand ziehen bzw. beurteilen, der spezifisches Wissen über diese Objekte hat (vgl. Gentner 1989). Inhaltsabhängige Schlussfolgerungen greifen an spezifischem Wissensbesitz an. Bereits Minsky (1975) und Schank & Abelson (1977) untersuchten den Einfluss von im Gedächtnis gespeicherten Schemata über Raumszenen und Ereignisse auf die Bildung von Erwartungen. Sie stellten fest, dass jeder, der hinter Bäumen und Sträuchern ein Dach sieht, erwartet, dass unter dem Dach auch ein Haus ist. Dies kann nicht nur aufgrund eines formal-logischen Schlusses gefolgert werden, sondern dies benötigt ein Vorwissen über Dächer und Häuser. Insofern

ist der Satz einzuschränken: nicht „jeder“, sondern nur Kenner von Häusern erwarten voraussichtlich, dass hinter den Bäumen auch ein Haus ist.

Bei den Sätzen: „Martin und Nanni waren im Supermarkt. Sie kauften Salat und Nudeln.“ schlussfolgert man spontan, dass sie den Salat in der Gemüseabteilung und die Nudeln im Nudelregal fanden, dass sie an der Kasse zahlten und dass sie sich wahrscheinlich Nudeln und Salat kochen werden. Es wird also auch umfangreicheres ereignisspezifisches Wissen mit aktiviert.

Die Ergebnisse von inhaltsabhängigen Inferenzen hängen a) vom Inhalt der Wissensstrukturen ab sowie b) von den Regeln und Gesetzmäßigkeiten, nach denen das Wissen aufgenommen, gespeichert und verarbeitet wird (vgl. Tversky 1981 zitiert nach Opwis & Lüer 1998, Strube 1984, Klix 1988, Dörner & van der Meer 1995, van der Meer 1998). Aufnahme, Speicherung und Verarbeitung von Wissen folgt allgemeinen, überindividuellen Regeln und Gesetzmäßigkeiten. Um inhaltsabhängige Inferenzen wie z.B. Erwartungen über den Ort eines gesuchten Objektes verstehen und vorhersagen zu wollen, bedarf es also sowohl der Kenntnis der Wissensinhalte als auch der Kenntnis der Regeln und Gesetzmäßigkeiten, nach denen das Wissen verarbeitet wurde und aktiviert wird.

3. *Prozessorientierte Taxonomien* thematisieren den Schlussprozess und unterscheiden:
- *automatische versus kontrollierte Inferenzen*, d.h. Schlussfolgerungen, die Informationen durch bloße Aktivierung des Wissensbesitzes bereitstellen bis hin zur Erzeugung von Informationen durch kontrollierte Prozesse, die über Denk-, Gedächtnis- und Wahrnehmungsstrukturen ablaufen,
 - *zeitlich rückwärts- und vorwärtsgerichtete Inferenzen*, d.h. Schlussfolgerungen, die zur Ursachenerklärung bzw. zur Vorhersage oder Antizipation neuer Informationen aus dem bisher verfügbaren Gedächtnisbesitz führen,
 - *elaborative versus reduktive Inferenzen*, d.h. Schlussfolgerungen, die zur Erweiterung oder aber zur Einschränkung relevanter Informationen führen.

Im Rahmen dieser Problemstellung sind vor allem *automatische* Inferenzen von Interesse. Die Unterscheidung von automatischen und kontrollierten Inferenzen (vgl. Engelkamp 1990, Klix 1988) spiegelt sich auch in den Ergebnissen der Explorati-

onsstudie wider. Die befragten Personen äußerten dort sehr deutlich einen Unterschied zwischen „spontanen“ Erwartungen und bewussten, „mühsam“ überlegten Erwartungen. Folgendes Zitat belegt dies: „Zuerst geh ich dorthin, wo ich es vermutet habe, wenn es da nicht ist, werde ich richtig in meinem Lauf gebremst ... dann geht der bewusste Denkprozess los.“ Spontane Erwartungen hinsichtlich des Ortes eines Objektes könnten demnach als automatische Inferenzen betrachtet werden.

Zusammenfassung Erwartungen, wo Objekte zu finden sind, werden im Kontext dieser Problemstellung als inhaltsabhängige, spontane Schlussfolgerungen definiert. Inhaltsabhängig deshalb, weil sie aus dem bestehendem Wissensbesitz abgeleitet werden. Spontan, weil sie automatisch durch Aktivierung des Wissensbesitzes bereitgestellt werden und nicht durch bewusst überlegte und kontrollierte Denkprozesse gebildet werden. Konkret für die Problemstellung dieser Arbeit heißt dies, dass eine spontane Erwartung bezüglich des Ortes eines Objektes im Raum davon abhängt, welches Vorwissen eine Person hat. Das im Gedächtnis gespeicherte Vorwissen ist dabei selbst ein Ergebnis der Erfahrungen sowie der Regeln und Gesetzmäßigkeiten, nach denen diese Erfahrungen im Gedächtnis gespeichert und verarbeitet wurden. Der Inhalt einer spontanen Erwartung ist demnach ein Ergebnis a) der Vorerfahrung sowie b) der Regeln und Gesetzmäßigkeiten, nach denen dieses Wissen abgespeichert, verarbeitet und wieder abgerufen wird.

Für die Problemstellung dieser Arbeit führt dies zu folgenden *Schlussfolgerungen*: im Hinblick auf eine benutzerorientierte, intuitive Anordnung von Objekten im Raum wird angenommen, dass eine Anordnung, die mit dem Wissensbesitz der Nutzer korrespondiert, Such- und Findeprozesse erleichtert. Eine Kenntnis a) des Wissens, das der Suchende hat, sowie eine Kenntnis b) der Regeln und Gesetzmäßigkeiten, nach denen dieser Wissensbesitz gespeichert ist, müsste treffsichere Prognosen liefern, welchen Kriterien eine nutzerorientierte Anordnung folgen sollte. Es wird daher im Folgenden untersucht, was aus kognitions-psychologischer Sicht unter Wissensbesitz verstanden wird und welche Art des Wissens sich hinsichtlich des Anordnungsproblems eignet, um dann weiter nach den Regeln und Gesetzmäßigkeiten zu suchen, nach denen dieses Wissen organisiert ist.

3.4.2 Prüfung verschiedener Formen der Wissensrepräsentation

Welche verschiedenen Formen der Wissensrepräsentation werden heutzutage angenommen? Und welche Wissensrepräsentation kann als Basis für die Bildung von Erwartungen für den Ort eines Objektes angenommen werden? Dies wird im Folgenden untersucht.

Die aktuelle kognitionspsychologische Forschung unterscheidet derzeit vier Wissensarten: *episodisches*, *prozedurales*, *anschaulich-analoges* und *begrifflich-semanticisches Wissen* (vgl. Opwis & Lüer 1998, Städler 1998; siehe Abb. 3.4).

Episodisches Wissen wird als autobiographisches Wissen einer Person definiert, das sich auf raum-zeitliche, einmalige Ereignisse bezieht. Das Wissen einer Person, wo sie z.B. in „ihrem“ Supermarkt, den sie normalerweise immer aufsucht, Butterkekse findet, stellt nach dieser Definition episodisches Wissen dar. Tulving definierte 1972 diese Wissensart, um es dem begrifflich-semanticischen Wissen gegenüberzustellen („Wo finden sich in Supermärkten generell Butterkekse?“). Eine Auseinandersetzung mit Theorien über episodisches Wissen wird als nicht zielführend für die Problemlösung angesehen, weil das Untersuchungsziel lautet, nutzerorientierte Anordnungen von Objekten im Raum für eine breite Nutzergruppe zu entwickeln. Die Untersuchung episodischen Wissens einer Person eignet sich u.U., um für eine einzelne Person eine maßgeschneiderte, mit ihren persönlichen Erfahrungen übereinstimmende Anordnung zu entwickeln.

Prozedurales Wissen wird als Handlungswissen definiert und bezieht sich auf die Art und Weise wie Handlungen durchgeführt werden (vgl. Hacker 1986). Problematisch ist dabei die unscharfe Definition dieser Wissensform: sie reicht von basalen, nicht reflektierbaren, kognitiven Prozeduren (wie Merkmalsvergleichsprozeduren) über sensu-motorisches Wissen (z.B. Reflexen oder automatisierten Bewegungen) bis hin zu komplexem, bewussteinfähigem Fertigkeitwissen (wie einer Maschinenbedienung). Wie schwierig eine Definition auch sein mag, hinsichtlich der Problemstellung dieser Arbeit eignet sich prozedurales Wissen nicht. Eine Untersuchung prozeduralen Wissens könnte allenfalls Erkenntnisse über Problemlösungs- und Handlungsstrategien von Personen in Suchsituationen liefern. Dies erscheint dann interessant, wenn man davon ausgeht, dass

die Findeleistung primär von einer effizienten Suchstrategie abhängt und man bei der Person ansetzen will, z.B. durch ein Training.

Äußerungen der befragten Personen in der Explorationsstudie deuteten – wie bereits dargestellt – darauf hin, dass die Person zunächst eine erste Erwartung quasi automatisch parat hat. Es wird daher angenommen, dass es spontane, automatisierte Erwartungen darüber gibt, wo etwas zu finden ist und dass komplexere kognitive Prozeduren im Sinne von Problemlösungsstrategien einsetzen, wenn die Person das Objekt nicht dort findet, wo sie es aufgrund der spontanen, intuitiven Erwartung nicht findet. Diese Annahme wird gestützt durch Befunde zu „automatischen“ bzw. autochtonen Inferenzen, die durch automatische Aktivierung verfügbarer Wissensstrukturen Informationen bereitstellen (van der Meer 1998). Ziel ist es, eine Anordnung zu entwickeln, die der ersten, spontanen Erwartung entspricht.

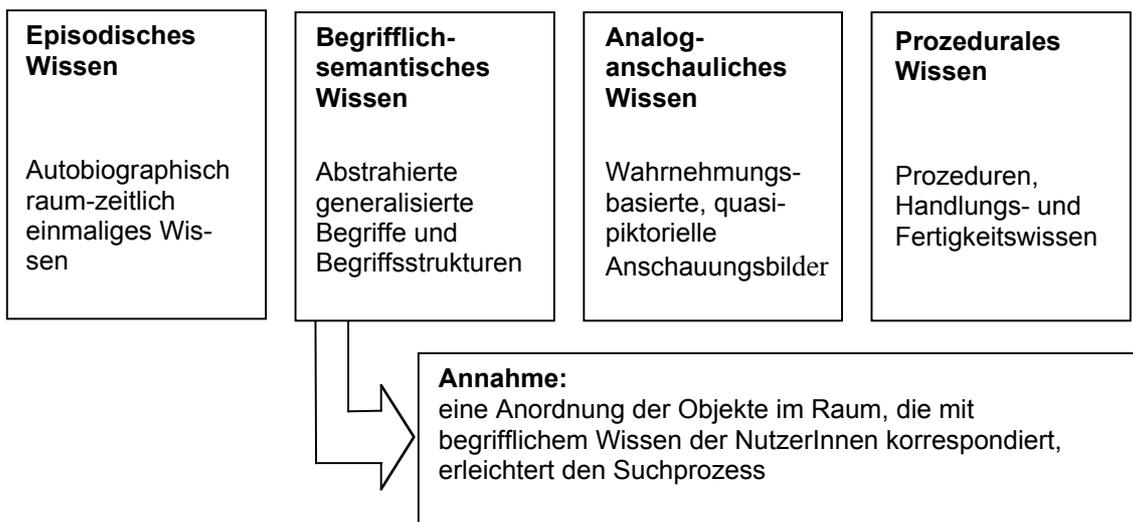


Abbildung 3.4. Eignung der Wissensarten zur Problemlösung (vgl. Opwis & Lürer 1998, Städtler 1998)

Anschaulich-analoges Wissen wird auch als wahrnehmungsbasiertes oder quasi-piktoriales Wissen definiert, das in Form von Vorstellungsbildern (Cooper & Shepard 1985, Engelkamp 1991, Logie 1995; zitiert nach Knauff 1997) und mentalen Modellen (Johnson-Laird 1983) repräsentiert ist. Die Fokussierung auf anschaulich-analoge Vorstellungen scheint zunächst nahe liegend. Kognitive Prozesse, die sich auf den Raum beziehen (Raumwahrnehmung, Raumorientierung) werden auch gerne im Rahmen die-

ses Paradigmas untersucht (vgl. Baddley 1986, Engelkamp 1991 zitiert nach Knauff 1997). Doch es gibt eine Vielzahl von Untersuchungen, die auch für eine begrifflich-semantische Repräsentation räumlicher Verhältnisse sprechen. Als Beleg ist z.B. der „Barriere-Effekt“ (Cohen, Baldwin & Sherman 1978) zu nennen: Versuchspersonen sollten die Entfernung zwischen ihnen bekannten Orten subjektiv schätzen. Dabei zeigte sich, dass nicht die euklidische Distanz in Metern, sondern der Aufwand, der für die Zielerreichung notwendig ist (z.B. verursacht durch einen Berg oder ein Gebäude als Barriere), einen Einfluss auf die subjektive Entfernung hat.

Von der Einzelsituation abstrahiertes, verallgemeinertes räumlich-konzeptuelles Wissen wird nicht mehr nur als anschaulich-analog sondern auch als begrifflich-semantisch repräsentiert angenommen (vgl. Minsky 1975, Klix 1984, Anderson 1996, Knauff 1997). Vorstellungsbilder und mentale Modelle sind dann nicht nur analog repräsentiert, sondern auch auf begrifflich repräsentiertes Wissen rückführbar. Vorstellungsbilder werden dann quasi bei Bedarf im Arbeitsspeicher (re-)konstruiert. Die aktuelle Diskussion geht derzeit dahin, dass sich anschaulich-analoges Vorstellungswissen sowie Mentale Landkarten auf spezifische, raumzeitlich einmalige Ereignisse beziehen. Es wird damit genauso beurteilt wie episodisches Wissen; diese Wissensart wird in Bezug auf das Anordnungsproblem, das überindividuelle Gültigkeit haben soll, als nicht geeignet bewertet. Es wird nicht weiterverfolgt.

Begrifflich-semantisches Wissen wird als von der Einzelsituation abstrahiertes, generalisiertes Welt- und Bedeutungswissen einer Person definiert, „bei dem der Bezug zur raum-zeitlich einmaligen Lernepisode verlorengegangen ist.“ (Städler 1998). Begriffliches Wissen konstituiert sich in Begriffen und Begriffsstrukturen. Was genau unter Begriffen verstanden wird, hängt von der jeweiligen Theorie ab. Allgemein werden Begriffe als Konzepte (mentale Entitäten) verstanden, die Beschreibungen/Merkmale von wichtigen Eigenschaften einer Objekt- bzw. Situationsklasse etc. zusammenfassen und repräsentieren. Begriffe können mit Wortmarken versehen sein, müssen dies aber nicht. Begriffsstrukturen werden als Beziehungsstrukturen von Begriffen definiert, die z.B. Szenen, Ereignisse und Situationen oder Ober-/Unterbegriffsbeziehungen repräsentieren. Begrifflich-semantisches Wissen z.B. über „einen Supermarkt“ „ein Gartencenter“ „ein Einkaufszentrum“ oder „einen Flughafen“ repräsentiert nicht Wissen über einen

einzelnen, einzigartigen Supermarkt, sondern abstrahiertes und generalisiertes Wissen über viele Supermärkte. Es eignet sich somit auch für einen interindividuellen Vergleich: gesucht wird ja nicht eine individuell passende, sondern eine „überindividuell passende“ Anordnung von Objekten, die für die gesamte Nutzergruppe, d.h. für alle Personen passt, die innerhalb dieser Anordnung Objekte suchen müssen. Eine weiterführende Auseinandersetzung mit Modellen und Theorien über begrifflich-semantisches Wissen erscheint daher als sinnvoll.

Exkurs: Regelbasierte Produktionssysteme. In der kognitiven Psychologie wurden Modelle entwickelt, die die gesamte Architektur des kognitiven Apparates modellieren sollen. Solche „regelbasierten Produktionssysteme“ nehmen einen Interpretierer und zwei Wissensspeicher an (vgl. Opwis 1988, Anderson 1993): der *deklarative* und der *prozedurale* Wissensspeicher. Der deklarative Wissensspeicher beinhaltet dabei das begrifflich-semantische, episodische und anschaulich-analoge „Fakten“-Wissen. Der prozedurale Wissensspeicher beinhaltet Verfahren und Prozeduren zur Konstruktion, Verknüpfung und Anwendung von deklarativem Fakten-Wissen. Die Auswahl, Verarbeitung und die Auswertung von Wissen wird dabei vom Interpretierer durchgeführt.

Das Konzept des prozeduralen und deklarativen Wissensspeichers wurde von den Computerwissenschaften übernommen, die lange Zeit (bis zum objektorientierten Programmieren) strikt zwischen Programmen und Daten unterschieden. Heute wird auch von verschiedenen Autoren der Psychologie die Auffassung vertreten, dass solche statischen Speicherstrukturen der flexiblen Anpassung menschlicher Gedächtnisleistungen an die unterschiedlichsten Kontexte nur schwerlich gerecht werden. (vgl. Hoffmann 1995, Barsalou 1989) Kritisiert wird, dass gerade für komplexere Fertigkeiten Handlungs- von Objektwissen nicht trennbar ist, weil es sich aufeinander bezieht.

Bewertung Globale Informationsverarbeitungsmodelle eignen sich m. E. (noch?) nicht zur Lösung des Anordnungsproblems. Sie haben beschreibenden Charakter und sind nicht elaboriert genug, um auf ihrer Basis Prognosen zu treffen. Vielmehr bedürfte es noch einer Integration vorhandener Theorien, z.B. zu den unterschiedlichen Wissensrepräsentationsformaten und ihrer Interaktion. Meiner Erfahrung nach verlocken solche Modelle dazu, die kognitiven Prozesse, die bei Nutzern vermutet werden, zu modellieren. Wahrscheinlich in der Hoffnung, durch eine Beschreibung Ansatzpunkte für eine

Problemlösung zu bekommen. In dieser Arbeit wird anders vorgegangen. Trotz des Wissens um die Komplexität und Interaktion der verschiedenen kognitiven Prozesse wird eine Fokussierung auf eine Wissensform – nämlich begriffliches Wissen – bevorzugt. Und es wird eine Theorie zur begrifflichen Wissensrepräsentation gesucht, die explizit Aussagen darüber liefert, nach welchen Regeln dieses Wissen repräsentiert ist.

3.4 Zusammenfassung und weiteres Vorgehen

Die Ergebnisse der Problemanalyse lassen folgende Schlussfolgerungen und untersuchungsleitende Annahmen formulieren:

1. Leitsysteme, Hinweisschilder und Übersichtspläne sind zwar übliche Maßnahmen, um Nutzer bei ihren Suchprozessen zu unterstützen. Doch sie bringen viele Probleme mit sich, dass dies die Vermutung bestätigt, dass das Problem tiefer liegt und die Anordnung der Objekte im Raum nutzerorientiert gestaltet werden sollte.
2. Für die Entwicklung „nutzerorientierter“ Objektanordnungen ist es sinnvoll, die Gruppierungen bzw. Nachbarschaften und die Relationen der Objekte zueinander nach nutzerorientierten Kriterien zu gestalten. (Was liegt nebeneinander und wie nahe liegt es zueinander?) Solche topologischen und relationalen Anordnungen sind unabhängig von einem absoluten Bezugspunkt sowie unabhängig von absoluten, metrischen Distanzen zwischen den Objekten. Relationale bzw. topologische Anordnungen können von den Gestaltern (z.B. Regalbeleger, Architekten etc.) in konkrete metrische oder perspektivische Anordnungen übersetzt werden. Der Aspekt einer anthropometrisch, nutzerorientierten Anordnung ist hier nicht Gegenstand der Untersuchung und muss bei der Realisierung separat geprüft werden.
3. In der Suchsituation bildet der Nutzer spontane Erwartungen darüber, wo er ein Objekt findet. Erwartungen werden durch Schlussfolgerungsprozesse (Inferenzen) auf der Basis von Wissensbeständen des Langzeitgedächtnisses gebildet. Diesem Wissen, d.h. den Inhalten sowie den Regeln und Gesetzmäßigkeiten, nach denen diese Inhalte verarbeitet und gespeichert wurden, wird eine besondere Bedeutung beigemessen.

4. In einer Analyse wurden die vier Wissensarten episodisches, anschauliches, prozedurales und begrifflich-semantisches Wissen untersucht und bewertet. Es wird angenommen, dass sich in den Gesetzmäßigkeiten, nach denen begrifflich-semantisches Wissen repräsentiert ist, Vorgaben für nutzerorientierte Anordnungen finden lassen. Ferner wird angenommen, dass eine Anordnung der Objekte im Raum, die mit den begrifflichen Wissensstrukturen der Nutzer korrespondiert, Suchprozesse erleichtert.

4. Auswahl einer geeigneten Theorie für das Anordnungsproblem

Im folgenden Kapitel werden Theorien zu Begriffen und Begriffsstrukturen dargestellt und untersucht. Ziel ist es, eine Theorie auszuwählen, die sich eignet, einen Lösungsansatz für das Anordnungsproblem zu entwickeln *und* Vorhersagen liefert, nach welchen Regeln Begriffsstrukturen aufgebaut sind. Denn es wird angenommen, dass eine Anordnung der Objekte im Raum dann erwartungsgerecht ist, wenn sie mit den Begriffsstrukturen der Nutzer korrespondiert. Im *ersten* Abschnitt erfolgt erst einmal eine Einleitung in die Welt der Begriffe. Im *zweiten* Abschnitt werden die Kriterien für die Auswahl einer geeigneten Theorie dargestellt. Im *dritten* Abschnitt werden fünf Begriffstheorien dargestellt und anhand der Kriterien auf Eignung geprüft. Im *vierten* Abschnitt erfolgt eine Darstellung und Prüfung für fünf Begriffsstrukturtheorien und es wird die geeignete ausgewählt. Im *fünften* Abschnitt erfolgt eine Zusammenfassung.

4.1 Einführung in die Welt der Begriffe

Was sind Begriffe? Was unterscheidet sie von Worten? Ideen? Kategorien? Oder sind sie gar identisch? Die Philosophie, die Linguistik, die Computerwissenschaften und die Psychologie haben durch eine bunte Vielfalt an Aussagen viel dazu beigetragen, den Gegenstandsbereich der Begriffe eher zu vernebeln als zu erhellen (vgl. Aebli 1988). Eine allgemeine Begriffsdefinition ist daher nur in Ansätzen möglich, da die Definition stark vom entsprechenden Modell abhängt. So wird zunächst eine der Minimal-Definition gegeben.

Ein Begriff ist definiert als eine Gedächtnisabbildung (=Repräsentation), die wesentliche Eigenschaften von Gegenständen zusammenfasst. Ein Beispiel soll dies verdeutlichen: Der Begriff Tisch repräsentiert nicht einen individuellen Tisch, sondern die Eigenschaften, die einen Gegenstand zu einem Tisch machen. Die Frage, welche Eigenschaften begriffsbestimmend (klassifizierungsrelevant) sind, ob diese Eigenschaften als Merkmale abstrahiert oder in Eigenschaftslisten zusammengefasst und im Gedächtnis gespeichert werden, ist Gegenstand der aktuellen Debatte und kann selbst von vielen Theorien nur unbefriedigend beantwortet werden. Gemeinsam ist aber allen Theorien,

dass Begriffe aufgrund von Gemeinsamkeiten (= Ähnlichkeiten, vgl. Rips, 1995), die die Objekte haben, gebildet werden. So sind die Ähnlichkeiten von Objekten, die unter den Begriff Tisch fallen: eine horizontale (oder leicht schiefe) Fläche, ihre bestimmten Höhen-, Längen- und Breitenverhältnisse und die Möglichkeit hierauf etwas abzulegen. Ob der Tisch vier oder drei Beine hat, ob er Schubladen hat, ob er 20, 70 oder 110 cm hoch ist und welches Dekor er hat, ist für den Begriff Tisch unwichtig und dient nur der Unterscheidung von Unterbegriffen: z.B. Couch-, Beistell-, Bistro-, Schreib- oder Barock-Tisch.

Tabelle 4.1. Beispiele für unterschiedliche Begriffsarten

Beispiele für unterschiedliche Begriffsarten	
o Lebewesen:	Käufer, Kaktus, Kolibri
o Objekte:	Sessel, Seife, Süßgebäck
o Ereignisse:	Frühstück, Flirt, Finanzskandal
o Handlungen:	Kaufen, Küssen, Kaffeetrinken
o Orte:	Mensa, Metzger, Meeting-Point
o Zustände:	Treue, Tollheit, Todessehnsucht
o Urteile:	Schüchtern, Schönheit, Schickimicki
o Denkprodukte:	Märchen, Mode, Mental Map
o Gedankenwege:	Planen, Suchen, Orientieren
o Funktionen:	Ein, Aus, Speichern, Löschen
o usw. usw.	

Begriffe können über alles gebildet werden: über Lebewesen, Gegenstände, Urteile, Funktionen wie auch über mentale Produkte etc. (siehe Tab. 4.1). Begriffe werden erworben durch das Handeln, die Sprache oder über Belehrung. Es gibt viele Begriffe, die mit einer Wortmarke versehen sind (wie Bäcker, Keks, Alpenveilchen). Es gibt aber auch Begriffe, die keine Wortmarke haben (z.B. der Zustand kurz nach dem Aufwachen oder Personen, die Zeitplanungssysteme verwenden). Und es kommt auch vor, dass Begriffe, die zunächst keine Wortmarke haben, dann eine erhalten (z.B. Yuppies, Gefilofaxte, Connected). Leider gibt es keine griffige Wortmarke für den Oberbegriff von Lebewesen, Gegenständen, Ereignissen, Funktionen etc. Deshalb wird im Folgenden – um den Leser und das Papier zu schonen – einfach nur noch stellvertretend von „Begriffen über Objekte“ geschrieben. Je nach Kontext können damit Begriffe über Gegenstände, Personen, Funktionen o.Ä. gemeint sein.

Ein Begriff ist also nicht gleichzusetzen mit einem einzelnen oder einer größeren Anzahl von Gegenständen. Ein Begriff steht – als mentale Repräsentation – stellvertretend für die wichtigen Eigenschaften von gleichartigen Gegenständen oder präziser formuliert: für die wichtigen Eigenschaften einer Objektklasse.

Ein Begriff ist auch nicht gleichzusetzen mit einem Wort! Im Deutschen wird leider der Begriff Begriff häufig auch für den Begriff Wort verwendet, um damit direkt in einer Begriffsverwirrung zu landen. Im Englischen wird eindeutiger zwischen "word" (Wort) und "concept" (Begriff) unterschieden. Im Folgenden wird Begriff synonym mit Konzept, Klassifikation oder Kategorie verwendet (vgl. Gardner 1989). Hiermit sind dann alle anderen mit gemeint.

4.2 Kriterien für die Auswahl einer geeigneten Theorie

Innerhalb der psychologischen Forschung liegen eine Vielzahl von Begriffstheorien vor. Erste empirische Untersuchungen erarbeiteten Külpe 1904, Hull 1920 und Ach 1921. Vor allem in den 1970er und 1980er Jahren kam es dann innerhalb der Linguistik, der kognitiven Psychologie, der Künstlichen Intelligenz etc. fast zu einer Inflation von Begriffsrepräsentationstheorien mit sehr unterschiedlichen Problemstellungen und Paradigmen (vgl. Velickovskij 1988, Städtler 1998). Bereits 1982 beklagte Boris Velickovskij das Fehlen einer systematischen Kritik und Wissensakkumulation und zitierte auch Klagen von anderen Forschern (z.B. Tulving, Leary, Estes). Die Situation hat sich m.E. bis heute nicht wesentlich verändert (vgl. Gardner 1989, Opwis & Lüer 1998). Für das vorliegende Problem bedeutet dies ein mühsames Unterfangen: es müssen verschiedene Theorien hinsichtlich ihrer Eignung für das Anordnungsproblem untersucht und bewertet werden.

Die Bewertung und Auswahl der Theorien erfolgt nach vier Kriterien, die nach wissenschaftlichen und den Untersuchungsgegenstand betreffenden Gesichtspunkten gewählt sind. Im Vordergrund steht also nicht eine umfassende, wissenschaftliche Bewertung der Theorien und Modelle. Dies würde den Rahmen dieser Arbeit sprengen.

1. Kriterium: Differenziertheit der Aussagen

Um Hypothesen für erwartungsgerechte Anordnungen zu entwickeln, braucht es eine Theorie, die differenzierte Prognosen darüber liefert, nach welchen Regeln Begriffe und Begriffsstrukturen aufgebaut sind. Es wird davon ausgegangen, dass die Kenntnis der Regeln und Gesetzmäßigkeiten, nach denen die Begriffe und Begriffsstrukturen gebildet wurden, zur Lösung einer nutzerorientierten Anordnung beitragen. So würde es nicht ausreichen, wenn eine Theorie z.B. lediglich behauptet, dass Begriffe Strukturen bilden. Vorzuziehen ist eine Theorie, die differenzierte Aussagen darüber liefert, welcher Art die Beziehungen zwischen den Begriffen sind, welche Begriffe in Beziehung treten, welche nicht, wie stabil diese Verbindungen sind etc.

2. Kriterium: Kompatibilität mit natürlichen Begriffen

Innerhalb der Begriffsforschung wird aus Gründen der internen Validität gerne mit künstlichem oder sinnfreiem Material wie sinnlosen Silben, geometrischen Mustern o.Ä. experimentiert. Vorzuziehen ist jedoch eine Theorie, die auch Untersuchungen an natürlichen Begriffen angestellt hat und somit auch die ökologische Validität berücksichtigt. Schließlich soll die Theorie auf Gegenstandsbereiche des alltäglichen Lebens (Süßgebäck, Zimmerpflanzen, Geschäfte etc.) angewendet werden.

3. Kriterium: Experimentell geprüft mit sprachlichem und anschaulichem Material

Es gibt viele Theorien, die ausschließlich mit sprachlich gebundenem Material geprüft wurden (in Satzverifikations-, Textverstehensaufgaben etc.). Vorzuziehen ist eine Theorie, die mit sprachlichem und mit anschaulichem Material arbeitet, um sie auf die Problemstellung dieser Arbeit zu übertragen. Entsprechend den Ergebnissen der Explorationsstudie ist anzunehmen, dass Suchprozesse auch vorsprachlich ablaufen und durch anschauliche Wahrnehmungseindrücke beeinflusst werden.

4. Kriterium: Explikative und prognostische Kraft der Theorie

Es gibt innerhalb der Psychologie leider viele Theorien, die Verhalten erklären, aber nicht vorhersagen können. Da es sich in dieser Arbeit um ein Anwendungsproblem handelt, bedarf es einer Theorie, die sich zur Vorhersage eignet. Es soll ja schließlich auf der Basis dieser Theorie vorhergesagt werden, welchen Gesetzmäßigkeiten die

Erwartungsbildung folgt, um ein Vorgehen abzuleiten, wie man eine nutzerorientierte Anordnung entwickelt.

4.3 Darstellung und Prüfung von Begriffstheorien

Im Folgenden werden fünf unterschiedliche Begriffstheorien analysiert und auf Eignung als Problemlösungstheorie bewertet. Die Analyse geschieht nach folgender Struktur:

Definitionen und Basisannahmen: Hier wird dargestellt, wie in der Theorie Begriffe definiert, wie sie inhaltlich und strukturell aufgebaut sind und welche Informationsklassen sie repräsentieren. Außerdem werden die wichtigen Basisannahmen genannt.

Experimente: Hier werden ein bis zwei Methoden oder Experimente skizziert, um zu zeigen, welches methodische Paradigma der Theorie zugrunde liegt, und welche empirischen Befunde diese Theorie erklären soll.

Kritik: Hier werden Kritik und widersprüchliche Befunde dargestellt.

Bewertung hinsichtlich des Anordnungsproblems: Hier erfolgt die Bewertung der Theorie nach den oben dargestellten Kriterien: 1) Differenziertheit der Aussagen hinsichtlich Regeln und Gesetzmäßigkeiten; 2) Kompatibilität für natürliche Begriffe und Begriffsstrukturen; 3) Experimentelle Untersuchungen mit anschaulichem und sprachlichem Material; 4) Erklärungs- und Vorhersagekraft der Theorie.

4.3.1 Exemplartheorien: Begriffe als Mengen erlebter Exemplare

Definition und Basisannahmen Theorien der Exemplar- oder Mengenrepräsentation stehen in der Tradition von Aristoteles. Er nahm an, dass Begriffe durch Induktion und Assoziation gebildet werden. Gedächtnisspuren sind die Basis der Begriffe. Begriffe werden als Zusammenfassungen von individuellen Erfahrungen mit Objekten einer Klasse definiert (vgl. Sternberg 1969, Smith & Medin 1981). Die Erfahrungen werden wie Listen oder Mengen im Gedächtnis repräsentiert. Der Begriff Tisch beispielsweise wäre

durch die Mengen an Gedächtnisspuren repräsentiert, die unterschiedliche Tische bei der Person hinterlassen haben. Mengenrepräsentationen nehmen also an, dass einzelne Erfahrungen gespeichert werden und dass – anders als bei Abstraktionstheorien – nicht durch Abstraktionsprozesse, sondern durch *Assoziationsprozesse* ein Begriff gebildet wird.

Beispiele und Experimente Sternberg (1969) stützte dieses Modell durch Ergebnisse aus seinen Wiedererkennungsexperimenten: Versuchspersonen mussten entscheiden, ob ein aktueller Reiz zu einer zuvor gelernten Menge von Reizobjekten gehört. Er stellte fest, dass die Wiedererkennungszeit von der Anzahl der in der Lernmenge enthaltenen Reize abhängt. Mit jedem zusätzlichen Lernelement stieg die Wiedererkennungszeit in seinen Experimenten um einen konstanten Betrag. Sternberg schlug also ein Erklärungsmodell für Begriffe vor, nachdem die Lernmenge im Gedächtnis wie eine Liste angelegt wird. Die begriffliche Identifikation eines neuen Reizes erfolgt dann durch einen Vergleich mit den in der Liste enthaltenen Einzelelementen. Die Zuordnung beispielsweise eines „Sperlings“ zum Begriff Vogel sollte also schneller gelingen, als die Zuordnung zum Begriff Tier, der einen größeren Umfang hat. Für viele dieser Beispiele ist eine Bestätigung gefunden worden (Collins & Quillian, 1972, Collins und Loftus 1975)

Kritik und Gegenbeispiele Für die Theorie Sternbergs werden Gegenbeispiele angeführt: So fanden Smith, Shoben und Rips (1974) z.B., dass die Versuchspersonen schneller bestätigten, dass Aluminium ein Metall ist, als dass Aluminium ein Leichtmetall ist. Entsprechend der Exemplartheorie müsste die Antwortzeit für Aluminium – Metall länger sein, da Metall der umfangreichere Begriff als Leichtmetall ist. Dieses Beispiel zeigt aber auch eine andere Problematik: Dem Fachmann ist der Begriff Leichtmetall aus eigenem Studium und Erleben bekannt, und er weiß, welche Eigenschaften Leichtmetall hat, wozu man es verwendet, dass Leichtmetall der Oberbegriff von Aluminium ist und ein Unterbegriff von Metall. Aber meiner Meinung nach ist stark zu bezweifeln, dass Laien oder die Teilnehmer von psychologischen Experimenten eindeutig über den Begriff Leichtmetall verfügen. Meiner Ansicht nach ist es wahrscheinlich, dass Effekte dieser Art auf das Vorwissen der Versuchsperson zurückzuführen sind. Doch es ist nicht beschrieben, ob das Vorwissen kontrolliert wurde. Auch aus

theoretischer Sicht wurde gegen dieses Modell argumentiert. Des Weiteren lassen diese Theorien Fragen zur Begriffsbildung und -differenzierung unbeantwortet.

Bewertung hinsichtlich des Anordnungsproblems Exemplartheorien sind hinsichtlich ihrer Aussagen, nach welchen Gesetzmäßigkeiten Begriffe aufgebaut sind, undifferenziert und haben nur eine geringe Vorhersagekraft: In Begriffen sei eben nur das zusammengefasst und gespeichert, was eine Person erlebt hat. In den Experimenten wurden natürliche Begriffe untersucht, aber es wurde nur mit sprachlichem Material gearbeitet.

Für das Anordnungsproblem hätte diese Theorie eine einschränkende bis entmutigende Konsequenz: Wenn Personen ihre Begriffe aufgrund ihrer Erfahrungen gebildet haben, dann müsste in der Konsequenz die Erwartung über die Lokation eines Objektes nach den bisher gemachten Erfahrungen gebildet werden: die Personen müssten ein Objekt dort erwarten, wo sie es bisher erlebt haben. Dies hieße in der Praxis, dass die optimale Anordnung der Objekte diejenige ist, die eine Person gut kennt. Wollte man dieser Schlussfolgerung folgen, könnte der Problemlösungsprozess bereits an dieser Stelle abgebrochen werden. Als praktische Konsequenz hieße dies: Ordne die Dinge immer in der gleichen Anordnung an. Diese Sichtweise wird jedoch angezweifelt! Und sie kann überprüft werden: Würden die Exemplarmodelle zutreffen, dann müsste die begriffliche Ordnung im Gedächtnis ein Abbild der realen Anordnung der Objekte im Raum sein. Diese Frage wird z.B. in der Süßgebäckstudie untersucht.

4.3.2 Prototypentheorien: Begriffe als Prototypen

Prototypentheorien liegt eine nahezu entgegengesetzte Modellannahme zugrunde (vgl. Franks & Bransford 1971, Rosch & Mervis 1975). Vor allem Eleanor Rosch hat mit ihren Arbeiten zur Erforschung natürlicher Begriffe die Mengentheorien stark ins Wanken gebracht und dies hat bis heute eine nachhaltige Wirkung für neuere Theoriebildungen. Nach Eleanor Rosch werden Begriffe nicht beliebig gebildet. Sie folgt der Auffassung Gibsons, dass Begriffe die Wahrnehmungsstrukturen des Betrachters reflektieren, sowie die Handlungen, die er mit dem Objekt ausführt. Objekte werden zu Begrif-

fen (Kategorien) zusammengefasst, die als ähnlich wahrgenommen werden und die zu ähnlichen Handlungen auffordern.

Definitionen und Basisannahmen In Prototypentheorien repräsentiert der Begriff nicht eine Liste von Objekterfahrungen, sondern er stellt eine eigene kognitive Einheit dar. Zentraler Kern eines Begriffes ist hier der Prototyp. Ein Prototyp ist ein repräsentatives Beispiel der Objektklasse, das die typischsten Erscheinungen einer Objektmenge, im Sinne einer Familienähnlichkeit oder zentralen Tendenz, beinhaltet. Er reflektiert am deutlichsten die Redundanzstruktur (Rosch et.al. 1975). Der Prototyp muss nicht real existieren. Prototypentheorien treffen Aussagen über Alltagsbegriffe. Natürliche Begriffe haben meist nicht eindeutige und scharfe, sondern mehr oder minder unscharfe Grenzen. Die Zuordnung eines Objektes zu einem Begriff wird dann nach der Familienähnlichkeit bestimmt, also wie sehr das Objekt dem Prototypen ähnlich ist oder nicht.

Beispiele und Experimente Labov (1973) zeigte in Kategorisierungsexperimenten mit Tassen unterschiedlichster Form die Wirkung des Wahrnehmungs- und Handlungskontextes. Er zeigte den Versuchspersonen Zeichnungen von Objekten, die alle einen Henkel hatten, sich aber in Form, Höhe und Durchmesser unterschieden. Die Versuchspersonen mussten die gleiche Zeichnung in vier unterschiedlichen Kontexten benennen: Wurden die Vpn instruiert sich einen Kaffee-trinken-Kontext vorzustellen, nannten sie die Figur eine Tasse, sollten sie sich Kartoffelbrei in der Figur vorstellen, nannten sie sie eine Schüssel, und sollten sie sich die Figur mit Schnittblumen gefüllt vorstellen, nannten sie sie eine Vase. Die Ergebnisse zeigen, dass Objekte je nach Handlungskontext unterschiedlichen Begriffen zugeordnet werden und dass Begriffe unscharfe Grenzen haben. Das Prototypenmodell erlaubt es, dass ein und dasselbe Objekt verschiedenen Prototypen ähnlich sein kann und damit verschiedenen Begriffen zugeordnet werden kann. Diese begriffliche Kreuzklassifikation entspricht der Realität natürlicher Objekte.

Eleonore Rosch konnte an zahlreichen Untersuchungen mit natürlichen Begriffen zeigen, dass es innerhalb von Begriffshierarchien typische Begriffe gibt (die so genannten Basis- oder Primärbegriffe). Versuchspersonen mussten Abbildungen (z.B. eine Gitarrenabbildung) und Worte unterschiedlichen Abstraktionsniveaus zuordnen. In die-

sem Beispiel konnten die Personen schneller das Wort Gitarre zuordnen als die Unter- oder Oberbegriffe Schlaggitarre oder Musikinstrument. Der Abbildung eines Hammers konnten die Versuchspersonen schneller das Wort Hammer zuordnen, als die Worte Zimmermannshammer oder Werkzeug. Die ersten Experimente ließen sie annehmen, dass immer Objekte mittlerer Begriffsebene die besonders typischen Vertreter einer Hierarchie sind. Doch diese Annahme konnte nicht durchgängig bestätigt werden. So wurde die Abbildung einer Eiche schneller als Baum beschrieben denn als Eiche oder gar Roteiche (Abb. 4.2). (Wobei hierbei wieder hinterfragt wird, ob die Versuchspersonen überhaupt einen Begriff von der Roteiche hatten, und ob dies nicht vielmehr auf ein mangelndes Vorwissen zurückzuführen ist.)

Unterbegriff	Mittelbegriff	Oberbegriff	
Schlaggitarre	Gitarre*	Musikinstrument	
Boskopp	Apfel*	Frucht	
Zimmermannshammer	Hammer*	Werkzeug	
Küchentisch	Tisch*	Möbel	
Sportauto	Auto*	Fahrzeug	
Roteiche	Eiche	Baum*	
Seebarsch	Barsch	Fisch*	
Bergadler	Adler	Vogel*	* Primärbegriff

Abbildung 4.2. Primärbegriffe (Prototypen) auf unterschiedlichen Hierarchieebenen (Rosch 1977, zitiert nach Hoffmann 1983 S. 66)

Untersuchungen zeigten, dass es in jeder Begriffsklasse einen oder mehrere Begriffe gibt, die typischer sind als andere, allerdings auf sehr unterschiedlichen Begriffsebenen: Trompete und Geige sind hochtypische Musikinstrumente, Brot und Kartoffeln sind hochtypische Nahrungsmittel (Hoffmann & Ziebler 1982). Solche Primärbegriffe oder Prototypen sind innerhalb von Begriffshierarchien diejenigen, zu denen die Versuchspersonen mehr anschauliche Merkmale assoziieren können als zu ihren Unterbegriffen. Außerdem scheinen sie als die effektivsten Begriffe in der begrifflichen Identifikation.

Kritik Doch sind z.B. Brot, Kartoffeln, Trompete, Geige auch Primärbegriffe bei asiatischen Völkern, in denen dies westliche Importprodukte sind? Dies wird in den Arbeiten

nicht diskutiert! Da bei asiatischen Völkern z.B. der Reis der wichtigste Kohlehydratlieferant auf dem Speiseplan ist und nicht das Brot, ist anzunehmen, dass Brot bei ihnen auch kein Primärbegriff ist, sondern wohl eher Reis. Da die Geige ein westliches Musikinstrument ist, ist zu bezweifeln, dass dieses Instrument dort großen Bevölkerungsschichten überhaupt bekannt ist. Es ist anzunehmen, dass die Primärbegriffe auch von den Handlungserfahrungen, von der Kultur und der Nutzergruppe abhängen. Dies wird jedoch nicht diskutiert. Eleonore Rosch et al. deckten jedoch mit der Prototypentheorie wesentliche Erkenntnisse über natürliche Begriffe auf, von denen hier noch einmal wichtige zusammengefasst werden:

- Natürliche Begriffe sind meist unscharfe Begriffe. Ein und dasselbe Objekt (z.B. Baumstamm) kann dabei – je nach Kontext und Handlungsziel – unterschiedlichen Begriffsklassen (Baumstumpf und Sitzgelegenheit) angehören. Dies nennt man eine Kreuzklassifikation.
- Begriffe verschiedener Begriffsebenen (Ober-, Mittel- und Unterbegriff, wie Lebewesen – Vogel – Rotkehlchen) haben eine unterschiedliche Typikalität. Es gibt Primärbegriffe (wie Vogel, Rot, Hammer, Geige), die schneller und eindeutiger identifiziert werden als ihre Ober- und Unterbegriffe. Dabei sind Primärbegriffe nicht immer auf der selben Begriffsebene.

Bewertung hinsichtlich des Anordnungsproblems Prototypentheorien machen keine differenzierten Aussagen nach welchen Regeln und Gesetzmäßigkeiten sich ein Begriff inhaltlich konstituiert: Die Aussage, dass ein Begriff am deutlichsten die Redundanzstruktur der Klasse als Ganzes reflektiert, bleibt dahingehend unbefriedigend, dass die Theorie nicht erläutert, was die Redundanzstruktur oder die Familienähnlichkeit ausmacht. Eleonore Rosch u.a. arbeiteten zwar in ihren Experimenten mit natürlichen Begriffen sowie mit sprachlichem und nicht-sprachlichem Material. Außerdem deckten die Prototypentheorien wesentliche Erkenntnisse über natürliche Begriffe auf, wie die Unschärfe von Begriffen, die Phänomene der Kreuzklassifikation und der Prototypen (Primärbegriffe). Leider können die Prototypentheorien diese Phänomene jedoch nur vage erklären und vor allem nicht vorhersagen. Der Prototyp einer Objektklasse kann auch nur im Experiment ermittelt werden (vgl. Hoffmann 1983). Dies kommt einem Zirkelschluss nahe und ist ein Beleg für die geringe Vorhersagekraft.

Hinsicht der Fragestellung dieser Arbeit machen die Prototypentheorien den Mund wässrig auf eine ideale, prototypische Anordnung: z.B. der Prototyp von einem Süßgebäckregal, einem Einkaufszentrum oder einem Flughafen. Doch sie können den Weg dorthin nicht aufzeigen. Deshalb werden die Prototypenmodelle für eine Problemlösung als nicht geeignet bewertet. Die Prototypenmodelle regen aber zu einer interessanten Überlegung an: wenn es innerhalb jeder Objektklasse Prototypen gibt, die prägnanter sind und schneller identifiziert werden als andere Objekte der Objektklasse, dann könnten sie als „Identifizierungsmarken“ für die Objektgruppe benützt werden, z.B. als Icons/ Symbole für ein Leitsystem.

4.3.3 Gegenstandsorientierte Merkmalsrepräsentationstheorien

Merkmalsrepräsentationstheorien folgen alle der Grundannahme, dass Begriffe durch Merkmale beschrieben werden können (Ach 1921, Rips, Shoben & Smith 1973, Klix 1984, Hoffmann 1986). Begriffe sind Zusammenfassungen von Objekten und Erscheinungen nach ihren invarianten, spezifischen Merkmalen. Der Bildung von Begriffen, also der Gewinnung von klassifizierungsrelevanten Merkmalen, liegen Abstraktionsprozesse zugrunde und keine Assoziationsprozesse, wie bei den Exemplartheorien. Merkmalstheorien lassen sich daher in zwei Gruppen einteilen: in *klassische, gegenstandsorientierte Merkmalsrepräsentationstheorien*, die das Gewicht auf sensorische Ähnlichkeit legen, und in *verhaltensorientierte Merkmalsrepräsentationstheorien*, die das Gewicht auf funktionale Ähnlichkeit legen. Sensorisch ähnlich sind Objekte, die z.B. ähnlich aussehen, riechen, kurz, ähnliche Sinnesempfindungen hervorrufen. Funktional ähnlich sind Objekte, die die gleiche Wirkung und Funktion im Verhalten haben. So sind z.B. ein Baumstamm und ein Stuhl nicht sensorisch ähnlich, aber sie können funktional ähnlich verwendet werden: als Sitzgelegenheit. Im Folgenden werden gegenstandsorientierte Theorien untersucht.

Definitionen und Basisannahmen Ach (1921) und Hull (1920) gelten als die Begründer der experimentellen Begriffsforschung. Bereits in ihren ersten empirischen Arbeiten definierten sie Begriffe durch Merkmalscharakteristika. Allerdings verwendeten sie ein

Untersuchungsdesign, bei dem die Versuchspersonen die Klassifizierungsregeln für künstliche, vom Versuchsleiter vorgegebene Begriffe entdecken sollten. Leider bezeichneten Ach und Hull dies „begrifflich“ irreführend als Begriffsbildungsexperimente, was zur Folge hatte, dass lange Zeit Begriffsbildungsprozesse verkürztmaßen mit der Entdeckung vorgegebener Regeln gleichgesetzt wurden. Und dies trifft für die Bildung natürlicher Begriffe nicht zu.

Experimente Bei ihren Untersuchungen verwendeten sie geometrische Muster (z.B. künstliche chinesische Zeichen oder verschiedenfarbige Vierecke, Kreise etc.) mit gut unterscheidbaren Merkmalen. Begriffe definierten sie als eindeutige, trennscharfe, sensorische Merkmalslisten (z.B. rote Kreise). Die Versuchspersonen sollten die vorgegebenen Klassifizierungsregeln entdecken. Die klassifizierungsrelevanten Merkmale waren eindeutig und die Zugehörigkeit zu einem Begriff erfolgte nach dem Alles-oder-Nichts-Prinzip. Dem Untersuchungsdesign lagen somit Annahmen über Begriffe im Sinne von klar umgrenzten Booleschen Mengen zugrunde. Diese Prinzipien widersprechen aber den Phänomenen natürlicher Begriffe, wie sie bei den Prototypenmodellen dargestellt wurden. Damit ist dieses Untersuchungsdesign weit von den Erkenntnissen über natürliche Begriffe entfernt.

Auch wenn Clark Hull bereits 1920 darauf hinwies, dass Begriffe nicht um ihrer selbst Willen gebildet werden, sondern vielmehr im Zusammenhang mit dem erfolgreichen Erreichen von Verhaltenszielen stehen, so wurden doch in der Folge Begriffe nicht unter natürlichen Bedingungen untersucht. Das von ihnen entwickelte Untersuchungsdesign zur Entdeckung von Klassifizierungsregeln blieb lange Zeit bestimmend in der Untersuchung von Begriffen. Und in der Folge gab es verschiedene Modelle, die versuchten, klassifizierungsrelevante Merkmalsklassen zu formulieren.

Weitere Modelle Edward Smith, Edward Shoben und Lance Rips (1974) unterschieden in ihrem "Featural Model for Semantic Decisions" "definierende" und "charakterisierende" Merkmale (*features*). Definierende Merkmale für einen Vogel sind z.B. „hat Flügel“, „hat Federn“. Charakterisierende Merkmale sind "kann singen", "brütet in Bäumen". Definierendes Merkmal für ein Rotkehlchen z.B. ist „roter Bauch“, charakterisierende Merkmale sind z.B. „ist Wildvogel“, "ist klein". Definierende

Merkmale sind demnach spezifische Merkmale, charakterisierende Merkmale geben eher allgemeinere Eigenschaften wieder (s. Abb. 4.2).

	Rotkehlchen	Adler	Vogel
Definierende Merkmale	$F_{i,R}$	$F_{i,A}$	$F_{i,V}$
	$F_{i,R}$	$F_{i,A}$	$F_{i,V}$
Charakterisierende Merkmale	$F_{i+1,R}$	$F_{i+1,A}$	$F_{i+1,V}$
	$F_{m,R}$	$F_{m,A}$	$F_{m,V}$
	$F_{m,R}$	$F_{m,A}$	$F_{m,V}$

Abbildung 4.2. Modellvorstellung für definierende und charakterisierende Merkmale
 F = Merkmale, R = Rotkehlchen, A = Adler, V = Vogel (Smith, Shoben und Rips 1974 S. 216)

Die Zuordnung eines Objektes zu einem Begriff soll in zwei Phasen ablaufen (Smith, Shoben & Rips 1974): Zunächst würden die Merkmalslisten der möglichen Begriffe aktiviert. Im Vergleich zwischen ihnen würde ein allgemeiner Ähnlichkeitswert zwischen den charakterisierenden Merkmalen bestimmt. Ist eine hinreichende Ähnlichkeit gegeben, erfolge eine Bejahung. Wenn nicht, erfolge ein schrittweiser Vergleich der definierenden Merkmale. Diese zweite Prozessstufe dauere länger und habe eine niedrigere Fehlerquote als die erste. Das Modell macht gute Vorhersagen für Ober- und Unterbegriffsbeziehungen. Aussagen für Neben- oder Synonymbegriffe werden nicht gemacht. Begriffsbeziehungen wie zwischen „Tasse“ und „Löffel“ kann das Modell nicht leisten (vgl. Klix 1988; S. 29). Es erfasst sie nicht, weil diese Begriffe keine gemeinsamen charakterisierenden oder definierenden Merkmale haben.

Denis (1982, zitiert nach Hoffmann 1986) führte neben der Klasse der wahrnehmbaren Merkmale eine Klasse mit funktionalen Merkmalen ein. Er unterscheidet figurative (= perzeptiv wahrnehmbare) Merkmale und nichtfigurative Merkmale, die durch ihre funktionale Einbettung definiert sind.

Kritik Die Arbeiten von Smith, Rips & Shoben wurden vielseitig kritisiert. Aebli (1988) kritisierte die „Kalifornier“ mit beißenden Kommentaren, nur „uninteressante Begriffe“ und „keine echten Denkprozesse“ zu untersuchen. An Fragen wie: Ist ein Apfel eine Frucht? Ist Fischen ein Sport? oder Ist ein Rotkehlchen ein Vogel? hätten allenfalls

mittelalterliche Scholastiker ihre Freude gehabt. „Die Versuchsperson muss ihr enzyklopädisches Wissen analysieren. Im besten Fall müssen Objektbegriffe des Alltagsdenkens verglichen werden, aber mit welchem Zweck? Es gibt keinen Zweck, kein Handlungsziel und kein Erkenntnisziel außer dem Vergleich oder der Nennung der Unter- oder Oberbegriffe. Genau dies hat man dem scholastischen Denken vorgeworfen.“ (Aebli 1988, S. 242) Klix (1988 S. 29) kritisiert, dass das Modell von Smith, Shoben und Rips bereits bei Synonym- oder ereignisbezogenen Begriffsrelationen versagt.

Klassischen, gegenstandsbezogenen Merkmalsrepräsentationstheorien wird überdies vorgeworfen, sie berücksichtigen die Ergebnisse über natürliche Begriffe nicht und können diese auch nicht erklären. Sie orientierten sich vielmehr an den grundlegenden Axiomen der Booleschen Algebra (trennscharfe Klassen, Zugehörigkeit nach dem Alles-oder-Nichts-Prinzip etc.) und einem Objektivitätsbegriff, der als „Vor-Heisenbergisch“ bezeichnet werden kann. Sie nehmen an, dass der Merkmalsatz eine Zusammenfassung objektiver, invarianter Merkmale ist. Auch die Einführung von Modellannahmen über unterschiedliche Merkmalsgewichtungen reichen hierfür nicht aus. Außerdem seien die Annahmen über die unterschiedlichen Merkmalsklassen nicht überprüfbar und die Vielzahl der verschiedenen Modelle weisen auf eine Beliebigkeit der Merkmalsklassen hin.

Bewertung hinsichtlich des Anordnungsproblems Die klassischen, gegenstandsbezogenen Merkmalstheorien eignen sich aus folgenden Gründen nicht: Sie arbeiteten in Experimenten nicht mit alltagsrelevanten Fragestellungen. Phänomene von natürlichen Begriffen (z.B. Unschärfe der Merkmalsklassen, Kreuzklassifikation) sind nicht in die theoretische Konzeption eingeflossen. Sie nennen zwar Merkmalsklassen, nach denen sich Begriffe konstituieren sollen. Sie treffen differenziertere Aussagen über die innere Struktur als z.B. Mengenrepräsentationsmodelle. Die Anzahl und Art der Merkmalsklassen muten aber sehr willkürlich an. Es bleibt die Frage offen, ob es nicht mehr oder weniger oder andere Merkmalsklassen gibt. (Warum gibt es z.B. bei Rips, Shoben & Smith (1972) nur definierende oder charakterisierende Merkmale, warum nicht auch andere?) Die Prognosekraft der Theorien ist gering, da keine Theorie die begriffsbestimmenden Merkmale eines Objektes voraussagen kann.

4.3.4 Verhaltensorientierte Merkmalsrepräsentationstheorie

Definitionen und Basisannahmen Verhaltensorientierte Merkmalstheorien nehmen an, dass Begriffe untrennbar mit dem Verhalten verbunden sind (Klix 1971, 1992, Hoffmann 1983, 1986, 1993). Klix und Hoffmann folgen damit den Annahmen anderer Begriffsforscher, wie Ach, Hull und Rosch, dass Begriffe nicht um ihrer selbst Willen gebildet werden, sondern dem Verhalten dienen und Ziele, Wahrnehmungsstrukturen und Handlungsfolgen beinhalten. Anders als die genannten Forscher setzen sie aber diese Annahmen zufriedenstellender in ihren Modellvorstellungen um.

Friedhart Klix (1971, 1992) und Joachim Hoffmann (1983, 1986) definieren Begriffe als mentale Repräsentationen, die invariante Merkmale einer Objektmenge zusammenfassen. Das Nervensystem filtert die stabilen Objektmerkmale während des Handelns mit dem Gegenstand heraus. Bei einem Begriffsbildungsprozess erfolgt durch kognitive Prozeduren und Operationen eine Abstraktion, Zusammenfassung und Speicherung dieser invarianten, "klassifizierungsrelevanten" Merkmale. Diese Merkmale gewinnt die Person also in der Handlungsabfolge mit dem Objekt durch die Wahrnehmung. Die Handlungsabfolge erfolgt aufgrund eines bestimmten Orientierungsanlasses (Klix 1971). Klix definiert dabei Orientierungsanlass sehr weit und versteht darunter nicht nur eine spontane Instinkt- oder Orientierungsreaktion, sondern auch ein Bedürfnis und ein Interesse (Klix 1971).

Der *Orientierungsanlass im Sinne von Motiv, Interesse, Bedürfnis* bildet einen Fixationspunkt für die Ordnungsbildung im Gedächtnis. Er hat einen wesentlichen Einfluss auf die innere Einstellung und die Wahrnehmung gegenüber dem Merkmalsangebot. Und er führt auch zu Handlungsabläufen, die wiederum bestimmte Merkmalsangebote kreieren: im Umgang mit dem Objekt erfährt die Person während des Handelns sehr spezifische Merkmale über das Objekt. Diese neu gewonnenen Merkmale führen ihrerseits wieder zu spezifischen Begriffsbildungsprozessen. „...Orientierungsanlass und Verhaltensfolge sind miteinander verbundene Größen, deren Bindung an die auftretenden Merkmale als wesentlich für die entstehende kognitive Strukturbildung gesehen wird.“ (Klix 1971, S. 537). Der passionierte Angler, der Pferdeliebhaber, der Verhaltensforscher gelangen – aufgrund ihrer Orientierungsanlässe – zu Merkmalsunterschei-

dungen, über die der Außenstehende, der „Laie“ nicht verfügt. Die Erkennungsverfeinerungen führen – den Hof des Interesses betreffend – zu Kenntnissen über Objekteigenschaften, die den Spezialisten vom Kenntnisarmen unterscheiden und zwar aufgrund differenzierterer Klassenbildungs- und folglich auch Erkennungsprozesse. (Klix 1971, S.548). Die Mächtigkeit des Begriffsfeldes einer Person ist somit abhängig von den Orientierungsanlässen im Sinne von Motiven, Bedürfnissen, Interessen der Person.

Klix (1971, 1988, 1992) und Hoffmann (1983, 1986) unterscheiden vier begriffskonstituierende Merkmalsklassen (siehe auch Abb. 4.3):

- *Sensorische Merkmale* Sensorische Merkmale repräsentieren Reizwirkungen, die über die Sinnesorgane wahrnehmbar und verhaltensrelevant sind. Dazu können Reizwirkungen gehören bezüglich Größe, Form, Farbe, Geruch, Oberflächentextur, Temperatur oder Geschmack. Begriffsrelevante Merkmale können elementare oder komplexe Merkmale sein. Sinneseindrücke wie braun, grün sind elementare Merkmale. Aufgliederbare Eindrücke wie Stamm, Ast, Ästchen, Blätter sind komplexe Merkmale, die meist selber wieder Begriffe darstellen.
- *Ereignisbezogene (= horizontale) Merkmalsrelationen* Sie verweisen auf andere Begriffe, die im gemeinsamen Kontext erlebt werden. Dazu gehören nach Klix (1988, 1992) Merkmalsrelationen zum Handlungsträger, zum Instrument, zu anderen Objekten im Kontext, zum Ort, zur Zeit, zu Zielen. Horizontale Beziehungsmerkmale bilden die Voraussetzungen, damit Begriffe untereinander Verbindungen eingehen können, wie sie im Abschnitt über Begriffsstrukturen dargestellt werden. Folgendes Beispiel soll die horizontalen Merkmalsrelationen des Begriffes Lehrer veranschaulichen: Handlungsträgerrelation (Lehrer), Finalitätsrelation (Wissen vermitteln), Rezipientenrelation (Schüler), Instrumentalrelation (Bücher, Tafel), Objektrelation (Schulbank), Ortsrelation (Schule, Klassenzimmer), Zeitrelation (Vormittags).
- o *Hierarchische Beziehungsmerkmale* Hierarchische Beziehungsmerkmale sind von ganz anderer Qualität: sie stellen Adressen dar für begriffliche Über- oder Unterordnungen. Während horizontale Ereignismerkmale auf Ereignisse und Situationen verweisen (z.B. Bild und Malen, Keks und Kaffeetrinken), verweisen hierarchische Merkmale auf übergeordnete, untergeordnete (Butterkeks – Keks – Süßgebäck) oder

auf nebengeordnete Begriffe (Butterkeks, Schokoladenwaffel). Hierarchische Beziehungsmerkmale bilden die Voraussetzung für Begriffe, hierarchische (= taxonomische) Begriffsstrukturen zu bilden (z.B. Apfel – Gartenobst – Obst – Nahrungsmittel). Klix (1992) unterscheidet folgende hierarchische Merkmalsrelationen: Ober-/Unterbegriffsrelation (Schüler – Lehrer), Nebenbegriffsrelation (Schüler – Klassenkamerad), Synonymrelation (Schüler – Schuljunge), Antonymrelation (Unwissender – Wissender), Komparativrelation (Anfänger – Fortgeschrittener).

- o *phonemisch-graphische Merkmale* Sprachmerkmale repräsentieren die graphemischen und phonemischen Eigenschaften des Wortes, das einen Begriff benennt. Wie eingangs beschrieben, kann ein Begriff eine Wortmarke haben, muss aber nicht. Für viele Begriffe gibt es keine Wortmarken (wie für die typische Form von Eichenblättern oder die „richtige“ Anordnung von Teller, Besteck, Glas und Serviette auf einem gedeckten Tisch).

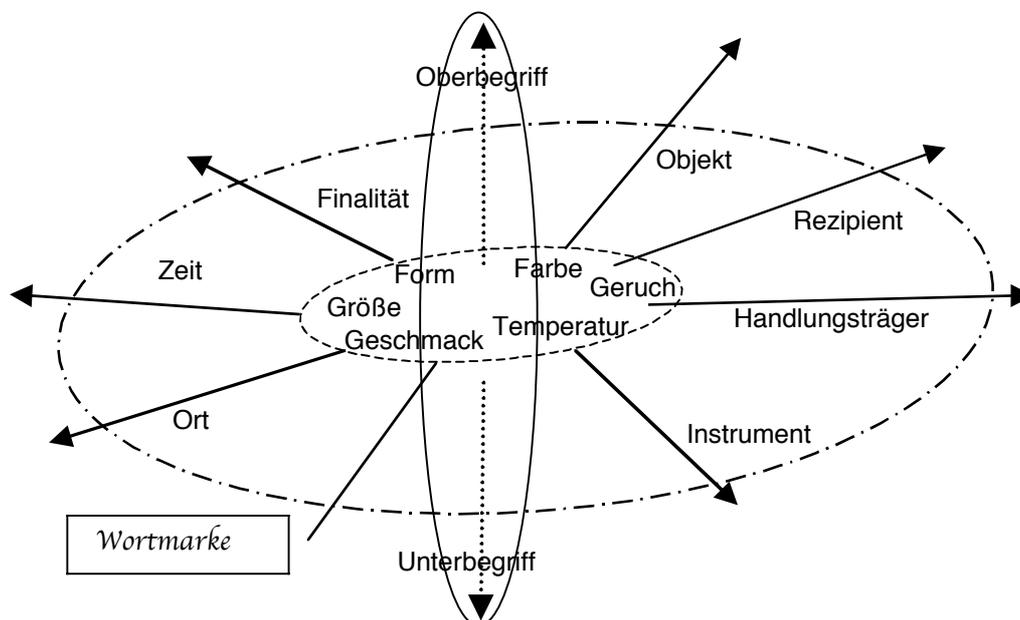


Abbildung 4.3. Modellvorstellung eines Begriffes mit den vier Merkmalsklassen: sensorische Merkmale (Mitte), ereignisbezogene Merkmalsrelationen (horizontale Ebene), hierarchische Beziehungsmerkmale (vertikale Ebene) und die Wortmarke (angehängt) (visualisiert von der Autorin nach Klix 1971, Hoffmann 1986)

In Untersuchungen bestätigen Hoffmann & Ziebler (1982) und Ziebler (1983) (zitiert nach Hoffmann 1986) die Merkmalsklassen von Klix, nehmen jedoch noch zwei zusätzlich Merkmalsklassen an:

- *Verhaltensmerkmale* Mit Verhaltensmerkmalen führt Hoffmann (1986) eine neue Merkmalsklasse ein. Verhaltensmerkmale verweisen für ihn auf entsprechende Verhaltensprogramme (z.B. Auto: fahren, beschleunigen, bremsen, voll packen, betanken etc.). Verhaltensprogramme sind zweifach repräsentiert: a) in Form von elementaren, neuronal organisierten Bausteinen für sensomotorische Verhaltensketten (wie das Ballen der rechten Faust), b) in Form von begrifflichen Einheiten, vor allem für komplexere Verhaltensprogramme (wie hämmern, wischen), die wiederum auf elementare, neuronal organisierte Bausteine für Verhaltensketten verweisen.

Verhaltensmerkmale hält er für notwendig, um z.B. die Unmittelbarkeit zu erklären, mit der viele Handlungsintentionen spontan in tatsächliches Handeln umgesetzt werden können. Bei dieser Modellannahme überträgt er analog die Repräsentationsformen von sensorischen Informationen bei der Informationsaufnahme (elementare und komplexere sensorische Reizwahrnehmung wie die Wahrnehmung der Farbe rot oder die Wahrnehmung einer Apfelform) auf die Repräsentation von Verhaltensprogrammen. Für eine solche begriffliche Repräsentationsform von Verhaltensfolgen sprechen Phänomene, wie sie auch bei anderen Begriffen eintreten, z.B.

- Kurzzeitgedächtnis-Effekte für Bewegungslernen: in Behaltensspannen von 10-20 Sekunden gelingt noch eine genaue Reproduktion, dann fällt die Genauigkeit drastisch ab. (Stelmach und Hughes 1984 zitiert nach Hoffmann 1986 S. 150)
 - Organisationseffekte bei der Reproduktion von Bewegungen: Ähnliche Bewegungen verbessern die Reproduktionsleistung (Stelmach und Hughes 1984; zitiert nach Hoffmann 1986 S. 150)
 - Wortmarken-Effekte für Bewegungslernen: die Benennung von Bewegungen verbessert deren Erlernbarkeit (Shea 1977; zitiert nach Hoffmann 1986)
- *Emotionale Merkmale* Emotionale Merkmale verweisen für Hoffmann auf emotionale Bewertungen. Dabei sind einige Begriffe überindividuell mit emotionalen Bewertungen verbunden (z.B. WITZ, TRAUER, URLAUB, KRIEG).

Experimente Klix, Hoffmann et al. arbeiteten in ihren Experimenten vor allem mit natürlichen Begriffen anhand von sprachlichem und anschaulichem Material. Gemessen wurden Reaktionszeiten und Antwortverhalten z.B. in Satz-/Bildverifikations-, Begriffsfindungs- und Analogiebestimmungsaufgaben. Hoffmann untersuchte die von Eleonore Rosch entdeckten Primärbegriffe, um sie mit Hilfe der Merkmalsrepräsentationstheorie zu erklären. Doch dies gelingt ihm erst vor dem Hintergrund seines Modells der antizipativen Verhaltenssteuerung.

Kritik Klix, Hoffmann et al. werden dahingehend kritisiert, dass sie nicht nur unzureichend erklären können, warum welche Merkmale eines Objektes denn nun klassifizierungsrelevant seien. Problem sei aber vor allem die Vorhersage, welche Merkmale eines Objektes klassifizierungsrelevant sein werden. Klix, Hoffmann et al. betonen zwar die Verhaltensrelevanz von Begriffen und Begriffsstrukturen, Klix schreibt bereits 1971, dass Begriffe im Verhalten erworben werden. Er bettete jedoch die Bildung von Begriffen nicht in Lernprozesse ein. Dies machte Hoffmann erst später in seiner Theorie der Antizipativen Verhaltenssteuerung (1993).

Bewertung hinsichtlich des Anordnungsproblems Die Merkmalsrepräsentationstheorie von Klix (1971, 1984, 1991, 1992), Hoffmann et al. (1983,1986) eignet sich aus folgenden Gründen: Sie liefert die differenziertesten Aussagen, nach welchen Regeln und Gesetzmäßigkeiten sich Begriffe konstituieren. Die Aussagen sind im Experiment mit künstlichen *und* natürlichen Begriffen, mit sprachlichem *und* anschaulichem Material untersucht. Die Theorie hat eine gute Erklärungs- und eine zumindest bedingte Prognosekraft: anhand der Theorie können zumindest differenzierte Hypothesen gebildet werden, durch welche Merkmalsklassen sich ein Begriff konstituiert. Außerdem regt sie zu differenzierteren Spekulationen an als die bisher dargestellten Theorien, welche Merkmale dann die begriffsrelevanten Merkmale sind. Damit hat die Theorie eine Aussagentiefe und eine heuristische Kraft, die es ermöglicht, Annahmen über die konstituierenden Merkmale von Begriffen zu formulieren.

Wenn Begriffe z.B. in Abhängigkeit vom Orientierungsanlass, Bedürfnis bzw. Interesse gebildet werden, dann müssten sich bei Personen mit unterschiedlichen Interessen und dementsprechend unterschiedlichen Handlungserfahrungen auch die Begriffe über

gleiche Objektklassen unterschiedlich konstituieren. Hier einige Beispiele, um das Gesagte zu veranschaulichen:

- o bei Zimmerpflanzen-Verkäuferinnen müsste sich der Begriff des Usambara-Weilchens aus anderen Merkmalen zusammensetzen als bei einer Kundin, die diese Pflanze kauft. Die ereignisbezogenen Merkmale müssten sich bei der Verkäuferin auf die Pflege und den Abverkauf im Verkaufsraum beziehen. Außerdem müsste die Verkäuferin wissen, welches der Ober- bzw. der Unterbegriff des Usambara-Weilchens ist. Bei der Kundin müssten sich die ereignisbezogenen Merkmale vielmehr auf das Kaufen, die Pflege daheim beziehen.
- o bei beruflich bedingt viel fliegenden Personen müsste sich z.B. der Begriff von einem Parkleitsystem anders konstituieren als Wenigfliegern. Die doch zum Teil unterschiedlichen Verhaltensabläufe und -konsequenzen müssten sich im Begriff, d.h. in den invarianten Merkmalen ausdrücken.

Die Merkmalsrepräsentationstheorie von Klix (1971, 1984, 1991, 1992), Hoffmann et al. (1983,1986) scheint sich entsprechend der Bewertungskriterien als Begriffstheorie zu eignen. Trotzdem wird noch die neueste Theorie von Hofmann untersucht, um zu prüfen, ob diese vielleicht noch leistungsstärker ist.

4.3.5 Begriffe in der Theorie der Antizipativen Verhaltenssteuerung

Basisannahmen der Theorie Hoffmann entwickelte in den vergangenen Jahren eine Theorie der Antizipativen Verhaltenssteuerung (1993). Er betrachtet Begriffe und Begriffsstrukturen innerhalb des psychischen Gesamtgeschehens. Er bettet sie in Lernprozesse ein und betont die verhaltenssteuernde Funktion von kognitiven Prozessen. Er betrachtet Verhaltensregelung im Sinne eines kybernetischen Regelkreises. Hoffmann arbeitet mit folgenden Basisannahmen: a) Informationsverarbeitung dient einer effektiven Verhaltenssteuerung; b) effektive Verhaltenssteuerung setzt Wissen voraus, unter welchen Bedingungen welche Verhaltensweisen zu welchen Konsequenzen führen; c) ein kognitiver „Antizipationsmechanismus“ stellt vor dem Verhalten antizipierte Konsequenzen (Soll-Werte) bereit; d) antizipierte Konsequenzen (Soll-Werte) werden mit

realen Konsequenzen (Ist-Werte) verglichen; e) dieser Feedback-Regelkreis (Soll-Ist-Vergleich) erlaubt Lernprozesse, die zu einer Erweiterung der Begriffe und Strukturen führen und damit zu effektiverem Verhalten.

Hoffmann nimmt an, dass begriffliches Wissen durch Verstärkerprozesse erlernt wird. Beim Start einer Handlung bilde die Person Erwartungen (Antizipationen) darüber, was die Konsequenzen ihres Handlungsentwurfes sein könnten. Stimmen die antizipierten Konsequenzen mit den realen Konsequenzen des Verhaltens überein, so wirke dies wie eine Verstärkung. Die Person lernt. Ein Beispiel (aus Hoffmann 1993) soll dies verdeutlichen: eine Person sieht einen Trinkbecher denkt, er sei aus Glas, greift entsprechend ihren Materialeinschätzungen zu und erwartet, dass der Becher dem Druck der Hand standhält. Ist der Becher aus Glas, wird die Erwartung bestätigt. Ist der Becher aus durchsichtigem Plastik, dann wird der Becher unter dem Druck deformiert. Die antizipierten Konsequenzen des Verhaltens stimmen nicht mit den realen überein und die Person muss ihr Verhalten in Abhängigkeit von den Gegebenheiten des Plastikbeckers ändern. Sie lernt dadurch zweifach: sie lernt einen durchsichtigen Plastikbecher von einem Glas zu unterscheiden. Und sie lernt den Druck der Hand zu modifizieren.

Hoffmann integriert mit diesem Modell: Kognitive Prozesse zur Bereitstellung des Wissens über die Ausgangssituationen, Verhaltensweisen und Konsequenzen des Verhaltens, Wahrnehmungsprozesse zur Analyse der Bedingungen vor, während und nach dem Verhalten und Handlungssteuerungsprozesse zur Durchführung eines angepassten Verhaltens.

Definitionen und Basisannahmen für Begriffe Kognitive Wissensstrukturen, vor allem Begriffe und Begriffsstrukturen, bilden die Klammer für die Wahrnehmung und die Handlungssteuerung. Hoffmann distanziert sich von der – vielen kognitionspsychologischen Modellen zumindest implizit zugrunde liegenden – Vorstellung, der Zweck der Informationsverarbeitung und Begriffsbildung bestehe im Aufbau eines möglichst korrekten Abbildes der Umwelt. Er sieht den Zweck von Begriffen in der Sicherung des Verhaltenserfolges unter wechselnden Bedingungen (Hoffmann 1996).

Hoffmann hebt damit die funktionale Bedeutung von Begriffen für das Verhalten hervor. Begriffe beinhalten für ihn das Wissen, unter welchen Bedingungen welches

Verhalten zu welchen Konsequenzen führt. Werden in frühen Merkmalsrepräsentationstheorien unter einem Begriff nur die Zusammenfassungen klassifizierungsrelevanter Merkmale durch Abstraktionsprozesse verstanden, so definiert er nun Begriffe als Zusammenfassungen von Erscheinungen, die durch ein bestimmtes Verhalten in invarianter und damit vorhersagbarer Weise verändert werden können. Er definiert Begriffe als *Klassen funktionaler Äquivalenz* oder anders formuliert: Begriffe sind Zusammenfassungen von Erscheinungen, die durch ein bestimmtes Verhalten in invarianter und damit vorhersagbarer Weise verändert werden können. ”Deshalb gehören auch so unterschiedliche Dinge, wie etwa ein Baumstumpf und ein Stuhl, zu einem Begriff Sitzgelegenheit, eben weil sie gleichermaßen die Gelegenheit zum sich setzen bieten.“ (Hoffmann 1992, S. 145) Hoffmann zieht den Schluss, Begriffsbildung sei nicht mehr primär als Abstraktion gemeinsamer sensorischer Merkmale zu verstehen. „Primär wird eine Bestimmung funktionaler Äquivalenzen durchgeführt, erst sekundär werden invariante Merkmale abstrahiert.“ (Hoffmann 1993, S. 151). Die Mechanismen einer antizipativen Verhaltenssteuerung leisten diese Äquivalenzbestimmung.

Hoffmann bietet hiermit erstmals eine lerntheoretische Definition an, was ”klassifizierungsrelevante”, invariante Merkmale sind. Es sind eben jene Merkmale, die wiederholt zur funktionellen Äquivalenz und zum Verhaltenserfolg beitragen. Die Person erlebt im Handlungsprozess bei der Realisierung eines Handlungszieles die Ausgangssituation, das Verhalten und die Konsequenzen mit seinen verschiedenen Merkmalsausprägungen. Diejenigen funktionalen Äquivalenzen, die bei einer erfolgreichen Handlungsausführung (also einer Übereinstimmung von antizipierten und realen Konsequenzen) erlebt werden, werden abstrahiert, zu Begriffen zusammengefasst und gespeichert.

Hoffmann versucht mit diesem Modell auch die Primärbegriffe zu erklären, allerdings fehlen noch systematische Untersuchungen. Bei Primärbegriffen – so nimmt er an – treffen Bedingungen zusammen, die die Äquivalenzklassenbildung fördern. So kann ein untergeordneter Begriff wie Rose in gleicher Weise Eigenschaften aufweisen wie der Basisbegriff Blume, weil Rosen gegenüber anderen Blumen erstens eine eigenständige funktionale Äquivalenz besitzen (als Zeichen besonderer Zuneigung) und zweitens sich in anderen Merkmalen (z.B. Dornen) von anderen Blumen hinreichend unterscheiden.

Kritik Hoffmann liefert kaum Aussagen über die innere Struktur von Begriffen. Begriffe beinhalten nach diesem Modell die invarianten funktionale Äquivalenzen a) der Ausgangssituation, b) des Verhaltens und c) der Konsequenzen für eine erfolgreiche Handlung. Den Begriffen ist somit die Ziel- und Zweckgerichtetheit inne, die die entsprechenden Objekte im Verhalten haben. Spezifischere Aussagen liegen bislang nicht vor. Die Annahmen sind (meines Wissens) bislang noch nicht empirisch überprüft.

Hoffmann unterscheidet auch nicht deutlich genug zwischen dem von ihm vorgeschlagenen Antizipationsmechanismus, der dafür zuständig ist begriffliches Wissen über antizipierte Verhaltenskonsequenzen bereitzustellen und dem begrifflichem Wissen selbst. Des Weiteren sind seine Darstellungen über diesen Antizipationsmechanismus bislang wenig differenziert bzw. elaboriert. Wie soll dieser Mechanismus geartet sein, der Äquivalenzbestimmungen durchführt? Wie ist seine innere Struktur?

Bewertung hinsichtlich des Anordnungsproblems Die Ausführung zu Begriffen, die Hoffmann im Rahmen seiner Theorie der antizipativen Verhaltenssteuerung liefert, sind noch zu wenig elaboriert, um sie für die Lösung des Anordnungsproblems zu benutzen. Auch steht eine experimentelle Überprüfung der Aussagen noch aus. Ebenso steht noch eine vertiefende Auseinandersetzung mit den Erkenntnissen seiner früheren Merkmalsrepräsentationstheorie aus. M. E. spricht jedoch vieles dafür, dass sich beide Theorien integrieren lassen: Hoffmann liefert mit seiner neuen Theorie eine Integration von kognitiven Prozessen und Lernprozessen. Er betont vor allem die erfolgreichen Verhaltenskonsequenzen als bedeutsam für die Begriffsbildung. So kann er ein erstes Kriterium für die Klassifizierungsrelevanz von Merkmalen nennen: Klassifizierungsrelevant sind jene Erscheinungen, die zu erfolgreichen Verhaltenskonsequenzen führen. Allerdings müsste diese Aussage näher spezifiziert werden, um z.B. Vorhersagen bilden zu können.

4.3.6 Überblick über das Ergebnis der Prüfung

Es wurden fünf Begriffstheorien untersucht. Für jede Theorie wurden die Basisannahmen und die jeweiligen Begriffsdefinition dargestellt, es wurden die wichtigsten Experimente sowie die Kritik und widersprechende Befunde angeführt. Abbildung 4.4 gibt

noch einmal Überblick über diese Theorien und nennt die wichtigsten Begriffsdefinitionen.

Exemplartheorien	Prototypentheorien	Gegenstandsbezogene Merkmals-theorien	Verhaltensorientierte Merkmals-theorie	Antizipative Verhaltenssteuerung
Sternberg 1969 Smith & Medin 1981	Rosch 1975 Brandsford 1972	Hull 1920 Rips et al. 1972	Klix 1971 Hoffmann 1986	Hoffmann 1993
Begriff = „Liste“ von Er- fahrungen	Begriff = Prototyp, mit der größten Familien- ähnlichkeit (=Redundanz- struktur) für die Objekt- klasse	Begriff = Merkmalsatz mit definierenden und charakt- erisierenden Merkmale	Begriff = Merkmalsatz mit - sensorischen - ereignis- bezogenen - kategorialen - phonemisch- graphischen Merkmalen	Begriff = Klasse funktionaler Äquivalenz welches Ver- halten unter welchen Be- dingungen zu welchen Kon- sequenzen führt

Abbildung 4.4. Überblick über die fünf untersuchten Begriffstheorien

Alle fünf Theorien wurden hinsichtlich ihrer Eignung zur Lösung des Anordnungsproblems nach vier Kriterien geprüft. Die Ergebnisse der Prüfung zeigt die Abbildung 4.5. Die Prüfung erbrachte, dass sich die verhaltensorientierte Merkmalsrepräsentationstheorie von Klix (1971, 1993) und Hoffmann (1983, 1986) hinsichtlich der praktischen Fragestellung dieser Arbeit am besten eignet.

Sie liefert die differenziertesten Aussagen auf die Fragen „Wie sind Begriffe inhaltlich aufgebaut?“, „Welche Informationsklassen beinhalten sie?“. Die Aussagen sind in Experimenten mit sprachlichem und anschaulichem Material, an natürlichen und künstlichen Begriffen untersucht. Die Ergebnisse scheinen übertragbar auf Begriffe des täglichen Lebens, wie Lebensmittel, Zimmerpflanzen, Geschäftstypen. Und sie lassen erste Annahmen formulieren zur Lösung des Anordnungsproblems: z.B. dass Personen mit unterschiedlichen Handlungszielen sich in ihren Begriffen unterscheiden: Personen, die z.B. viel fliegen haben mehr Begriffe bezüglich Flughafeneinrichtungen als solche, die nicht fliegen, und selbst gleiche Begriffe unterscheiden sich in ihren Merkmalsätzen.

So wird der Merkmalsatz für die Begriffe Treffpunkt oder Ticketschalter bei Personen, die viel fliegen, anders sein, als bei Personen, die nicht fliegen.

Theorien und Modelle	Exemplartheorien	Prototypentheorien	Gegenstandsorientierte Merkmals-theorien	Verhaltensorientierte Merkmals-theorien	Antizipative Verhaltenssteuerung
Beurteilungskriterien	Sternberg 1969	Rosch 1975 Frank & Bransford 1972	Hull 1920, Rips et al. 1972	Klix 1971 Hoffmann 1986	Hoffmann 1993
1. Differenzierte Begriffsdefinition	nein	nein	nein	ja	nein
2. Kompatibel für natürliche Begriffe	nein	ja	nein	ja	ja
3. Experimentell geprüft	ja	ja	ja	ja	?
4. Erklärungs- und Vorhersage geeignet	nein	nein	nein	bedingt	ja

Abbildung 4.5. Übersicht über die Prüfung der Theorien anhand der vier Kriterien

Es folgt eine Auseinandersetzung mit Theorien bzw. Modellen über Begriffsstrukturen mit dem Ziel, eine geeignete Theorie auszuwählen. Es werden die gleichen Kriterien angelegt.

4.4 Darstellung und Prüfung von Begriffsstrukturtheorien

Im vorhergehenden Abschnitt wurden fünf Begriffstheorien geprüft und es wurde eine ausgewählt, die als geeignet für das Lösen des Anordnungsproblems erscheint. Im folgenden Abschnitt werden fünf Theorien bzw. Modelle zu *Begriffsstrukturen* dargestellt und bewertet, um ebenfalls eine geeignet erscheinende Theorie auszuwählen.

Alle fünf Begriffsstrukturtheorien gehen von der Annahme aus, dass begrifflich-semantisches Wissen nicht in isolierten Begriffen repräsentiert wird, sondern dass die

Begriffe miteinander in definierter Beziehung stehen. Der semantische Gehalt, d.h. die Bedeutung eines Begriffes ist durch seine Beziehungen zu anderen Begriffen festgelegt (zur Kritik des Urnenmodells des Gedächtnisses vgl. Strube 1984).

Es wird davon ausgegangen, dass Kenntnisse über die Regeln und Gesetzmäßigkeiten, nach denen Begriffe miteinander in Beziehung stehen, den Schlüssel für die Entwicklung von such- und findefreundlichen Anordnungen liefern. Wie in der Problemanalyse dargestellt, wird eben davon ausgegangen, dass Nutzer ihre Erwartungen, wo sie ein Objekt finden, auf der Basis ihrer Begriffsstrukturen bilden. Es wird ferner davon ausgegangen, dass eine Kenntnis über die Regeln und Gesetzmäßigkeiten, nach denen Begriffe miteinander in Beziehung stehen, Hypothesen entwickeln lässt, wie Gegenstände im Raum anzuordnen, sprich in Beziehung zu bringen sind, damit sie den Erwartungen der Nutzer entsprechen. Ziel dieses Abschnittes ist es deshalb, eine Theorie (bzw. ein Modell) zu finden, die differenzierte, empirisch gestützte Beschreibungen und Erklärungen liefert auf die Fragen: Nach welchen Regeln und Gesetzmäßigkeiten sind die Begriffsstrukturen aufgebaut? Welche Relationen repräsentieren Begriffsstrukturen?

Seit Beginn der 1970er Jahre entstand eine Vielzahl von Begriffsstruktur-Theorien. Sie unterscheiden sich a) in Art und der Anzahl der angenommenen Begriffsbeziehungen (Relationen), b) in den Untersuchungsmethoden sowie c) in dem Ausmaß, in dem sie Vorhersagen erlauben. Im Rahmen dieser Arbeit ist es nicht möglich, alle Theorien und Modelle darzustellen und zu bewerten. Es werden nur fünf untersucht. Sie decken die Unterschiedlichkeit der Modelle gut ab und sind überdies in der Literatur häufig erwähnt worden. Zur Zeit populäre Modelle (vgl. Städtler 1998, S. 584) der (neo-) konnektionistischen Netzwerke werden hier nicht untersucht, weil sie (bislang) nicht in der Lage sind die Bildung komplexer Begriffe zu beschreiben.

Drei der folgenden fünf Begriffsstruktur-Modelle bauen nicht automatisch auf den Modellen über Begriffe auf. Das Modell der semantischen Netzwerke (Collins & Quillian 1972), das Modell der propositionalen Repräsentation (Anderson & Bower 1973) bzw. das HAM, die ACT-Modelle von Anderson oder die Schemata-Modelle (Minsky 1975, Schank & Abelson 1977) haben keine expliziten Modelle über Begriffe. Das ver-

haltensorientierte Merkmalsrepräsentationsmodell von Klix (1971, 1992) und Hoffmann (1983, 1986) hingegen baut systematisch auf der Theorie über Begriffe auf. Die Analyse geschieht entsprechend der Analyse der Begriffstheorien nach folgender Struktur: Basisannahmen - Experimente – Kritik – Bewertung hinsichtlich des Anordnungsproblems nach den vier oben genannten Kriterien.

4.4.1 Semantische Netzwerke

Eines der ersten Modelle zur Repräsentation begrifflichen Wissens wurde von Collins & Quillian (1972) mit dem Modell der Semantischen Netzwerke unterbreitet (Hoffman 1986, Anderson 1996 und Opwis & Lürer 1997). Quillian hatte sein Modell zunächst als lauffähiges Computerprogramm zur Speicherung und zum Abruf begrifflicher Informationen konzipiert; einen psychologischen Anspruch vertrat das Modell zu Anfang nicht (Quillian 1968, S. 216, zitiert nach Knauff 1997).

Basisannahmen Erste Basisannahme des Modells ist, dass das Wissen in Begriffen gespeichert ist, die durch Relationen untereinander verbunden sind. Es entsteht somit ein Netzwerk, dessen Knoten (tokens), die Begriffe und dessen Kanten die Relationen (types) sind. Collins & Quillian unterscheiden zwei Relationsarten: Ober-Unterbegriffsrelationen (superordinate associations) und Eigenschaftsrelationen (property associations). Als Folge entsteht ein hierarchisches Netzwerk, das begriffliches Wissen in Knoten repräsentiert und durch die Kanten die semantische Ähnlichkeit zu anderen Knoten repräsentiert. *Zweite Basisannahme* ist das Speicherökonomie-Prinzip: Dieses Prinzip nimmt an, dass alle Informationen, die ableitbar sind, nicht stationär gespeichert, sondern im Bedarfsfall auch abgeleitet werden. Merkmale, die einen Oberbegriff definieren, gelten dementsprechend auch für den Unterbegriff. Sie werden nicht noch einmal separat beim Unterbegriff repräsentiert. Sie sind nur durch die bestehende Relation repräsentiert. *Dritte Basisannahme* ist die Voraktivierung, die im Priming-Paradigma mündet: Bei der Aktivierung eines Knotens werden die durch Kanten verbundenen anderen Knoten mit aktiviert.

Beispiele Das Gesagte soll am Beispiel der Begriffsstruktur von Tieren verdeutlicht werden (vgl. Abbildung 4.6). Die einzelnen Knoten stellen Begriffe dar, wie Tier, Vogel, Fisch oder Kanarienvogel. Jedem Knoten sind Eigenschaften zugeordnet. Der Begriff TIER repräsentiert Eigenschaften, z.B.: hat eine Haut, kann sich bewegen, frisst, atmet. Der Begriff VOGEL repräsentiert Eigenschaften: kann fliegen, hat Federn und Flügel. Zwischen den Knoten bestehen Verbindungen, z.B. eine Ober-Unterbegriffsbeziehung, wie sie etwa zwischen Tier, Vogel und Kanarienvogel besteht. Die Eigenschaften, die den Vogel als Tier definieren, sind nicht noch einmal separat beim Vogel repräsentiert, sondern nur durch die Oberbegriffsrelation zum Begriff Tier. Eigenschaften, die für eine höhere Ebene zutreffen, gelten somit auch für die Ebenen darunter. Dem Modell entsprechend benötigen Unterbegriffe auch längere Erkennungszeiten als ihre Oberbegriffe.

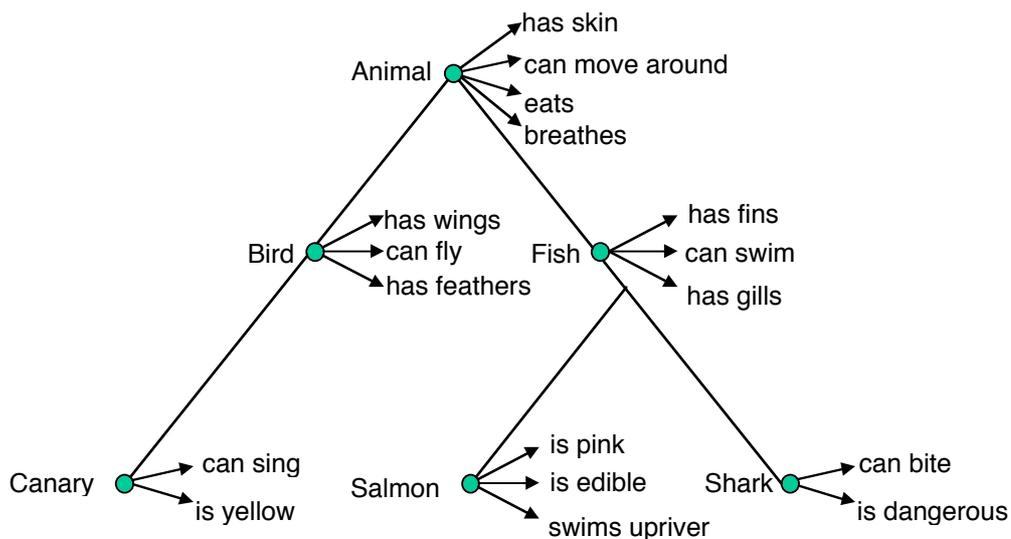


Abbildung 4.6. Visualisierung der Repräsentation von Wissen in einem semantischen Netzwerk (Collins & Quillian 1972)

Experimente Das Modell der semantischen Netzwerke wurde auf Ergebnissen von Satzverifikationsaufgaben gegründet: Versuchspersonen verifizierten die Aussage „Ein Vogel hat Federn“ schneller als die Aussage „Ein Kanarienvogel hat Federn!“ oder die Aussage „Ein Kanarienvogel ist gelb.“ schneller als die Aussage „Ein Kanarienvogel kann fliegen.“ Der Zeitunterschied wurde dahingehend interpretiert, dass die Eigenschaften auf unterschiedlichen Begriffsebenen gespeichert werden: Spezifische Eigenschaften des Kanarienvogels werden beim Begriff Kanarienvogel abgespeichert. Eigen-

schaften, die für Vögel im Allgemeinen gelten, sind – entsprechend dem postulierten Speicherökonomie-Prinzip – beim Begriff Vogel abgespeichert. Die Prüfung der Aussage „Ein Kanarienvogel kann fliegen.“ dauert deshalb länger als die Prüfung „Der Kanarienvogel ist gelb.“, weil diese Information aus dem Oberbegriff Vogel abgeleitet werden muss.

Neben Satzverifikationsaufgaben traten Priming-Aufgaben, denen das bereits erwähnte Priming-Paradigma zugrunde liegt: Die Buchstabenfolge H-a-i wird schneller als Wort „Hai“ identifiziert, wenn vorher das Wort „Fisch“ dargeboten wird. Dies wird damit erklärt, dass die Begriffe im Gedächtnis eng miteinander verbunden sind. Bei der Aktivierung des Begriffes Fisch wird der Begriff Hai mitaktiviert. Je schneller die Personen nach der Darbietung des Prime-Reizes den Zielreiz (target) identifizieren können, desto stärker sind die entsprechenden Begriffe im Gedächtnis miteinander verbunden.

Kritik Das Modell des semantischen Netzwerkes wurde stark kritisiert. So konnte das Modell der semantischen Netzwerke zwar gut darstellen, dass ereignisbezogene Begriffsbeziehungen wie Chirurg - Skalpell näher repräsentiert sind als Skalpell - Taschenmesser. Doch es konnte nicht mehr erklären, warum Stein schneller als Oberbegriff zu Diamant erkannt wird als Juwel, oder Tiger schneller als Oberbegriff für Maus erkannt wird als Nager (vgl. Klix 1988). Ein anderes, wichtiges Argument bezieht sich darauf, dass das Modell der semantischen Netzwerke als Computerprogramm nur für sprachliche Inhalte (Wortbedeutungen) entwickelt wurde, aber nicht für anschauliche, vorsprachliche Prozesse.

Der kanadische Psychologe Tulving kritisierte, dass das Modell individuelles Wissen unberücksichtigt lasse. Er formulierte als Erster den Begriff des „episodischen“ Wissens (Tulving 1972, zitiert nach Velickovskij 1988). Episodische Gedächtnisinhalte beziehen sich auf einmalige, raum-zeitliche Ereignisse, die im Gegensatz zu semantischen Wissensstrukturen einen ausdrücklichen autobiographischen Bezug haben; z.B. das persönliche Wissen über den eigenen Schreibtisch. Dieses Wissen kann das Modell der semantischen Netzwerke nicht integrieren. Hoffmann (1993) kritisierte, dass die Art der Relationen (Kanten) nicht eindeutig definiert und empirisch untersucht sei.

Eine weitere Kritik bezieht sich auf das *Abruf-Interferenz-Paradoxon*: Entsprechend diesem Paradoxon müssten Netzwerke umso langsamer arbeiten, je mehr Informationen in dem Netz gespeichert sind (Klimesch 1988). Dies stimmt jedoch mit experimentellen Befunden nicht überein. Wie bereits dargestellt, haben Primärbegriffe die kürzesten Erkennungszeiten, unabhängig davon, ob sie mit Ober- oder Unterbegriffen verglichen werden. Die Ergebnisse von Eleanor Rosch bezüglich Primärbegriffen lassen sich somit in das Modell des semantischen Netzwerkes nicht einordnen, sie widersprechen auch noch den Modellannahmen.

Collins und Loftus (1975) revidierten daher mit ihrem *Modell der Ausbreitenden Aktivierung* (spreading-activation-theory) das Modell der semantischen Netzwerke. Sie führten die zusätzliche Annahme ein, dass die verbindenden Relationen zwischen den Knoten unterschiedlich stark bzw. eng sein können. Sie nahmen an, dass bei der Aktivierung eines Knotens – wie in einer Welle – benachbarte Knoten mitaktiviert werden, entsprechend der Stärke der Verbindung. Und sie revidierten die Annahme des Speicherökonomie-Prinzips: sie ließen Eigenschaftsrepräsentationen von Begriffen sowohl beim Oberbegriff als auch beim Unterbegriff zu. Somit konnten sie zwar die Erkenntnisse von Eleanor Rosch beschreiben, aber sie konnten sie nicht erklären.

Hoffmann (1986) kritisiert, dass diese Netzwerkmodelle (z.B. Collins & Quillian 1969, Collins & Loftus 1975) davon ausgehen, dass a) Beziehungen zwischen Begriffen nur bestehen, wenn sie als Kanten gespeichert sind und dass b) die Begriffe nach dem Hierarchieprinzip immer ihren Oberbegriffen zugeordnet sind. Es erscheint aber nahezu unmöglich, dass alle möglichen Arten von Beziehungen, die zwischen Begriffen herstellbar und auch erfindbar sind (z.B. auch metaphorische Beziehungen wie „eine Haut wie ein Pfirsich“) in einem Gewirr eines Netzes repräsentiert sind, in dem von einem Knoten Hunderte von Kanten zu anderen Knoten gehen (Hoffmann 1986).

Bewertung hinsichtlich des Anordnungsproblems Das Modell der semantischen Netzwerke ist in Hinblick auf das Anordnungsproblem wenig geeignet. Die Untersuchungen wurden nur mit sprachlichem Material, wenngleich auch mit natürlichen Begriffen gemacht. Phänomene wie Kreuzklassifikationen/ Mehrfachzuordnungen, wie sie bei natürlichen Begriffen vorkommen, kann das Modell nicht erklären. Es ist wenig differen-

ziert und scheint nur lexikalisches, kategoriales Wissen zu modellieren. Es spezifiziert die Art der Verbindungen zwischen den Begriffen nicht. Des Weiteren erlaubt es keine Vorhersagen über die begriffliche Nähe von so alltäglichen und vielschichtigen Begriffen wie Lebensmitteln oder Geschäftstypen.

Das Priming-Paradigma ist im Hinblick auf das Anordnungsproblem von Bedeutung. Bei der Darbietung eines Prime-Reizes werden semantisch nahe Begriffe voraktiviert. Die Kürze der Pause zwischen Prime- und Targetreiz sowie die Schnelligkeit, mit der Versuchspersonen dann den Target-Reiz identifizieren, wird als Maß für eine begriffliche Nähe gewertet. Kurze Zeiten sprechen für eine begriffliche Nähe. Dies wird als Beleg gewertet, dass Begriffe nicht isoliert oder in beliebiger Beziehung, sondern in einer definierten Beziehung stehen.

Die Priming-Befunde sprechen dafür, dass Begriffe über eine definierte semantische Nähe bzw. Distanz verbunden sind. Entsprechend der Annahme, dass eine Anordnung von Objekten im Raum durch relative räumliche Distanzen beschrieben werden kann, lässt dies die Schlussfolgerung zu, dass über Priming-Experimente die relative semantische Nähe bzw. Distanz zwischen Begriffen ermittelt werden könnte und als relationales Maß für die räumlichen Distanzen der Objekte dienen könnte. Der Problemlösungsansatz könnte somit lauten, dass die relative räumliche Nähe von Objekten der semantischen Nähe der Objektbegriffe entsprechen sollte. Das Priming-Paradigma könnte somit einen Problemlösungsweg aufzeigen. Es liefert aber keine prognostischen Hypothesen darüber, durch welche Regeln und Gesetzmäßigkeiten die semantische Nähe von Begriffen bestimmt wird. Es müssten daher alle Objekte, die im Raum anzuordnen sind, in Paarvergleichen durch ein Priming-Experiment verglichen werden, um die begrifflich semantische Nähe zu ermitteln. Dies erfordert Laborbedingungen und einen experimentellen Paarvergleich der zu untersuchenden Begriffe. Ein unökonomisches, umständliches und unelegantes Vorgehen. Es wird daher weiter nach einer Theorie gesucht, die Aussagen darüber trifft, nach welchen Regeln Begriffsstrukturen aufgebaut sind. Diese Regeln sind notwendig, um Prognosen über die Nähe von Begriffen zu formulieren.

4.4.2 Propositionale Netzwerke und Andersons ACT-Modelle

Anderson & Bower (1973), Kintsch (1974) und Lindsay & Norman (1972) entwickelten Netzwerkmodelle, nach denen Wissensinhalte propositional repräsentiert sind. Das Modell der propositionalen Netzwerke basiert auf Untersuchungen zum Sprachverstehen und zum Sprachgenerieren. Es fehlt ihm ein eindeutiges, empirisch fundiertes Modell über Begriffe.

Basisannahmen und Beispiele Propositionale Netzwerke sollen zum Beispiel erklären, warum Versuchspersonen schon nach kurzer Zeit nicht mehr unterscheiden können, welchen der beiden Sätze sie gehört haben (zitiert nach Anderson 1996):

o „Nixon schenkte Breschnew, dem Staatschef der UdSSR, einen schönen Cadillac.“

o „Breschnew bekam von Nixon einen Cadillac geschenkt.“

Propositionen beziehen sich – wie das Beispiel zeigt – auf den (deklarativen) Inhalt einer sprachlichen Äußerung, der unabhängig von der Syntax oder Grammatik ist. Propositionen sind abstrakte Wissenseinheiten. Sie sind keine Worte, Wortfolgen und auch keine Sätze.

Der Begriff der Proposition wurde aus der Formalen Logik (Prädikatenlogik) und der Linguistik übernommen und geht zurück auf die im Rahmen der generativen Semantik entwickelten Kasusgrammatik von Fillmore (1968 zitiert nach Velickovskij 1988). Hier sind propositionale Darstellungen entwickelt worden, um das Bedeutungsverstehen von Sprache zu erklären. Mittels Grammatik und Syntax (vgl. Chomsky) kann z.B. nicht erklärt werden, warum Personen sich weniger den Wortlaut von Aussagen merken, als vielmehr die konzeptuelle Information (vgl. Sachs 1967 zitiert nach Velickovskij 1988). Oder wie Personen Bedeutungen von Worten und Aussagen erschließen und z.B. identische Worte, je nach Kontext in ihrer Bedeutung, unterschiedlich interpretieren.

Jede Proposition besteht aus Begriffen, die miteinander verbunden sind. Dem Satz “Der Hund spielt mit dem Ball” liegt z.B. die folgende Proposition zugrunde: Spielen (Hund, Ball). Der zentrale Begriff ist das “Prädikat” der Proposition (Spielen) – meist ein Verb. Ein Prädikat ist durch Relationen mit den Argumenten (Hund, Ball) verbunden. Die Argumente können einfache oder komplexe Begriffe sein, die selbst wieder als

Propositionen darstellbar sind. Eine Proposition wird deshalb auch Prädikat-Argumenten-Struktur genannt. Ein Prädikat lässt sich als Rahmen beschreiben, der eine Anzahl von Leerstellen für Argumente bereitstellt und diese einschließen kann. Es lassen sich folgende Leerstellen für folgende Argumenttypen unterscheiden (Norman & Lindsay 1972): Agent (Verursacher bzw. Handelnder), Rezipient (Empfänger, Person, auf welche das Ereignis einwirkt), Objekt (Gegenstand, auf den eingewirkt wird), Instrument (Ursache bzw. Stimulus des durch das Verb ausgedrückten Ereignisses), Lokation (Ort, an dem das Ereignis stattfindet) und Zeit (Zeitpunkt, an dem das Ereignis stattfindet). Eine empirische Bestätigung der Argumente Agent, Rezipient, Objekt und Instrument gibt Shafto (1973). Er analysierte die Argumente mit Hilfe einer multidimensionalen Skalierung und fand heraus, dass sich diese vier Argumente als unterschiedliche Kombinationen auf den beiden Dimensionen "Belebtheit" und "Aktivität" interpretieren lassen.

Propositionen können ihrerseits ein Netzwerk bilden. Folgendes bekannte Beispiel von Anderson (1996) soll eine propositionale Netzwerkstruktur veranschaulichen. Der Satz „Nixon schenkte dem Staatschef der UdSSR einen schönen Cadillac.“ lässt sich in drei Propositionen zerteilen: Nixon, Breschnew (Staatschef, USSR), schenken (Cadillac, Nixon, Breschnew). Diese drei Propositionen lassen sich in einer propositionalen Netzwerkstruktur darstellen, wie Abbildung 4.7 zeigt:

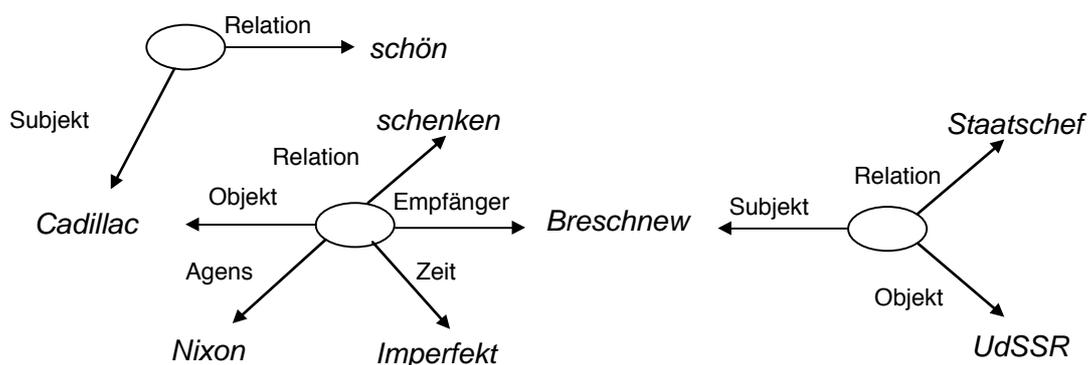


Abbildung 4.7. Netzwerkrepräsentation der Propositionen, die dem Satz zugrunde liegen „Nixon schenkte dem Staatschef der UdSSR einen schönen Cadillac.“ (zitiert nach Anderson 1996)

Es wird angenommen, dass beim Verstehen Propositionen wie Rahmen bereitstehen, die durch Gesprächsinformationen ausgefüllt werden. Die semantische Struktur, also die Bedeutung, wird durch die Verbindungen zwischen Propositionen bestimmt.

Kritik Eine wesentliche Kritik zielt auf die mangelnde Begründung der Relationen zwischen den Begriffen. Die Theorie der Propositionalen Netzwerke liefert keine Belege für die Art und Anzahl der unterschiedlichen Relationen. Hoffmann kritisiert (1993 S. 136), dass die Netzstrukturen lediglich auf Plausibilitätsüberlegungen beruhen würden, und dass - obwohl angenommen wird, dass sich die Kanten unterscheiden - nicht spezifiziert wird, worin die Unterschiede bestehen. Hoffmann (1993 S.137) führt an, dass nur zwei Arbeitsgruppen Untersuchungen zur empirischen Unterscheidung begrifflicher Relationen erarbeitet haben: Chaffin & Herrmann (1984, 1987) sowie Klix und van der Meer (1984, 1988).

Bewertung hinsichtlich des Anordnungsproblems Propositionale Netzwerkmodelle betonen ebenso wie semantische Netzwerke den Strukturaspekt von Wissensrepräsentationen. Die Befunde, die im Rahmen dieses Modells gemacht wurden, schließen sich der Erkenntnis an, dass Wissensinhalte nicht isoliert, sondern hoch geordnet repräsentiert sind. Die propositionalen Netzwerkmodelle treffen in begrenztem Umfang Aussagen über den Aufbau der Wissensstruktur: Der zentrale Begriff einer Proposition ist das Prädikat (meist ein Tätigkeitswort oder aber Zustandswort), das durch Relationen mit Argumenten (Begriffen über Personen, Objekten, Ort, Zeit etc.) und mit Ober- und Unterbegriffen verbunden ist. Tätigkeiten haben somit in propositionalen Netzwerken eine stark organisierende Bedeutung für die horizontale, begriffliche Strukturbildung. Oder anders formuliert: Die strukturelle Nähe von „Begriffen“ (Argumenten) wird durch Tätigkeiten (Prädikate) definiert. Inwieweit dieser Satz jedoch im Vergleich zu hierarchischen Beziehungen zutrifft, bleibt unklar. Es wird nicht untersucht, ob alle Relationen in Funktion und Wirkung gleich zu bewerten sind, oder ob es zwischen den Relationen Unterschiede gibt. Insbesondere hinsichtlich horizontaler und vertikaler (Ober-, Unterbegriffe) Begriffsbeziehung wäre diese Frage von Interesse: Haben Ober-, Neben-, Unterbegriffsrelationen (z.B. MEHLPRODUKTE, MEHL, GRIES, NUDELN, SPAGETTI) eine größere semantische Nähe als die Relationen von Objektbegriffen, die in einer Proposition über einen Tätigkeitsbegriff verbunden sind, wie Kuchen backen: MEHL-ZUCKER-

BACKPULVER sowie Spaghetti kochen: SPAGHETTI – TOMATENMARK – DOSENTOMATEN? Folgt die semantische Nähe bzw. Distanz zwischen Begriffen definierten Regeln oder bestimmt sich die semantische Nähe, sprich die Stärke der Relation, bei jeder Begriffsbeziehung individuell? Das Modell der propositionalen Netzwerke liefert zwar mehr Aussagen auf die Frage, wie Begriffsstrukturen aufgebaut sind als das Modell der semantischen Netzwerke, doch es tauchen neue, wichtige Fragen auf, die nicht beantwortet werden, weil die Relationen nicht untersucht werden.

Das Modell beruht auf Untersuchungen über natürliche Begriffe im sprachlichen Kontext (Sprachverstehen etc.). Es kann aber Phänomene wie Kreuzklassifikationen und Primärbegriffe nicht erklären. Die Erklärungsschwäche beruht vielleicht darauf, dass das Modell zu sehr an sprachlichen, linguistischen Problemstellungen verhaftet ist, wie dem Problem des Textverstehens. Daher werden noch weitere Modelle untersucht, die auch mit anschaulichem Material experimentierten und Phänomene natürlicher Begriffe (z.B. Kreuzklassifikation und Primärbegriffe) erklären können und noch differenziertere Aussagen bezüglich der inneren Struktur von Begriffen machen.

Prüfung von Andersons „Addaptive Control of Thought“-Modell (ACT)

Anderson entwickelte diverse Varianten seines Adaptive Control of Thought Modells. Das Ziel dieser Modelle ist es, die Arbeitsweise der Wissensstrukturen im psychischen Gesamtgeschehen aufzudecken. Es sind so genannte globale Netzwerktheorien (hybride, regelbasierte Produktionssysteme), die das komplexe Zusammenwirken von Gedächtnis-, Wahrnehmungs- und Handlungsprozessen erklären sollen. Newell & Simon führten 1972 als eErste solche Modelle ein, um die Architektur des menschlichen kognitiven Apparats zu modellieren. Neben Anderson haben auch andere Forscher regelbasierte Wissensmodelle (vgl. PRISM von Langley & Neches, 1981, GRAPES von Sauer & Farrell 1982, beide zitiert nach Opwis & Lürer 1998, Kompartimentmodell von Klix 1984) formuliert, die aber hier nicht aufgeführt werden. Sie würden den Rahmen der Arbeit sprengen und sind hinsichtlich des Anordnungsproblems von untergeordneter Bedeutung. Andersons ACT-Modelle werden hier dargestellt, weil sie die bekanntesten sind und das Bild abrunden.

Anderson & Bower entwickelten 1973 das HAM-Modell (Human Associative Memory Modell), das auf propositionalen Netzwerken basiert. Nach diesem Modell sind sämtliche Wissensbestände im Langzeitgedächtnis durch ein weitläufiges Netzwerk propositionaler Bäume repräsentiert. Das HAM-Modell konnte – als Computerprogramm realisiert – zahlreiche Sätze erfolgreich repräsentieren (vgl. Mehrebenen- betrachtung von Theorien: Formalisierung, Implementation und Materialisierung; Opwis & Lür 1998). Experimente konnten diverse Hypothesen bestätigen: z.B. die Hypothese, dass die Wiederholung einer Proposition die Behaltensfähigkeit der Elemente der Proposition verbessert. Die Befunde zu natürlichen Begriffen, wie Kreuzklassifikationen und Primärbegriffe unterstützt aber das HAM-Modell von Anderson nicht. Anderson überarbeitete in der Folge das Modell und entwickelte diverse ACT-Modell-Varianten, die im Folgenden dargestellt werden. Zwischen 1972 und 1976 entwickelte Anderson die Modelle ACTE, ACTF und ACT*. Anderson überarbeitete auch das ACT*-Modell, weil es integrierte Prozesse von Wahrnehmung und Denken wie das Problem des begrifflichen Entdeckens nicht löste (Anderson 1993). 1993 veröffentlichte er das Modell ACT-R.

Basisannahmen Entsprechend seinen früheren Modellen hält Anderson an der von Ryle (1949) aus den Computerwissenschaften in die kognitive Psychologie übernommenen, aber nicht unumstrittenen Unterscheidung (z.B. Hoffmann 1993, Barsalou, 1989) von zwei Speichern im Langzeit-Gedächtnis fest, dem *deklarativen und dem prozeduralen Wissensspeicher*: im deklarativen Speicher ist das Faktenwissen gespeichert; davon getrennt sind im prozeduralen Speicher die Handlungsprogramme in Form von Produktionsregeln und Zielwissen gespeichert. Bei kognitiven Prozessen greifen die Produktionsregeln aus dem prozeduralen Speicher auf das deklarative Wissen zu.

Im Mittelpunkt dieses Modells steht die Behauptung, dass kognitive Fähigkeiten durch Produktionsregeln aus dem prozeduralen Speicher realisiert werden (Anderson 1993). Die Produktionsregeln sind die kleinsten Einheiten von Handlungsprogrammen. Sie stehen im Dienste der Ziele und damit des Handelns. Produktionsregeln sind organisiert in Wenn-Dann-Paaren (conditional-action pairs) im Sinne einer Vorschrift: „Wenn diese Bedingung vorliegt, dann handele so!“ Die Produktionsregeln greifen auf das deklarative Faktenwissen zu. Anderson (1993) beschreibt dies am Beispiel des Summierens von Zahlen: „Wenn Du Zahlen zusammenzählen willst, dann schreibe sie unterein-

ander. Wenn die Summe der Einer größer als 10 ist, dann schreibe nur die Einer und summiere den Zehner in der Zehnerspalte.“ Das Ziel zu Summieren zerfällt damit in mehrere Unterziele. Die Produktionsregeln sind um diese hierarchisch strukturierten Handlungsziele herum angeordnet. Hiermit betont Anderson den Handlungszielbezug der prozeduralen Wissensinhalte sowie der Informationsverarbeitungsprozesse.

Die Annahmen über die Organisation des Wissens im deklarativen (Fakten-) Wissensspeicher sind beim ACT-R Modell knapp und allgemein gehalten. Der Ausdruck Begriffe (bzw. Concepts) taucht in seiner ACT-R Modelldarstellung nicht auf. Anderson formuliert in seinem Modell, dass das Wissen im deklarativen Speicher in Chunks (= kognitiven Einheiten vgl. Miller 1956) zusammengefasst ist. Ein Chunk besteht aus einer beschränkten Anzahl von Elementen (er nimmt dabei ca. 3 Elemente an), die in einer spezifischen Relation zueinander stehen. Die Elemente können selber wieder Chunks sein. Komplexere Wissensstrukturen sind repräsentiert als Hierarchien von Chunks (Anderson 1993). Dabei sind die Chunks mit Spezifitäts- und Aktivationsparametern assoziiert.

Kritik Im Hinblick auf die Fragestellung dieser Arbeit sind die Annahmen über begriffliche Wissensstrukturen (chunks) zu undifferenziert. Andersons Modelle werden aber auch an anderen Punkten kritisiert: Eine Kritik zielt auf die Unterscheidung in einen deklarativen und prozeduralen Wissensspeicher. Sie sei (früher) für Computerverarbeitungsmodelle von Nutzen (gewesen), jedoch fragwürdig für psychologische Repräsentationsformen. Sie widerspricht diversen Befunden, die für eine Verflechtung von z.B. begrifflichem Wissen und kognitiven Prozeduren wie Handlungsprogrammen sprechen (vgl. Hoffmann 1993, Barsalou 1989, Kintsch 1988 zitiert nach Hoffmann 1995). Des Weiteren ist die Spezifikation bezüglich Produktionsregeln undifferenziert. Klix (1993) liefert sehr viel differenziertere Befunde über Produktionsregeln. Er unterscheidet vier Prozeduren (Vergleichen, Verdichten, Verketteten, Verkürzen) und sechs Operationen (Aktivieren, Hemmen, Ersetzen, Verändern, Abbilden und Umkehren), die an begrifflichen Wissensstrukturen ansetzen und neues Wissen generieren.

Bewertung hinsichtlich des Anordnungsproblems Das regelbasierte Wissensmodell ACT-R von Anderson wird hinsichtlich des Anordnungsproblems als wenig interessant

betrachtet: Dieses Modell versucht zwar das Zusammenspiel von Handlungs- (= prozeduralem Wissen) und Faktenwissen (deklarativem Wissen) zu erklären. Und so könnte man zunächst vermuten, Anderson könnte einen Beitrag auf die Frage liefern, nach welchen Regeln die „Produktionsregeln“ am Faktenwissen ansetzen, um Erwartungen über den Ort eines gesuchten Objektes zu bilden. Diese Hoffnung wird aber nicht erfüllt. Anderson zielt auf ein computer-simulationsfähiges Modell ab und liefert Aussagen für reduzierte Probleme (z.B. des Addierens). Die Erkenntnisse regen meiner Meinung nach nicht dazu an, sie auf komplexere, alltägliche Wissensstrukturen zu übertragen.

4.4.3 Schemata-Modelle

Schemata-Modelle wurden entwickelt, um alltägliche Wissensstrukturen über stereotype Ereignisse (EIN BAD NEHMEN, ESSEN GEHEN) oder räumliche Umgebungen (Badezimmer, Restaurant) zu beschreiben. Untersuchungen legten z.B. den Schluss nahe, dass es kategoriale Wissensstrukturen gibt, die beim Verstehen von Texten etc. mitwirken (Bartlett 1932). Bartlett verstand unter „Schematisierung“ eine aktive Organisation vergangener Erfahrung. Folgendes Beispiel soll verdeutlichen, wie sich beim Lesen von Texten Bedeutungen und Schlussfolgerungen automatisch erschließen, obwohl sie nicht explizit beschrieben werden. *„Peter und Susi gaben der Stewardess ihre Boardkarte. Dann gingen sie die Treppen hinunter zu dem Bus, der sie zu ihrem Flugzeug und in den Urlaub bringen sollte.“* Automatisch wird man beim Lesen schlussfolgern, dass Peter und Susi ein Ticket vorweisen mussten, um eine Boardkarte zu bekommen, dass sie ihr Urlaubsgepäck bereits aufgegeben haben etc. Der kundige Vielflieger vermutet außerdem, dass Peter und Susi wohl mit einer kleinen Fluggesellschaft oder an einem kleinen Flughafen abfliegen, wenn sie ein Bus ans Flugzeug bringt. Das Beispiel zeigt, dass beim Lesen Wissen aktiviert wird. Dieses Wissen repräsentiert nicht einen konkreten, zeitlich-räumlich einmaligen Flughafenbesuch, sondern repräsentiert allgemeines, abstrahiertes Wissen von Flughafennutzungen. Es sind die Voraussetzungen, um besagte Schlussfolgerungen zu ziehen.

Basisannahmen Ereignisbezogene Begriffsstrukturen werden – entsprechend der Annahmen – durch wiederholt erlebte Ereignisse geprägt. Sie werden als quasi-stabil be-

trachtet, weil sie auch wieder veränderbar sind. Bartlett (1932) nannte solche allgemeinen Wissensstrukturen *Schemata* (vgl. auch Anderson 1996). Der Schematabegriff stammt ursprünglich aus der Künstlichen Intelligenz und den Computerwissenschaften und entspricht dort bestimmten Typen von Datenstrukturen. In der Psychologie wurden Schemata differenziert in *Frames* (Minsky 1975), die statische, räumliche Situationen und Objektszenen repräsentieren (z.B. Haus, Zimmer, Stadt, Küche etc.) und *Scripts* (Schank & Abelson 1977), die Ereignisse und Ereignisfolgen repräsentieren (z.B. Restaurantbesuch, Kochen, Einkaufen etc.).

Die innere Struktur von Schematas wird wie folgt definiert (vgl. Anderson 1996): Schemata bestehen aus Attributen und Werten. Die Attribute fungieren dabei wie Leerstellen (=slots), die im konkreten Fall mit „Werten“ ausgefüllt werden. Das Beispiel in Abbildung 4.8 soll dies verdeutlichen (zitiert nach Anderson 1996).

Attribute/slots	Werte für das frame HAUS
Oberbegriff:	Gebäude
Teile:	Zimmer
Material:	Holz, Stein
Funktion:	Wohnraum für Menschen
Form:	rechteckig

Abbildung 4.8. Attribute und Werte für das frame Haus

Anderson unterscheidet Attribute, die den Charakter von Propositionen haben (z.B. dass Häuser dem Menschen als Wohnung dienen) und Attribute, die perzeptuellen Charakter haben (z.B. die Größe, die Farbe). Eine besondere Stellung misst er den Oberbegriff-Slots zu. Sie geben an, zu welcher übergeordneten Kategorie ein Schema gehört. Solange keine widersprüchliche oder einschränkende Information gespeichert wird, „erbt“ ein Schema die Attribute des Oberbegriffs (vgl. Speicherökonomieprinzip der Semantischen Netzwerke). Folgendes Beispiel soll dies verdeutlichen: beim frame Haus erbt das Haus vom Oberbegriff Gebäude die Attribute, dass es Wände und ein Dach besitzt, und dass es auf dem Erdboden steht. Diese Informationen sind im Schema Gebäude repräsentiert und nicht mehr im Schema Haus. Sie können durch den Oberbegriffs-Slot vom Schema HAUS abgeleitet werden.

Beispiele und Untersuchungen Scripts stellen stereotype Ereignisse dar (vgl. Schank & Abelson, 1977). In solchen Scripts treten stereotype Handlungssequenzen mit einer zeitlichen Dimension auf. Jedes Script besteht aus einer Reihe von Akten oder Szenen, die ihrerseits in kleine Einheiten zerfallen. Die Struktur und Verwendung von Scripts wird gerne am Restaurantbesuch von Schank & Abelson (1977) dargestellt.

Abbildung 4.9 zeigt die einzelnen Szenen dieses Scripts. Versuchspersonen wurden gebeten, Restaurantbesuche zu beschreiben. Die Szenen wurden durch Inhaltsanalyse gewonnen. Dabei zeigte sich eine hohe Übereinstimmung der verschiedenen Szenen. Das Restaurant-Script von Schank und Abelson zeigt aber die Abhängigkeit des Scripts vom kulturellen und zeitlichen Hintergrund: in Europa bleibt man zum Bezahlen am Tisch sitzen und in modernen Restaurants verwenden die Bedienungen bereits computergestützte Dateneingabegeräte, die per Infrarot-Übertragung die Bestellung in die Küche senden.

<p>Szene 1: Eintreffen Gast betritt Restaurant Gast sucht sich einen Tisch Gast entscheidet sich für einen Platz Gast geht zum Tisch Gast nimmt Platz</p>	<p>Szene 3: Essen Koch gibt Essen an Bedienung Bedienung bringt Essen zum Gast Gast isst</p>
<p>Szene 2: Bestellung Gast nimmt Speisekarte Gast liest Speisekarte Gast entscheidet sich für ein Gericht Gast ruft Bedienung Bedienung kommt zum Tisch Gast bestellt Essen Bedienung geht zum Koch Bedienung gibt Bestellung an Koch Koch bereitet Essen</p>	<p>Szene 4: Gehen Bedienung schreibt Rechnung Bedienung geht zum Gast Bedienung überreicht Rechnung Gast gibt Bedienung Trinkgeld Gast geht zur Kasse Gast bezahlt an der Kasse Gast verlässt Restaurant</p>

Abbildung 4.9. Das Script eines Restaurantbesuchs nach Schank & Abelson (1977) zitiert aus Anderson (1996)

Bower u.a. (1979 zitiert nach Velickovskij 1988) bestätigen in ihren Untersuchungen die organisierende Rolle von Scripts beim Erzählen von Restaurantgeschichten. Sie lie-

Ben ihre Versuchspersonen die wichtigsten Ereignisse bei einem Restaurantbesuch aufzählen und werteten die Ergebnisse mittels einer Inhaltsanalyse und Häufigkeitsauszählung aus. Auch Bartlett (1932) ließ seine Versuchspersonen Geschichten erzählen bzw. wiedererinnern. Morton & Byrne (1975 zitiert nach Velickovskij) ließen Versuchspersonen aus dem Gedächtnis Aktivitäten memorieren. So sollten Hausfrauen aufzählen, was sie brauchen, um ein Gericht zuzubereiten. Brewer & Treyens (1981) arbeiten mit anschaulichem Untersuchungsmaterial und zeigten, dass das Schema-Modell auch auf nicht-sprachliche Phänomene übertragbar ist. Sie ließen ihre Versuchspersonen in einen (Büro-) Raum gehen und danach das Gesehene niederschreiben. Es zeigte sich, dass die Personen sich auch an Gegenstände (wie Papier) „erinnerten“, die in dem konkreten Raum nicht vorkamen, die aber üblicherweise in einem Büroraum vorkommen.

Kritik Die Kritik an den Schemata-Modellen zielt vor allem darauf, dass ihnen keine gemeinsame Theorie zugrunde liegt. Sie seien zu unpräzise, zu statisch, um Weltwissen mit dem angemessenen Auflösungsgrad abzubilden und sie haben beschreibenden, aber keinen erklärenden oder gar prognostischen Wert (Herrmann 1982, 1993, Klix 1984, Hoffmann 1986, Velickovskij 1988, Opwis & Lür 1998). Aufgrund der Schemata-Modelle ist zwar eine Vielzahl an Ergebnissen über den Einfluss begrifflichen Wissens auf die Wahrnehmung, die Gedächtnisorganisation, das Verstehen und den Wissensabruf von Alltagsereignissen und -szenen untersucht worden. Die Schemata-Modelle können aber selbst nicht erklären oder gar vorhersagen, wie Schemata entstehen, warum bestimmte Attribute (z.B. Gegenstände) relevant für das Schema sind und andere nicht. Probleme liefert der Modellansatz bezüglich des Abstraktionswechsels: Klix (1992 S. 325) kritisiert, dass das Schema-Modell kein „Zooming“ in der Arbeit des Gedächtnisses erlaubt: Spielt die Art der Vorspeisen eine Rolle beim Restaurantschema? Spielt die Menükarte eine Rolle und was ist, wenn die Person vom Ober die möglichen Speisen erzählt bekommt oder die Person, ohne die Karte zu lesen, ihre Lieblingsspeise bestellt?

Probleme bereitet auch die Anzahl der Schemata: Haben Personen neben einem Restaurant-Schema auch ein Fastfood-, Kantinen-, Feinschmeckerrestaurant-Schema? Wie interagieren sie? Schank (1982) beantwortete die Kritik durch die Entwicklung seines MOP-Modell (Memory-Organization-Packages-Modell) und geht von einer hierarchischen Organisation aus. Es ist auch nicht geklärt, wie Handlungs- und Faktenwissen zu-

einander stehen. Wenn Scripts Ereignisse repräsentieren (z.B. Restaurantbesuch), wie sind dann Handlungsprogramme und Faktenwissen repräsentiert? Klix (1984) kritisiert demzufolge auch die ungenügende Trennung von Repräsentations- und Prozessaspekten. Auch Anderson (1996) kritisiert die unzureichende Trennung von prozeduralem und deklarativem Wissen. Beide gehen von einer strikten Trennung bei der Repräsentation von Handlungs- und Faktenwissen aus, weil sie einen prozeduralen und deklarativen Wissenspeicher unterstellen. Fasst man die Kritik zusammen, dann regen die Schemata-Modelle zu wichtigen Untersuchungen an, weisen aber in ihren Modellierungen große Mängel und geringen explikativen und prognostischen Wert auf.

Bewertung hinsichtlich des Anordnungsproblems Auch die empirischen Befunde, die im Rahmen der Schemata-Modelle entstanden, belegen, dass Wissen organisiert, in Strukturen repräsentiert ist. Somit bestehen neben hierarchischen Begriffsstrukturen auch Begriffsstrukturen von stereotypen Ereignissen („Scripts“) und Situationen („Frames“). Diese „Scripts“ und „Frames“ haben wiederum einen organisierenden Einfluss auf die Wahrnehmung, den Wissensabruf und das Verstehen.

Die Schemata-Modelle muten zunächst als geeignet hinsichtlich des Anordnungsproblems an. Vor allem „Frames“ könnten Aufschluss über die Anordnung von Objekten im Raum vermuten lassen. Auch sind die Untersuchungen an natürlichen Begriffen mit sprachlichem und anschaulichem Material gemacht worden. Das Modell liefert jedoch wenig differenzierte Aussagen, nach welchen Regeln diese Frames aufgebaut werden. Hinsichtlich einer Anwendung der Schemata-Modelle auf das Anordnungsproblem entstehen Probleme, die die Schwäche des Modells zeigen: Innerhalb von Schemata-Modellen ist nicht nur nicht geklärt, wie mehrere Scripts (z.B. Supermarkt-, Verbrauchermarkt-, Feinkost- und Tante Emma-Laden-Script) integriert sind; sie sagen auch nichts darüber aus, wie Frames (z.B. Supermarkt) und Scripts (z.B. Essen einkaufen) zueinander repräsentiert sind. Im Sinne eines Speicher-Ökonomie-Prinzips ist eine mehrfache Repräsentation wenig wahrscheinlich. An dem Beispiel zeigt sich, welche Probleme das Modell aufgrund der geringen Differenzierung mit sich bringt. Doch auch wenn man praktisch vorgeht und die typische Methode des Schemata-Modells verwendet und die Personen mittels Erzählaufgaben hinsichtlich ihrer Scripts und Frames z.B. zu Süßgebäck untersucht, die Ergebnisse integriert bzw. gegeneinander abgleicht,

würden sich Probleme ergeben. Dieses Vorgehen wäre in der Auswertung stark subjektiv belastet, es wäre aufwändig und zweifelhaft in den Ergebnissen aufgrund der modell-immanenten Probleme und der zu erwartenden methodischen Schwierigkeiten (Untersuchungsökonomie, Objektivität, Reliabilität und Validität der Ergebnisse). Deshalb wird von dem Schemata-Modell als theoretische Grundlage für das Anwendungsproblem Abstand genommen.

4.4.4 Hierarchische und ereignisbezogene Begriffsstrukturen (Klix et al.)

Friedhart Klix (1971, 1976, 1984, 1992, 1996) und Mitarbeiter entwickelten eine stringente und umfassende Theorie, sowohl über Begriffe als auch Begriffsstrukturen. Basis der Theorie sind dabei die bereits dargestellten, experimentell untersuchten Annahmen über Begriffe als Bausteine der Begriffsstrukturen. Damit unterscheidet sich diese Theorie deutlich von den bisher dargestellten Modellen, denen explizite Modellannahmen über Begriffe sowie deren experimentelle Untersuchung fehlen. Es wird noch einmal kurz an die wesentlichen Annahmen von Klix über Begriffe erinnert. Ein Begriff besitzt nach Klix (1984,1992) sensorische, hierarchische, ereignisbezogene, graphisch-phonemische Merkmale.

Über hierarchische (=innerbegriffliche) Beziehungsmerkmale kann ein Begriff Beziehungen zu anderen Begriffen eingehen, um hierarchische (taxonomische) Begriffsstrukturen zu bilden. (Begriffsklärung: Das Wort „innerbegrifflich“ meint nicht die Merkmalsstruktur innerhalb eines Begriffes, sondern die Beziehungen innerhalb einer Begriffshierarchie wie Baum, Birke, Eiche. Sie alle sind durch Ober- und Nebenbegriffsrelationen miteinander verbindbar, gehören zu dem Oberbegriff Pflanze und werden daher innerbegrifflich genannt.) Über ereignisbezogene (= zwischenbegriffliche) Merkmalsrelationen geht ein Begriff Beziehungen zu anderen Begriffen ein, um „Ereignisbegriffe“ (Klix nennt dies auch Geschehenstypen) über stereotype Ereignisse oder Situationen zu repräsentieren (sie sind mit Schemata vergleichbar). Im Folgenden werden zunächst die innerbegrifflichen Begriffsstrukturen (Taxonomien) dargestellt, daran anschließend die zwischenbegrifflichen Begriffsstrukturen (Ereignisbegriffe).

Basisannahmen und Beispiele für Hierarchische Begriffsstrukturen Begriffe beinhalten durch ihre Beziehungsmerkmale die Voraussetzung für die Bildung taxonomisch, hierarchischer Wissensstrukturen. Klix unterscheidet folgende Beziehungsmerkmale: Oberbegriffsrelationen (z.B. Dackel–Hund, Birke-Baum), Nebenordnungsrelationen (z.B. Eiche–Birke, Tisch-Stuhl, Koch–Bäcker), Kontrastrelationen (z.B. Himmel–Hölle, Berg-Tal, Hoch-Tief), Komparativrelationen (z.B. krank-siech, Wind-Sturm, laufen-rennen)

Klix (1984) nimmt an, dass diese hierarchischen Begriffsbeziehungen nicht direkt gespeichert, sondern nur indirekt über Adressen repräsentiert sind und im Bedarfsfall auf der Basis der innerbegrifflichen Beziehungsmerkmale abgeleitet werden. Im Gegensatz dazu seien Beziehungen zu Ereignisbegriffen explizit, quasi-stationär gespeichert. Klix stützt diese Unterscheidung von indirekter und direkter Speicherung auf zwei Argumente: Einmal auf das *Speicherökonomie-Prinzip*: Dieses Prinzip nimmt an, dass alle Informationen, die ableitbar sind, nicht stationär gespeichert, sondern im Bedarfsfall abgeleitet werden. Im Bedarfsfall wird also durch Vergleichs- bzw. Prüfprozeduren (Ähnlichkeitsbestimmung der Merkmale) taxonomisches, hierarchisches Wissen abgeleitet. Zum anderen führt er diverse experimentelle Untersuchungen an (van der Meer 1978 zitiert nach Klix 1984, Küchler 1991): Küchler (1991) konnte in Priming-Experimenten zeigen, dass Effekte für ereignisgebundene Begriffe bereits nach 200 ms Voraktivierungen nachweisbar sind; bei Ober-Unterbegriffs-Wortpaaren zeigte sich bei einer Darbietungszeit des Prime-Reizes von 800ms kaum ein Effekt. Diese Ergebnisse bestätigen die Annahme, dass ereignisgebundene Wortpaare einander assoziativ zünden, weil sie quasi-stationär repräsentiert sind. Unter-Oberbegriffserkennungen müssen jedoch erst hergeleitet werden. Sie beruhen auf Merkmalsvergleich, was prozedural aufwändiger ist, denn nach der Aktivierung müssen noch Vergleichsschritte durchgeführt werden.

Basisannahmen und Beispiele für ereignisbezogene Begriffsstrukturen Der Leichtigkeit und Flexibilität in der Schaffung immer neuer, flüchtiger, hierarchisch-taxonomischer Verbindungen zwischen Begriffen steht die Festigkeit der überdauernden Ereignisbegriffe gegenüber. Klix und Mitarbeiter haben äquivalent zum Schemata-Konzept das Konzept des Ereignisbegriffes (synonym: Geschehenstyp) entwickelt. Auch Ereignisbegriffe repräsentieren stereotype Ereignisse und Situationen. Es ist ein weitaus differen-

zierteres Modell über ereignisbezogene Begriffsstrukturen als die oben beschriebenen Schemata-Modelle (vgl. Klix 1984, van der Meer 1991).

Begriffsklärung Eigentlich müsste der Ereignisbegriff „Ereignisbegriffsstruktur“ heißen, weil er immer aus mehreren aufeinander bezogenen Begriffen besteht. Insofern kann der Begriff „Ereignisbegriff“ zu einer Begriffsverwirrung führen. Andererseits ist diese Bezeichnung gerechtfertigt, weil der Ereignisbegriff selbst eine Einheit bildet, die Informationen in einem hohen Maß organisiert, verdichtet und für kognitive Verarbeitungsprozesse zur Verfügung stellt. Ereignisbegriffe sind nach Klix (1984, 1992) direkt repräsentiert in Form ereignisbezogener Netzwerke. Ereignisbegriffe sind nach Klix (1984,1993) Klassenbildungen über raum-zeitlich kohärente Ereignisse bzw. Situationen. Ereignisbegriffe sind z.B. „Einkaufen“, „Lehren“, „Heiraten“, sowie „Gewitter“, „Börsenkrach“, „Konferenz“. Ereignisbegriffe haben im Kern meist ein Verb (nicht immer). Dieses Verb steht in Beziehung mit einer Anzahl von Begriffen über Personen und/oder Gegenständen. Dies entspricht in etwa dem Aufbau einer Proposition, die sich in ein Prädikat und in Argumente gliedert. Grundlage für die Bildung von Ereignisbegriffen sind – genau wie bei der Bildung von „einfachen“ Begriffen – Abstraktionsprozesse über invariante Merkmale und Merkmalskonstellationen bzw. Begriffe und Begriffskonstellationen.

Ein Ereignisbegriff besteht aus (vgl. Klix 1992, van der Meer 1991): a) einem zentralen Begriff = semantischen Kern, meist ein Verb, b) zwischenbegrifflichen (semantischen) Relationen, c) Objektbegriffen und d) einer assoziativ gebundenen Wortmarke. Dies soll am Beispiel des Unterrichtens verdeutlicht werden. Abbildung 4.10 zeigt den Ereignisbegriff Lehren. Der semantische Kern >Lehren< ist über zwischenbegriffliche Relationen verbunden mit ereignis-typischen Objektbegriffen: Handlungsträgerrelation Hat (Unterrichten-Lehrer), Finalitätsrelation (Unterrichten - Wissen Beherrschen), Rezipientenrelation (Unterrichten - Schüler), Instrumentalrelation (Unterrichten Lehren - Bücher), Objektrelation (Unterrichten - Schulbank), Ortsrelation (Unterrichten - Schule), Zeitrelation (Unterrichten -Vormittags).

Die Voraussetzung dafür, dass Begriffe untereinander Beziehungen eingehen können, um stereotype Ereignisse zu repräsentieren, schreiben Klix und Mitarbeiter den zwi-

schenbegrifflichen, ereignisbezogenen Merkmalsrelationen von Begriffen zu. Sie verweisen auf andere Begriffe, die im gemeinsamen Kontext erlebt werden.

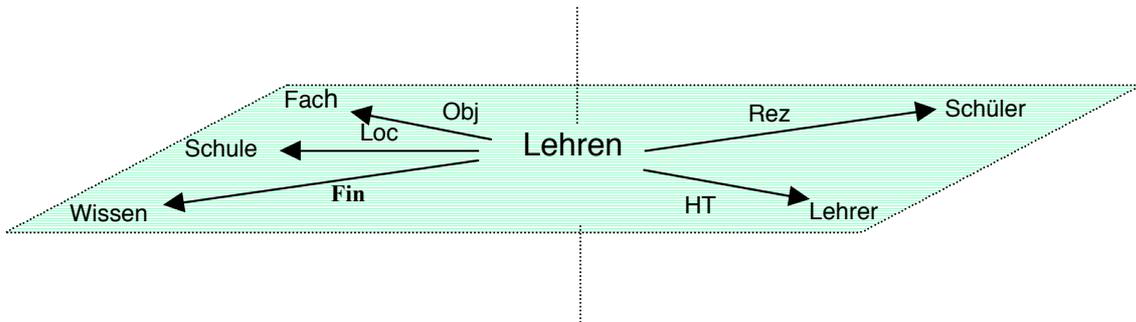


Abbildung 4.10. Visualisierung des Ereignisbegriffes „Lehren“ nach Klix 1992
HT = Handlungsträger, Rez = Rezipient, Obj = Objekt, Loc = Ort, Fin = Finalität

Hierarchien und Folgen von Ereignisbegriffen Ereignisbegriffe bilden Hierarchien, aber auch über Reihung (Sequenzen) Ereignisfolgen. Beide Fälle sollen kurz dargestellt werden: Sie bilden Hierarchien, über die den Begriffen immanenten hierarchischen Beziehungsmerkmale. Abbildung 4.11 zeigt, wie die drei unterschiedlich konkreten Ereignisbegriffe (Lehren), (Musik unterrichten), (Cello unterrichten) miteinander in einer hierarchischen Beziehung stehen.

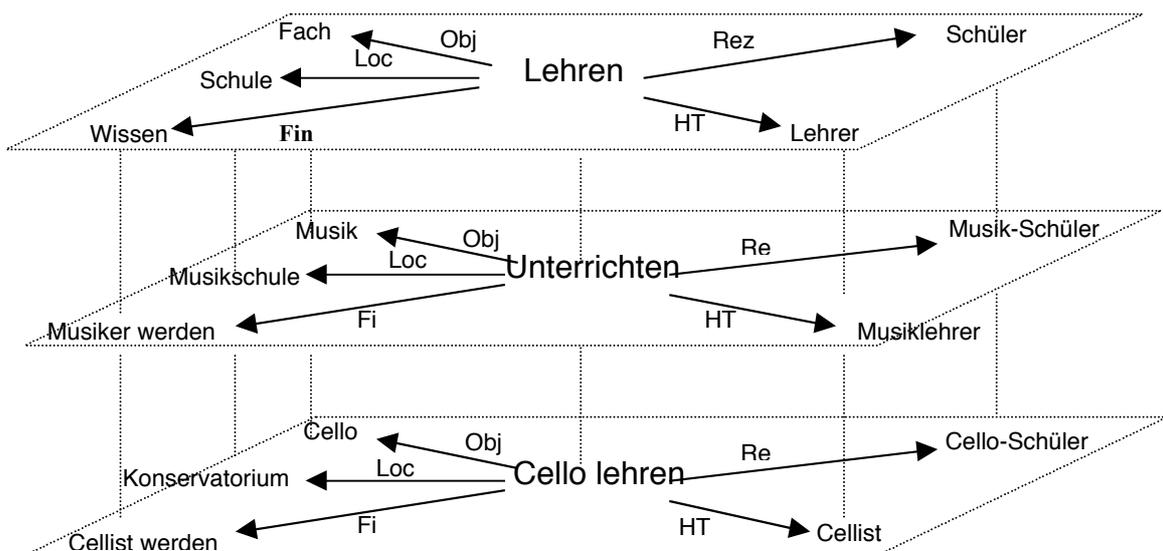


Abbildung 4.11. Beispiel für Ereignisbegriffe, die in einer hierarchischen Beziehung stehen. Die Übergänge zu den Unter- bzw. Oberbegriffen stellen die gepunkteten Linien dar. (zitiert nach Klix 1992 S. 243) HT = Handlungsträger, Rez = Rezipient, Obj = Objekt, Loc = Ort, Fin = Finalität

Ereignisbegriffe können auch miteinander verkettet sein, um Ereignisfolgen zu repräsentieren. Ereignisbegriffe sind dann über Relationen verbunden, die eine deutliche Zeitrichtung beschreiben. Klix unterscheidet 4 Relationstypen, um Ereignisfolgen zu bilden: 1) *Finalitätsrelation*, um das Ziel und den Zweck zu beschreiben; 2) *Konditionalitätsrelation*, um Vorbedingungen zu beschreiben; 3) *Kausalitätsrelation*, um Ursache-Folge-Beziehungen zu beschreiben; 4) *Konsekutionsrelation*, um das Eintreten von Konsequenzen zu beschreiben. Abbildung 4.12 zeigt eine Folge von Ereignisbegriffen, die durch Kausalitäts- und Konditionalitäts-Relationen miteinander verbunden sind:

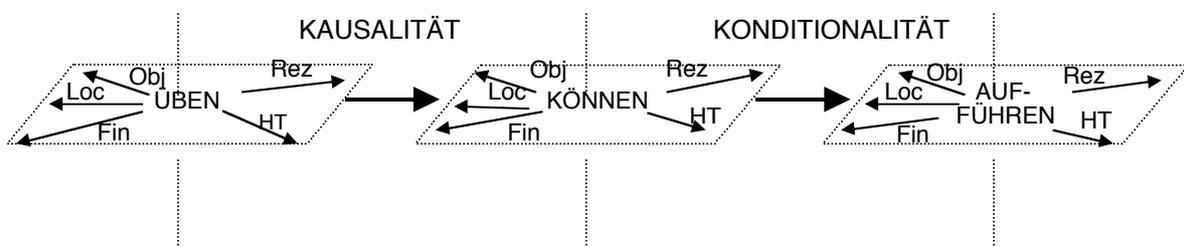


Abbildung 4.12. Beispiel für eine Verkettung von drei Ereignisbegriffen über Kausalitäts- und Konditionalitätsrelationen. Gestrichelte Linien nach oben bzw. unten zeigen Übergänge zu Ober- bzw. Unterbegriffen (Beispiel entwickelt nach Klix 1992, S. 247) HT = Handlungsträger, Rez = Rezipient, Obj = Objekt, Loc = Ort, Fin = Finalität

Ereignisbezogene Begriffsstrukturen (Anderson 1996) bzw. Ereignisbegriffen (=Geschehenstypen von Klix 1984) gelten als bedeutungsgebend, z.B. in der Wahrnehmung (bei der Objektidentifikation vgl. Hoffmann 1986), im Sprachverstehen (Schank & Abelson 1977), bei Problemlösungsprozessen (vgl. Klix 1971, Dörner 1995). Im Prozess des Erkennens interagieren Ereignisbegriffe (=Geschehenstypen) mit unmittelbar aufgenommenen Sinnesinformationen und lassen über top-down-Prozesse neue Reize begrifflich erfassen, d.h. in ihrer Bedeutung erkennen und verstehen. Um das Gesagte an einem Anordnungsproblem deutlich zu machen: Eine KundIn in einem Supermarkt erkennt und identifiziert demnach Produkte vor dem Hintergrund ihrer im Gedächtnis gespeicherten Begriffsstrukturen. Die Begriffsstrukturen haben demnach Einfluss auf die Erkennung und Identifizierung der Produkte. Die Auswahl von Ereignisbegriffen und das Verbinden ihrer Attribute mit den aktuellen Bedingungen einer beobachteten

Szene wird als Voraussetzung für ein Verstehen angesehen (Norman & Rumelhart 1975, Klix 1984, 1993, Anderson 1996).

Experimente Die unterschiedlichen inner- und zwischenbegrifflichen Relationen als beziehungsstiftende Relationen für taxonomische Begriffsstrukturen und Ereignisbegriffe und -folgen wurden an sprachlichem und visuell-bildlichem Material untersucht. Damit entzieht sich das Modell einerseits dem Vorwurf nur an linguistisch, sprachlichen Phänomenen überprüft worden zu sein. Andererseits wurden Relationen zwischen Begriffen nicht nur postuliert, sondern auch tatsächlich experimentell untersucht. Damit unterscheiden sich diese Modellaussagen grundlegend von denen anderer Modelle, wie den Schemata-Modellen, dem Semantischen und Propositionalen Netzwerk-Modell.

Analogieerkennungsaufgaben (mit und ohne semantischem Priming) waren dabei eine häufig verwendete Untersuchungsmethode, um die unterschiedlichen Relationsarten empirisch zu überprüfen (Analogieerkennungs-Paradigma). In sprachlich orientierten Experimenten wurden Versuchspersonen z.B. Wortpaare vorgelegt, die in unterschiedlicher Beziehung zueinander stehen (z.B. Oberbegriffs- versus Nebenbegriffsrelation: z.B. Eiche-Baum versus Eiche-Birke oder Instrumental- versus Objektrelation: Schreiben-Stift versus Schreiben-Brief). Die Versuchspersonen mussten die verschiedenen Relationen unterscheiden. Gemessen wurde die Reaktionszeit. Die Dauer wurde als Indikator für den kognitiven Aufwand interpretiert. Kürzere Zeiten werden dahingehend interpretiert, dass die Begriffsrelation gespeichert ist; längere, dass die Begriffsrelation erst gebildet werden muss (vgl. van der Meer 1978, Kukla 1981 zitiert nach Klix 1984). Die Relationsarten wurden auch an visuell-bildhaftem Material untersucht (z.B. Vergleichsprozesse zwischen analogen Mustern; Reschke 1977, Ziebler 1982 zitiert nach Klix 1984) oder in Experimenten mit Satz-Bild-Vergleichen: Die Versuchspersonen bekamen Sätze wie „Die Tür ist offen“ und ein Bild vorgelegt, das eine offene Tür darstellt und sollten die Übereinstimmung beurteilen. Gemessen wurde die Reaktionszeit.

Kritik Besondere Beachtung hat Klix zu seinen Arbeiten über Ereignisbegriffe (= Geschehenstypen; vgl. Städtler 1998) erhalten sowie für die experimentellen Untersuchungen zu den begrifflichen Relationen. Eine andere Arbeitsgruppe um Chaffin & Herrmann (1984, 1987 zitiert nach Hoffmann 1993), die auch begriffliche Relationen expe-

rimentell untersucht, gelangte ebenfalls zu der empirischen Unterscheidung von Kontrastrelation (Tag - Nacht), Ähnlichkeitsrelation (Wind - Sturm), Oberbegriffsrelation (Rose - Blume), Handlungsträger-Relation (Chauffeur - Auto) und Teil-Ganzes-Relation (Lenker - Fahrrad).

Bewertung hinsichtlich des Anordnungsproblems Die Theorie von Klix und Mitarbeitern liefert im Vergleich zu den bisher dargestellten Theorien und Modellen die differenziertesten Regeln und Gesetzmäßigkeiten bezüglich Begriffsstrukturen. Die Aussagen sind untersucht durch Laborexperimente u.a. mit Analogieerkennungsaufgaben (mit und ohne semantisches Priming), an natürlichen sowie künstlichen Begriffen, mit sprachlichem sowie anschaulichem Material. Die Befunde, die im Rahmen dieser Theorie erarbeitet wurden, belegen (ebenso wie die der anderen Theorien), dass Begriffe nicht isolierte Wissenseinheiten sind, sondern in definierter Beziehung stehen und Begriffsstrukturen bilden. Darüber hinaus formuliert die Theorie aber auch Gesetzmäßigkeiten über die unterschiedlichen Arten von Begriffsbeziehungen. Diese Gesetzmäßigkeiten lassen sich auf das Anordnungsproblem anwenden. Es lassen sich überdies Annahmen formulieren, wie eine nutzerorientierte Anordnung der Objekte zu erzielen ist.

Zentraler Begriff eines Ereignisbegriffes ist das Prädikat (meist Verb), das über zwischenbegriffliche Merkmalsrelationen mit anderen Begriffen quasi-stabile Beziehungen eingeht. Diese Beziehungen repräsentieren stereotype Handlungsereignisse und Handlungssituationen mit ihren typischen Gegenständen und Personen, dem Handlungsziel und -zweck sowie dem Handlungsort und -zeit. Als Beziehungen zwischen den Begriffen lassen sich Handlungsträger-, Finalitäts-, Rezipienten-, Instrumental-, Objekt-, Orts- und Zeitrelationen unterscheiden. In der Begriffsstruktur sind somit jene Begriffe in einer typischen strukturellen Beziehung zueinander repräsentiert, wie sie typischerweise von der Person erlebt werden. Da die zwischenbegrifflichen Beziehungen quasi-stabil gespeichert sind, wirken die sich ergebenden ereignisbezogenen Begriffsstrukturen selbst wie ein Begriff. Ereignisbegriffe können untereinander in Beziehungen stehen, um Ereignisfolgen zu repräsentieren. Auch hierfür lassen sich bestimmte Relationen im Experiment belegen.

Über hierarchische (= innerbegriffliche) Beziehungsmerkmale verweisen die Begriffe auf die entsprechenden Ober-, Unter-, Neben-, Steigerungs- und Gegensatzbegriffe und repräsentieren somit hierarchisch-taxonomisches Wissen. Diese hierarchisch-taxonomischen Wissensstrukturen scheinen implizit und nicht explizit gespeichert zu sein. Sie können bei Bedarf von den Begriffen abgeleitet werden.

Das Modell erlaubt eine Beschreibung und Erklärung der beiden prinzipiellen begrifflichen Ordnungsdimensionen: der ereignis-handlungsbezogenen Ordnungsstruktur und der hierarchisch-taxonomischen Ordnungsstruktur.

4.4.5 Begriffsstrukturen im Modell der Antizipativen Verhaltenssteuerung

Auch im Modell der Antizipativen Verhaltenssteuerung werden Annahmen über Begriffsstrukturen formuliert. Sie basieren auf Annahmen über Begriffe (siehe Abschnitt 4.2), sind aber wie diese nicht sehr differenziert.

Basisannahmen Hoffmann versucht mit seinen Modellannahmen sowohl die Ausbildung taxonomischer, d.h. hierarchischer Begriffsstrukturen zu erklären, als auch die Ausbildung von partonomischen, d.h. ereignisbezogenen Begriffsstrukturen:

Taxonomische, hierarchische Begriffsstrukturen bilden sich durch zunehmend präzisere *Verhaltensabsichten bzw. Zielvorstellungen* heraus und einer damit einhergehend präziseren Unterscheidung der Verhaltenskonsequenzen. „Wenn es bspw. nur darum geht, den Durst zu löschen, ist jedes Getränk recht. Geht es jedoch darum, die passenden Weine für ein Menü zu wählen, sind genauere Unterscheidungen zu treffen.“ (Hoffmann 1993 S. 149). Je präziser die Zielvorstellungen sind, desto mehr werden die Äquivalenzklassen eingeschränkt. Von den Getränken zu den Weinen, von den Weißweinen zu den französischen Weißweinen usw. Ebenso muss etwa ein Florist zwischen verschiedenen Blumenarten ständig unterscheiden, wo wir als gelegentliche Blumenkäufer mit einem Dutzend Kategorien auskommen (vgl. Zimmer 1984). Je unspezifischer die Verhaltensabsichten sind, desto großzügiger fallen die erlebten Äquivalenzen aus. Die Folge sind mächtige Äquivalenzklassen, wie die der Getränke, der Blumen etc. Je präziser die Verhaltensabsichten und -ziele sind, desto differenzierter sind Äquivalenzklassen und desto differenzierter werden die hierarchischen Begriffsstrukturen.

Partonomische (=ereignisbezogene) Begriffsstrukturen entstehen durch Verhaltensabfolgen, wobei ein Begriff als Ausgangsbedingung für den anderen steht. Als Beispiele für partonomische Relationen nennt er Begriffspaare wie Messer-Griff, Fahrrad-Klingel oder Operation-Skalpell. Die Entstehung von partonomischen Relationen, also Beziehungen zwischen Ganzheiten und ihren Teilen, begründet er mit folgenden Überlegungen: „Einzelne Verhaltensakte bewirken in den meisten Fällen nur lokale Veränderungen einer gegebenen Situation. Zumeist werden nur Details verändert. Unter diesen Bedingungen führt eine antizipative Verhaltenssteuerung zwangsläufig dazu, dass globale Charakteristika der Ausgangssituation mit lokalen Effekten der in ihr realisierten Handlungen antizipativ verbunden werden.“ (Hoffmann 1992 S. 149)

Kritik Im Rahmen der Theoriebildung sind die Aussagen über Begriffe und Begriffsstrukturen bislang wenig differenziert und meines Wissens empirisch noch nicht überprüft. Sie haben spekulativen Charakter, bieten jedoch interessante heuristische Ansätze, vor allem für die Erklärung natürlicher Begriffe und impliziter Begriffsbildung. Es fehlt eine differenzierte Auseinandersetzung und Integration bisheriger Erkenntnisse der Begriffsforschung. Hoffmann scheint seine theoretischen und empirischen Erkenntnisse über Begriffe und Begriffsstrukturen, die er vor dem Formulieren der Theorie der Antizipativen Verhaltenssteuerung entwickelt hat, noch nicht eingearbeitet zu haben. Meines Erachtens wäre dies möglich und könnte in eine interessante empirische Forschung münden.

Bewertung in Hinblick auf das Anordnungsproblem Eine Bewertung ist insofern kritisch, da das Modell noch wenig elaboriert ist. Die Annahmen über die Ausdifferenzierung taxonomischer Wissensstrukturen sind interessant: Durch einen Wechsel von Verhaltenszielen und -absichten kommt es zu einer Ausdifferenzierung des Wissens. Im Hinblick auf die Fragestellung dieser Arbeit ist daher anzunehmen, dass die Mächtigkeit der taxonomischen Begriffsstrukturen abhängig ist von den Handlungszielen (und Aufgaben), die eine Person hat. Oder anders gesagt: die Anzahl der Ober-, Unter-, Nebenbegriffe über einen Gegenstandsbereich ist von den Handlungszielen und Aufgaben der Person abhängig. Annahmen über partonomische Begriffsstrukturen sind karg. Hoffmann setzt diese Strukturen in ihrer Funktion mit den Ereignisbegriffen von Klix (1984) gleich (vgl. Hoffmann 1993 S. 149), die dort differenzierter dargestellt und überdies

empirisch untersucht sind. Dennoch liefert Hoffmann zumindest eine – wenn auch vorerst nur spekulative – Erklärung für eine lern- und verhaltensabhängige Bildung von Ereignisbegriffen. Im Hinblick auf diese Arbeit könnte daher gerade diese Verhaltensabhängigkeit von Interesse sein, nämlich die Annahme, dass Personen abhängig von ihren Aufgaben und Verhaltensfolgen mehr oder minder differenzierte und ausgebildete Ereignisbegriffe haben. Die Aussage, dass z.B. Personen, die nicht kochen, einen weitaus undifferenzierten Ereignisbegriff von Kochen, Küche etc. haben als Personen, die kochen, scheint zunächst banal, ist aber in ihren tatsächlichen Implikationen von großer Bedeutung. Entsprechend den lernabhängigen Annahmen von Hoffmann müsste daher eine Person, die kocht, weitaus differenziertere Ereignisbegriffe haben.

4.4.6 Überblick über das Ergebnis der Prüfung

Es wurden fünf Theorien bzw. Modelle über Begriffsstrukturen untersucht (siehe Abb. 4.13). Für jede der fünf Theorien bzw. Modelle wurden die Basisannahmen und die jeweilige Begriffsdefinition dargestellt, es wurden die wichtigsten Experimente sowie die Kritik und widersprechende Befunde angeführt.

Semantische Netzwerke	Propositionale Netzwerke	Schemata Modelle	Ereignisbezogene & hierarchische Begriffsstruktur	Antizipative Verhaltenssteuerung
Collins & Quillian 1972	Anderson & Bower 1973	Minsky 1975 Schank & Abelson 1977	Klix 1984	Hoffmann 1993
Netzwerk aus Kanten und Knoten Speicher-Ökonomieprinzip Priming-Paradigma	Proposition = Rahmen aus Prädikat & Kontext, Ereignis, Objekt, Ort	Schemata, Frames = Stereotype Ereignisse, bzw. Szenen	Geschehens-typen = Ereignisbezogene Strukturen & Hierarchische Strukturen = kategoriale Strukturen	Taxonomische Strukturen & Partonomische Strukturen

Abbildung 4.13. Überblick über die fünf untersuchten Theorien bzw. Modelle

Alle fünf Theorien wurden hinsichtlich ihrer Eignung zur Lösung des Anordnungsproblems nach vier Kriterien geprüft. Die Ergebnisse der Prüfung zeigt Abbildung 4.14. Die Prüfung erbrachte, dass sich die ereignisbezogenen und hierarchischen Begriffsstrukturen von Klix et al.(1971, 1993), Hoffmann (1983, 1986) et al. hinsichtlich der praktischen Fragestellung dieser Arbeit am besten eignen: Sie liefern differenzierte Aussagen auf die Fragen „Nach welchen Regeln und Gesetzmäßigkeiten sind Begriffsstrukturen inhaltlich aufgebaut?“, „Welche Informationsklassen beinhalten sie?“. Die Aussagen sind in Experimenten mit sprachlichem und anschaulichen, an natürlichen und künstlichen Begriffen weitestgehend belegt. Die Theorie scheint auf Begriffe des täglichen Lebens, wie Lebensmittel, Zimmerpflanzen, Geschäftstypen anwendbar. Und die Theorie ermöglicht nicht nur Erklärungen, sondern auch Vorhersagen über Begriffsstrukturen.

<div style="text-align: center;">Theorien und Modelle</div> <div style="text-align: right; padding-right: 10px;"> Beurteilungs- kriterien </div>	Seman- tische Netz- werke Collins & Quillians 1972	Propositio- nale Netz- werke Anderson & Bower 1973 Kintsch 1974	Schemata- Modelle Minsky 1975 Schank & Abelson 1977	Ereignis- und hierar- chische Be- griffs- strukturen Klix 1984	Antizipative Verhaltens- steuerung Hoffmann 1993
1. Differenzierte Begriffsdefinition	nein	ja	bedingt	ja	nein
2. Kompatibel für natürliche Begriffe	nein	ja	ja	ja	ja
3. Experimentell geprüft	bedingt	bedingt	ja	ja	?
4. Erklärungs- <u>und</u> Vorhersage geeignet	nein	nein	nein	bedingt	bedingt

Abbildung 4.14. Übersicht über die Prüfung der fünf Begriffsstruktur-Theorien

4.5 Zusammenfassung

In diesem Kapitel wurde nach einer Theorie gesucht, die sich eignet einen Lösungsansatz für die Entwicklung erwartungsgerechter Anordnungen zu formulieren und Vorhersagen für die semantische Ähnlichkeit von Objekten liefert. Hierfür wurden fünf Begriffstheorien untersucht: Exemplartheorien (z.B. Sternberg, 1969, Collins & Quillians 1972), Prototypentheorien (Franks & Bransford 1971, Rosch & Mervis 1975), gegenstandsorientierte Merkmalsrepräsentationstheorien (z.B. Smith, Shoben & Rips 1974), verhaltensbezogene Merkmalsrepräsentationstheorien (z.B. Klix 1992, Hoffmann 1983, 1986) und die Theorie der Antizipativen Verhaltenssteuerung (Hoffmann 1993). Des Weiteren wurden fünf Begriffsstruktur-Theorien untersucht: Semantische Netzwerke (Collins & Quillian 1969, Propositionale Netzwerke (Anderson & Bower 1973), Schemata-Modelle (Minsky 1975, Schank & Abelson 1977), ereignisbezogene und hierarchische Begriffsstrukturen (Klix 1984), partonomische und taxonomische Begriffsstrukturen (Hoffmann 1993).

Jede Theorie wurde nach folgenden Punkten dargestellt: Basisannahmen, Definitionen, Experimente, Kritik. Und jede Theorie wurde daraufhin geprüft: 1) wie differenziert die Definition von Begriff bzw. Begriffsstruktur ist, 2) ob die Theorie auch für natürliche Begriffe anwendbar ist, 3) ob die Theorie an sprachlichem und anschaulichem Untersuchungsmaterial experimentell überprüft wurde und 4) ob sich die Theorie auch zur Vorhersage eignet. Eine Übersicht über die Ergebnisse der Prüfung zu Begriffstheorien findet sich in Abschnitt 4.3.6 (S. 89-91) und zu Begriffsstrukturtheorien in Abschnitt 4.4.6 (S. 118-119).

Die verhaltensorientierte Merkmalsrepräsentationstheorie von Klix (1971, 1992) und Hoffmann (1983, 1986) zeigte sich als die geeignetste. Sie wird im nächsten Kapitel auf das Anordnungsproblem angewendet.

5. Anwendung der Theorie auf das Anordnungsproblem

Im folgenden Kapitel wird die Begriffs- bzw. Begriffsstrukturtheorie von Klix (1971, 1984, 1993), Hoffmann (1983, 1986) et al. auf das Anordnungsproblem angewendet: Im *ersten* Abschnitt werden drei Gesetzmäßigkeiten, nach denen begriffliches Wissen verarbeitet wird und die für das Anordnungsproblem von Bedeutung sind, a) herausgearbeitet, b) anhand von Experimenten belegt und c) auf das Anordnungsproblem angewendet. Ziel und Zweck dieses Vorgehens ist es, ein theoriebasiertes Untersuchungsdesign zu entwickeln und Prognosen in Form von Hypothesen für das Verhalten der Versuchspersonen zu entwickeln. Diese hier formulierten Gesetzmäßigkeiten finden sich nicht in gleichen Worten bei Klix und Hoffmann. Sie formulierten zwar eine Vielzahl von Aussagen, liefern aber nicht prägnant Gesetzmäßigkeiten. Deshalb ist es nötig, aus der Theorie für das Problem Gesetzmäßigkeiten zu extrahieren und spezifisch anzuwenden. Im *zweiten* Abschnitt erfolgt eine Zusammenfassung des Untersuchungsansatzes und der Hypothesen.

5.1 Anwendung von drei Gesetzmäßigkeiten auf das Anordnungsproblem

Im Folgenden werden drei Gesetzmäßigkeiten dargestellt und mit Experimenten belegt. Dann werden Konsequenzen für die Lösung des Anordnungsproblems gezogen.

1. Gesetzmäßigkeit: Begriffe haben eine spezifische Nähe bzw. Distanz

Die Annahme einer spezifischen semantischen Ähnlichkeit (Nähe) zwischen den Begriffen wurde durch eine große Anzahl von Experimenten bestätigt (vgl. Rips 1989, Städtler 1998). Alle in dieser Arbeit dargestellten Begriffsstruktur-Theorien arbeiten mit dieser Gesetzmäßigkeit und liefern empirische Belege hierzu. Es ist also keine Gesetzmäßigkeit, die allein von Klix et al. formuliert wird, aber sie *ist grundlegender Ausgangspunkt* ihrer Theorie. Im Folgenden soll diese Gesetzmäßigkeit belegt werden.

Empirische Belege Bereits die Ergebnisse erster Assoziationsexperimente (vgl. Galton 1879 zitiert nach Strube 1984; Ebbinghaus 1885, 1971) zeigten, dass die Nennungen der

Versuchspersonen bestimmten Regelmäßigkeiten folgen. Untersuchungen von Clark (1972; zitiert nach Klix 1992) ergaben, dass dann nach den Nennungen aus situationalen Kontexten diverse Phasen folgen, bei denen nur Unterbegriffe oder nur Oberbegriffe oder nur Gegensatzbegriffe (Antonyme) oder nur Reimworte assoziiert werden. Und auch aktuellere Untersuchungen zeigen, dass bei freien Assoziationen die Versuchspersonen inhaltlich vorzugsweise zunächst Begriffe nennen, die nahe um den Schlüsselreiz liegen. Nach diesen spontanen, automatischen Nennungen tritt eine Pause ein, um dann Begriffe aus weiteren situationalen Kontexten (vergleichbar mit Schemata, Skript, Frame bzw. Ereignisbegriffen) zu assoziieren (vgl. Strube 1984, Walker und Kintsch 1985 zitiert nach Klix 1992).

Die unterschiedliche semantische Nähe von Begriffen ist auch Basis des *Priming-Paradigmas* (Meyer & Schvaneveldt 1976 zitiert nach Städtler 1998), ein bis heute wichtiges Untersuchungs-Paradigma in unterschiedlichen psychologischen Forschungsbereichen. Bei einem Priming-Experiment erhält die Versuchsperson zuerst einen Prime-Reiz (=Anregungsreiz; z.B. ein Abbild, ein Wort oder eine Buchstabenfolge). Die Zeit der Darbietung wird so dosiert, dass ein Lesen und Bedeutungsverstehen möglich ist. Nach der Darbietung erfolgt eine kurze Pause – SOA: Stimulus-Onset-Asynchronie – von z.B. 200ms, 1000ms oder 2500ms. Dann erfolgt der kritische Ziel- oder Target-Reiz oder ein Distraktor (Kontrollreiz). Die Versuchsperson muss nun so schnell wie möglich mit „ja“ oder „nein“ entscheiden, ob der zweite Reiz die Bedingungen der Instruktion erfüllt, z.B. ob die dargebotenen Buchstaben ein Wort darstellen. Festgehalten werden die Reaktionszeit und die Fehler. Variiert werden der Prime- und der Target-Reiz und die Pausendauer (SOA-Dauer). Hat der Prime-Reiz eine semantische Ähnlichkeit mit dem Target-Reiz, dann wird der Erkennungs- und Entscheidungsprozess beschleunigt. Tritt ein Effekt bereits bei einer sehr kurzen SOA-Pause auf, dann spricht dies für sehr geringe kognitive Aktivitäten. Ist die Pause länger als einige Sekunden, dann verschwindet der Einfluss des Prime-Reizes auf den Target-Reiz.

Dieser Priming-Effekt wird je nach Begriffsstruktur-Modell anders erklärt: Semantische Netzwerkmodelle erklären den Effekt durch eine automatische Voraktivierung (Spreading-Activation-Theory) der betreffenden semantischen Begriffsstruktur (Collins & Loftus 1975). Durch die Identifikation des ersten Wortes wird das betreffende se-

mentische Umfeld in eine erhöhte Aktivität versetzt, so dass der Target-Reiz dann – sofern er dem gleichen semantischen Umfeld entspringt – schneller erkannt wird. In Propositionalen Netzwerkmodellen wurde das Phänomen durch das Konzept assoziativer Vernetzung und der assoziativen Übergangswahrscheinlichkeit erklärt: Begriffe, die durch ein Reizwort gebahnt werden, werden durch Vernetzung und eine höhere Übergangswahrscheinlichkeit erklärt; Begriffe, die nicht gebahnt werden, mit einer geringeren Vernetzung und niedrigeren Übergangswahrscheinlichkeit.

Auch Klix, Hoffmann et al. verwendeten Priming-Experimente (insbesondere Priming-Analogieerkennungs-Experimente), um Hypothesen bezüglich Relationen zwischen Begriffen zu überprüfen. Die Ergebnisse einer Vielzahl von Experimenten bestätigen nicht nur die Hypothese, dass Begriffe durch Relationen in Beziehung stehen. Die Ergebnisse konnten darüber hinaus auch Hypothesen bestätigen, die unterschiedliche Relationsarten behaupten, wie hierarchisch-kategoriale Begriffsrelationen (Ober-, Unter-, Neben-, Kontrastbegriffsrelationen) und ereignisbezogene Begriffsrelationen (wie Handlungsträger-, Rezipient-, Objektrelation etc.) (vgl. Klix, Kukla & Klein 1976, zitiert nach Hoffmann 1986; Krause, Lohmann & Tschecke 1980 zitiert nach Hoffmann 1986; Klix 1984, Preuß 1985 und Schmieschek 1989 zitiert nach Klix 1992; Küchler 1989). Aufgrund der Ergebnisse lehnt Klix (1984, 1991) das Konzept einer gleichmäßigen Vernetzung wie bei konnektionistischen Netzwerken und somit eines stationär aktivierbaren Begriffsfeldes ab. Die unterschiedlichen Erkennungszeiten sprechen für einen sehr unterschiedlichen Vernetzungsgrad der Begriffe und sehr unterschiedliche Qualitäten der Relationen. Welche Qualitäten diese Relationen haben, wird in der 3. Gesetzmäßigkeit näher dargestellt. Hier soll es genügen, dass die Begriff untereinander in einer spezifischen semantischen Nähe bzw. Distanz stehen. Daraus lassen sich Konsequenzen für nutzerorientierte Anordnung entwickeln.

Annahmen bezüglich des Anordnungsproblems In den empirischen Untersuchungen sollen nutzerorientierte Anordnungen für Süßgebäck, Zimmerpflanzen, Flughafeneinrichtungen und Geschäftstypen entwickelt werden. Jede dieser Gegenstandsklassen besteht aus einer Vielzahl von Objekten, über die die NutzerInnen aufgrund ihrer Handlungserfahrungen Begriffe gebildet haben. Dabei wird von folgender *Annahme* ausgegangen: *Begriffe über alltägliche Objekte haben eine unterschiedliche, spezifische se-*

semantische Nähe bzw. Distanz (Ähnlichkeit). Die Beziehungen zwischen diesen Begriffen sind nicht gleich. Um es an einem Beispiel zu erläutern: Es wird angenommen, dass Begriffe wie Butterkeks, Löffelbiscuit, Schoko-Nussgebäck für die KäuferInnen von Süßgebäck nicht gleich ähnlich (nah) sind, sondern dass die Begriffe eine sehr unterschiedliche semantische Nähe haben. Butterkeks und Löffelbiscuit sind wahrscheinlich semantisch ähnlicher (näher) als Butterkekse und Schoko-Nussgebäck.

Auf der Basis der Gesetzmäßigkeit, dass Begriffe in definierter semantischer Nähe bzw. Distanz zueinander repräsentiert sind und der Tatsache, dass Anordnungen von Objekten durch eine räumliche (relationale oder topologische) Nähe bzw. Distanzen beschrieben werden können (vgl. Kapitel 3), wird folgende *Schlussfolgerung* gezogen:

- Wenn die relative räumliche Nähe bzw. Distanz der Objekte der semantischen Nähe bzw. Distanz der betreffenden Begriffe entspricht, dann wird diese Anordnung den spontanen Erwartungen der Nutzer entgegenkommen und somit nutzerfreundlicher sein als eine Anordnung nach anderen Kriterien, d.h. das Such- und Findeverhalten der Nutzer müsste besser, d.h. zeitlich und von der Wegstrecke kürzer sein. Um das Such- und Findeverhalten der Nutzer zu unterstützen, müsste also das, was für die Nutzer begrifflich nahe beieinander ist, dementsprechend auch räumlich nahe beieinander platziert werden.

Schlussfolgerungen für das Untersuchungsdesign Für die Lösung des Anordnungsproblems ergibt sich daraus folgendes Vorgehen: die Begriffsstruktur, d.h. die Beziehung zwischen Begriffen wird als semantische Ähnlichkeit operationalisiert. Als *Erstes* müsste – mittels einer geeigneten Methode – die begrifflich-semantische Nähe (=Ähnlichkeit) für die Begriffe der anzuordnenden Objekte an den entsprechenden NutzerInnen ermittelt werden; Als *Zweites* müsste dann die semantische Nähe der Begriffe in eine räumliche Anordnung der Objekte übersetzt werden: semantisch ähnliche Begriffe sind entsprechend räumlich nah, semantisch unähnliche Begriffe sind entsprechend räumlich entfernt anzuordnen. Dabei ist die räumliche Anordnung nicht als metrische, sondern nur als relationale bzw. topologische Anordnung zu verstehen. Gestalter können dann im konkreten Gestaltungsprozess diese relationale Anordnung in die Gestaltungslösung übersetzen.

Für diese Schlussfolgerung wäre, wie eingangs festgestellt wurde, ein Heranziehen der Theorie von Klix, Hoffmann et al. nicht notwendig gewesen, da diese Gesetzmäßigkeit ja auch von anderen Theorien formuliert wird. Doch anhand der Theorie von Klix, Hoffmann et al. lassen sich weitere Gesetzmäßigkeiten formulieren, die eine Prognose erlauben, welchen Kriterien die semantische Ähnlichkeit (Nähe) von Begriffen folgt.

2. Gesetzmäßigkeit: Begriffe und Begriffsstrukturen sind abhängig vom Handeln und repräsentieren früher gemachte Handlungserfahrungen

Begriffe und Begriffsstrukturen werden als untrennbar mit dem Verhalten verbunden definiert. Sie entstehen während des Verhaltens (Klix 1971, Hoffmann 1986). Und sie sind wichtig für eine effektive Verhaltenssteuerung. Klix et al. definieren Begriffe als Zusammenfassungen von invarianten Merkmalen. Diese Merkmale gewinnt die Person in einer Handlungsfolge mit einem Gegenstand. Diese Handlungsfolge folgt einem Orientierungsanlass. Klix (1971, 1992) definiert Orientierungsanlass im Sinne von Motiv, Interesse, Bedürfnis und schreibt ihm eine wesentliche, Ordnung bildende Funktion zu. Hoffmann schloss sich dieser Ansicht nicht explizit an (1983, 1986) und untersucht derzeit im Rahmen seiner Theorie der Antizipativen Verhaltenssteuerung die Bedeutung des Feedbacks und des Handlungserfolges (Stock & Hoffmann 2001).

Im Rahmen dieser Studie scheint es zunächst auch nicht wichtig zu sein, welche der beiden Positionen die „Richtige“ ist. Wesentlich ist, dass Begriffe und Begriffsstrukturen generalisiertes Wissen darstellen im Sinne von abstrahiertem, invariantem Wissen. Und, dass sie kein generelles Wissen darstellen, das alle Personen „automatisch“ haben. Sondern dass Begriffe und Begriffsstrukturen vielmehr davon abhängig sind, welche Erfahrungen eine Person bisher gemacht hat, welche Handlungsabfolgen sie in welchen Situationen gemacht hat und zu welchen Konsequenzen dies geführt hat. Dementsprechend haben Personen, die unterschiedliche Erfahrungen mit den gleichen Objekten hatten, unterschiedliche Begriffe und Begriffsstrukturen über diese Objekte. Interessant ist dabei, wie groß diese Unterschiede sind.

Empirische Belege Folgende beiden Laborexperimente belegen diese Annahme: Kukla & Sydow (1970, zitiert nach Klix 1971 S. 337) untersuchten *die Veränderbarkeit von*

Begriffsklassifizierungen in Abhängigkeit vom Lernziel. Zunächst mussten Versuchspersonen Objekte bezüglich ihrer Ähnlichkeit beurteilen. Dann wurden Lernversuche durchgeführt: Die Versuchspersonen mussten nun lernen, die Objekte nach einem Merkmal zu klassifizieren, das sich bei der ersten Ähnlichkeitsskalierung als unbedeutend zeigte. Die neue Klassenzuordnung war den Versuchspersonen also spontan nur wenig eingängig und sie mussten diese Klassifizierung lernen. Nach dem Lernversuch wurde erneut eine Ähnlichkeitsskalierung durchgeführt. Die erlernte Klassenbildung spiegelte sich nun in der Veränderung der Ähnlichkeitsbeurteilung wieder. Dieser Effekt deutet auf eine Veränderbarkeit der Klassifizierungsleistung. Er weist darauf hin, dass die Begriffsbildung vom Lernziel abhängig ist.

Zimmer (1984) konnte in einem *Experiment die unterschiedliche Verfügbarkeit der Begriffe von Experten im Vergleich zu Laien zeigen* und somit den Unterschied in der Merkmalsanalyse. Floristen und Laien wurde erst ein Wort vorgesprochen und 1 sec später ein Bild gezeigt. Sie mussten nun entscheiden, ob beides übereinstimmt. Die insgesamt 27 Bilder stellten Strichzeichnungen oder Farbfotos von Blumen sowie Kontroll-Items (Gemüse, Tiere etc.) dar. Die Worte unterschieden sich in ihrem begrifflichen Abstraktionsniveau: Einmal bezeichneten sie spezifisch die dargestellte Blume (Rose, Veilchen, Tulpe etc.) oder das Kontrollobjekt (Himbeere, Maikäfer etc.), ein anderes Mal waren es die Oberbegriffe (Blume, Beere etc.). Experten und Laien unterschieden sich nicht bei der Entscheidung, ob die Strichzeichnung mit dem Wort übereinstimmte. Strichzeichnungen sind schon auf begriffsrelevante Merkmale reduziert und erleichtern auch dem Laien die Merkmalsanalyse. Beide Gruppen entschieden schneller, wenn das Wort abstrakter war. Abstraktere Worte, wie „Blume“, erfordern einen geringeren Aufwand im Merkmalsvergleich mit der Darstellung einer Tulpe, als wenn die Personen das Wort Tulpe mit einer Tulpenstrichzeichnung vergleichen müssen. Die beiden Versuchspersonengruppen unterschieden sich jedoch bei der Beurteilung der Farbfotos. Die Experten beurteilten die Farbfotos mit spezifischen Bezeichnungen genauso schnell wie Farbfotos mit abstrakten Bezeichnungen, während die Laien bei spezifischen Bezeichnungen mit Farbfotos signifikant länger brauchten als bei abstrakten. Zimmer schloss aus den Ergebnissen, dass Experten durch ihre berufliche Tätigkeit über differenziertere und schneller verfügbare Begriffe verfügen.

Auch Quasi-Experimente bestätigen eine Abhängigkeit der Begriffsstruktur von den Handlungsvorerfahrungen: Nancy Henley (1969), Christine Storm (1980) und Andreas Hejj und Gerhard Strube (1988) untersuchten z.B. das Begriffsfeld Säugetiere bei Kindern und Zoologen. Storm untersuchte u.a. mittels Sortieren und Paarvergleich 12 Tiernamen an Versuchspersonen an der kanadischen Ostküste. Für die Versuchspersonen der 7. und 11. Klasse sowie für die Studenten waren sensorische, insbesondere äußerlich sichtbare Merkmale (z.B. Größe) klassifizierungsrelevant. Zoologen strukturierten dagegen deutlich nach Ernährung und Reproduktionsrate. Hejj und Strube (1988) replizierten Storms Studie in München und konnten diese Wissens- und Tätigkeitsabhängigkeit bestätigen. Sie kamen zu der Schlussfolgerung, dass Experten nicht nur mehr Faktenwissen haben, sondern grundlegend tiefere Modelle über die Objekte. Dies zeigt sich in konsistenteren Beurteilungen und einer geringeren Streuung.

Entsprechend dieser Erkenntnisse müssten z.B. Vielflieger die semantische Ähnlichkeit von Flughafen-Serviceeinrichtungen anders beurteilen als Wenigflieger, weil sie andere Erfahrungen, andere Motive, Interessen, Handlungsabfolgen mit diesen Einrichtungen haben. Vielflieger fliegen beruflich bedingt innerdeutsche Ziele an, während die Wenigflieger in der Regel als Touristen vor allem ausländische Urlaubsorte anfliegen. Deshalb müssten z.B. die Vielflieger die Schalter der innerhalb Deutschlands fliegenden Fluggesellschaften und der Charterfluggesellschaften anders beurteilen als die Wenigflieger.

Hypothese bezüglich des Anordnungsproblems: Für das Anordnungsproblem lässt sich folgende Annahme ableiten:

- Wenn Begriffe und Begriffsstrukturen abhängig sind von früheren Handlungserfahrungen, dann haben Personen, die ähnliche Handlungserfahrungen mit Objekten gemacht haben, auch ähnliche Begriffsstrukturen über diese Objekte. Personen, die unterschiedliche Handlungserfahrungen mit den Objekten hatten, haben unterschiedliche Begriffsstrukturen über die gleichen Objekte.

Schlussfolgerungen für das Untersuchungsdesign Für die Entwicklung einer nutzerorientierten Anordnung bedeutet dies: Die Begriffsstruktur von Objektbegriffen muss an der entsprechenden Nutzergruppe erhoben werden, d.h. an jenen Personen, die sich in

der Anordnung später auch intuitiv zurechtfinden sollen und die mit den Objekten bisher ähnliche Handlungserfahrungen gemacht haben. Ein Beispiel soll dies verdeutlichen: Wenn es darum geht, eine nutzerfreundliche Anordnung z.B. für Zimmerpflanzen in einem Verkaufsraum zu erarbeiten, müssen echte KundInnen untersucht werden und nicht z.B. FachverkäuferInnen oder andere Personen. Es ist anzunehmen, dass KundInnen von Zimmerpflanzen andere Begriffe und Begriffsstrukturen über Zimmerpflanzen haben als z.B. FachverkäuferInnen, Gärtner oder Ladenplaner, die ihr eigenes Fach- und Erfahrungswissen haben. Um zu prüfen, ob diese Prognose tatsächlich zutrifft, werden z.B. bei der Zimmerpflanzen-Studie KundInnen und Fachkräfte untersucht und verglichen. Entsprechend der oben genannten Gesetzmäßigkeit müssten sich Unterschiede in der Ähnlichkeitsbeurteilung von Zimmerpflanzen zeigen.

3. Gesetzmäßigkeit: Ereignisbezogene Begriffsbeziehungen sind schneller verfügbar als hierarchische Begriffsbeziehungen

Klix et al. nehmen zwei Arten von Begriffsbeziehungen an: Über die *hierarchischen* Beziehungsmerkmale steht ein Begriff in Beziehungen zu Ober-, Unter-, Neben-Begriffen etc.. Über *ereignisbezogene* Merkmalsrelationen steht ein Begriff in Beziehungen zu „ereignis-typischen“ Begriffen, wie Handlungsträger-, Objekt-, Finalitäts-, Ortsbegriffen etc. Zentraler Begriff einer ereignisbezogenen Begriffsstruktur ist ein Tätigkeits- bzw. Handlungsbegriff. Klix et al. nehmen an, dass *ereignisbezogene Begriffsbeziehungen schneller verfügbar sind als hierarchische Begriffsbeziehungen*. Sie nehmen an, dass ereignisbezogene Begriffsstrukturen explizit gespeichert sind, während hierarchische Beziehungen erst bei Bedarf hergestellt werden. Ein Herleiten solcher Ober-Unter-Begriffsbeziehung erfordert damit mehr Zeit.

Klix widerspricht hiermit den Annahmen anderer Modelle. Bei den Semantischen Netzwerkmodellen wird die Ober-Unter-Begriffsbeziehung wie auch ihre Umkehrung (Ober-Unter-Begriffsbeziehung) als explizit gespeichert angenommen (Collins & Quillian 1972). Klix führt zum einen folgendes Plausibilitätsargument an: „Nehmen wir an, in einem erwachsenen Gedächtnis seien nur 10.000 Begriffe mit Namen und Merkmalsätzen gespeichert. Nehmen wir ferner an, die Ähnlichkeiten zwischen ihnen seien symmetrisch, was in der Regel nicht zutrifft. ... Also, die Ähnlichkeitswerte seien von

einem Begriff A zu B die gleichen wie von B zu A. Dann gibt das bei expliziter Speicherung einen Speicherbedarf für annäherungsweise $1/2 \cdot (10.000!)$ Daten. Dieser Wert ist größer als die Anzahl der Atome im Weltall geschätzt wird.“ (Klix 1992, S. 305). Er nennt darüber hinaus Experimente, die die Annahme unterstützen, dass ereignisbezogene Relationen explizit gespeichert sind und assoziativ aktiviert werden (z.B. van der Meer 1978, zitiert nach Klix 1984, Preuß 1991 und Kuchler 1991).

Empirische Belege Kuchler (1991) führte ein Priming-Experiment durch, um diese Hypothesen zu prüfen. Für das Priming-Experiment wählte sie drei Arten von Wortpaaren:

Ober-Unterbegriffs-Wortpaare: z.B. Pudel–Hund, Cobra–Schlange, Blume–Pflanze

Ereignisbezogene Wortpaare: z.B. Tisch–Decken, Lied–Singen, Pfeife–Rauchen

Composita-Wortpaare: z.B. Fenster–Rahmen, Tee–Kanne, Butter–Brot

Der Prime-Reiz war entweder eines der ersten Worte oder nur ein Signal. Dann folgte eine Pause (SOA) von 200ms oder 1000ms. Dann folgten der Target-Reiz, der das zweite Wort des Begriffspaares war, entweder richtig geschrieben oder falsch („songen“ statt singen, „Pflonze“ statt Pflanze etc.). Die Versuchspersonen sollten nach der Darbietung des Target-Reizes so schnell wie möglich entscheiden, ob der Target-Reiz richtig geschrieben war oder falsch. Die Ergebnisse zeigten, dass bei 200ms SOA-Pause der mit einem ereignisbezogenen Begriff geprimte Target-Reiz signifikant schneller bestätigt wurde als der mit einem hierarchischen Begriff geprimte Target-Reiz. Die Identifizierung der richtigen oder falschen Schreibweise des Target-Reizes konnte nur bei den ereignisbezogenen Prime-Reizen verbessert werden. Die Leistungen bei einem bloßen Signal waren signifikant schlechter. Dahingegen gab es nahezu keinen Unterschied zwischen dem Signal und einem hierarchischen Prime-Reiz. Bei beiden Darbietungen brauchten die Versuchsperson gleich lang, um die richtige bzw. falsche Schreibweise zu identifizieren, die wie gesagt länger war als bei den ereignisbezogenen Wortpaaren. Bei einer SOA-Pause von 1000ms zeigten sich diese Effekte nicht mehr. Die längere Zeitdauer reichte aus, um auch hierarchische Wortpaare vorab zu aktivieren. Die Ergebnisse stützen die Annahme, dass ereignisbezogene und hierarchische Begriffsbeziehungen unterschiedlich repräsentiert sind, und dass ereignisbezogene Begriffsbeziehungen schneller verfügbar sind als hierarchische.

Hypothesen bezüglich des Anordnungsproblems Da angenommen wird, dass die Erwartung, wo ein Objekt zu finden ist, eine spontane, „automatische“ Schlussfolgerung ist, wird daher weitergehend angenommen, dass ereignisbezogene Begriffsstrukturen die Basis für die Erwartung sind. Für das Anordnungsproblem lassen sich folgende zwei Hypothesen formulieren:

- Erhebt man an Nutzern die begrifflich-semantische Nähe von Objekten, dann werden sie diejenigen Objekte als semantisch ähnlicher beurteilen, die sie bisher in einem Ereignis- bzw. Handlungskontext verwendet haben. Sie werden die Objekte nach handlungsbezogenen und nicht nach hierarchisch-kategorialen Kriterien beurteilen.
- Die Begriffsstruktur, d.h. die Beziehung semantischer Ähnlichkeitsbeurteilung ist stabil und nicht durch ein aktuelles Handlungsziel veränderbar. Die Beurteilung der semantischen Nähe von Begriffen ist unabhängig davon, welches „aktuelle“ Ziel sie verfolgt.

Dabei ist Folgendes zu berücksichtigen: Bei den anzuordnenden Objekten oder Funktionen (Einzelhandelsgeschäfte, Zimmerpflanzen, Süßgebäckartikel oder Flughafen-Einrichtungen) handelt es sich um Objekte, die als Kontext- und als Nebenbegriffe betrachtet werden können. Es wird aber dennoch angenommen, dass sie nach handlungsbezogenen Kriterien beurteilt werden, also danach, in welchem Handlungskontext sie verwendet werden.

Schlussfolgerungen für das Untersuchungsdesign Um die zweite Hypothese zu prüfen, müsste in einem Ähnlichkeitsbeurteilungs-Experiment die Variable „aktuelle Zielorientierung“ variiert werden. Die Variation dieser Variable könnte dadurch erfolgen, dass zwei Versuchspersonengruppen über entsprechende Instruktionen in unterschiedliche Zielorientierungen versetzt werden.

5.2 Zusammenfassung: Problemlösungsweg und Hypothesen

Die Anwendung der Theorie von Klix et al. auf das Anordnungsproblem führt zu einem *Problemlösungsweg* und zu *Prognosen* im Sinne von untersuchungsleitenden Hypothesen, die in einer Input- und Evaluations-Studie überprüft werden. Der *Problemlösungs-*

weg zur Entwicklung einer nutzerorientierten Objktanordnung im Raum lautet wie folgt:

1. In einem *ersten Schritt* ist für die anzuordnenden Objekte die Begriffsstruktur an der entsprechenden Nutzerstichprobe im Sinne einer *Input-Studie* zu erheben.
 - 1.1 Zunächst sind die Gegenstände/Objekte/Funktionen festzulegen, für die eine nutzerorientierte Anordnung entwickelt werden soll (z.B. Artikel in einem Supermarkt, Flughafeneinrichtungen, Funktionen auf einer Bedienoberfläche).
 - 1.2 Außerdem ist die Nutzergruppe zu bestimmen, für die diese nutzerorientierte Anordnung entwickelt werden soll. Die Nutzergruppe definiert sich durch ähnliche Handlungen (Handlungsziele, -abfolgen, -erfahrungen) mit den Objekten.
 - 1.3 Dann ist an dieser Nutzerstichprobe die Begriffsstruktur für die Objekte zu erheben. Die Begriffsstruktur wird dabei definiert als begrifflich-semantische Ähnlichkeit, d.h. als die semantische Nähe bzw. Distanz zwischen den Begriffen. Es ist daher eine Methode zu finden, die sich eignet, die begrifflich-semantische Ähnlichkeit von Objekten an NutzerInnen zu erheben.
 - 1.4 Es werden *folgende Prognosen im Sinne von untersuchungsleitenden Hypothesen* bezüglich der begrifflichen Ähnlichkeitsbeurteilung der Objekte aufgestellt:
 - a) Die begrifflich-semantische Ähnlichkeit von Objekten ist abhängig von den Nutzungsvorerfahrungen. Für Personen einer Nutzergruppe lassen sich ähnliche Ähnlichkeitsbeurteilungen für die entsprechenden Objektbegriffe finden.
 - b) Die begrifflich-semantische Ähnlichkeit folgt handlungsbezogenen und nicht kategorialen Kriterien. D.h. es werden solche Begriffe semantisch als ähnlich beurteilt, die in einem gemeinsamen Handlungskontext verwendet wurden und nicht solche, die in einer hierarchisch-kategorialen Begriffsbeziehung stehen.

- c) Die begrifflich-semantische Ähnlichkeitsbeurteilung ist nicht vom aktuellen Handlungsziel abhängig.
2. In einem *zweiten Schritt* ist die an der Nutzergruppe erhobene begrifflich-semantische Ähnlichkeitsbeurteilung in eine räumliche Anordnung der Objekte zu übersetzen und in einer *Evaluations-Studie* zu überprüfen.
 - 2.1 Die empirisch gefunden begrifflichen Ähnlichkeitsrelationen sind auf die Objektmenge zu übertragen. Diese begriffs-strukturierte Anordnung der Objekte ist den Gestaltern (z.B. Architekten, Regalbelegern etc.) zu vermitteln.
 - 2.2 Die Gestalter übersetzen die Anordnungsvorgabe in eine konkrete räumliche Anordnung, die dann noch z.B. auf anthropometrische Kriterien (körperliche Erreichbarkeit) geprüft werden muss.
 - 2.3 Es wird folgende *Prognose im Sinne einer untersuchungsleitenden Annahme* bezüglich einer begriffsstruktur-orientierten Anordnung der Objekte gestellt:

Nutzer werden in Objektanordnungen, in der die räumliche Nähe bzw. Distanz zwischen den Objekten der eigenen begrifflich-semantischen Nähe bzw. Distanz der Objektbegriffe entspricht ein besseres Such- und Findeverhalten zeigen, als in Anordnungen, die anderen Kriterien folgen.

Die Ausführungen zeigen, dass die Theorie es ermöglicht, einen Problemlösungsweg zu beschreiben und Prognosen zu liefern. Im nächsten Kapitel werden Problemlösungsweg und Prognosen in Hypothesen und in ein passendes Experimentaldesign übersetzt.

6. Entwicklung des Untersuchungsdesigns

Im folgenden Kapitel wird das Untersuchungsdesign zur Erhebung der semantischen Ähnlichkeit von alltäglichen Objekten an Nutzern geklärt. Im *ersten* Abschnitt werden Untersuchungsvariablen und -stichproben operationalisiert. Im *zweiten* Abschnitt werden mögliche Erhebungsmethoden dargestellt und es werden die geeigneten ausgewählt. Im *dritten* Abschnitt werden passende Auswertungsmethoden dargestellt. Im *vierten* Abschnitt werden Methoden für den Vorversuch, also für die Auswahl der zu untersuchenden Objekte dargestellt und im *fünften* Abschnitt erfolgt eine Zusammenfassung.

6.1 Untersuchungsvariablen und Stichprobenanforderungen

Untersuchungsziel der empirischen Studien ist es, an der betreffenden Nutzergruppe, d.h. an den Personen, die dann auch die Objekte suchen und finden müssen, die Begriffsstruktur über die Objekte zu erheben. Die in Kapitel 5 entwickelten Prognosen werden für die Untersuchung wie Hypothesen verwendet.

Operationalisierung der Untersuchungsvariablen Es werden zwei *abhängige Variablen* untersucht: a) die *Begriffsstruktur* Sie wird entsprechend den theoretischen Ausführungen in Kapitel 5 als begrifflich-semantische Ähnlichkeit operationalisiert. Sie wird über Ähnlichkeitsbeurteilungsverfahren erhoben. b) die *Klassifizierungskriterien*: Die Ähnlichkeitsbeurteilungen werden auf handlungsbezogene versus kategoriale Kriterien untersucht. Handlungsbezogene Kriterien sind solche, die sich auf die Handlung, in der die Objekte verwendet werden, beziehen (z.B. Handlungsziel, Handlungskontext). Kategoriale Kriterien sind solche, die die hierarchischen, kategorialen (Ober-, Unter- oder Nebenbegriffs-)Beziehungen der Objekte wiedergeben (z.B. fachliche Systematiken).

Es werden zwei *unabhängige Variablen* untersucht: a) die *Handlungsvorerfahrung*: Die Handlungsvorerfahrung wird operationalisiert über die NutzerInnen bzw. KäuferInnen, die die zu untersuchenden Gegenstände nutzen. Es wird angenommen, dass aufs Geratewohl bzw. zufällig gezogene, typische NutzerInnen bzw. KäuferInnen ähnliche

Handlungsziele mit den Objekten haben, und damit auch sehr ähnliche Handlungsvorerfahrungen haben. Werden zwei unterschiedliche Nutzergruppen untersucht, werden unabhängige Stichproben gezogen, wobei jede Stichprobe unterschiedliche Aufgaben bezüglich der identischen Untersuchungsobjekte hat; b) die *aktuelle Zielorientierung*: Die aktuelle Zielorientierung wird operationalisiert über die Instruktion. Es werden zwei gleiche Stichproben gezogen, wobei jede Gruppe über die Instruktion ein unterschiedliches Ziel für die Ähnlichkeitsbeurteilung der Objekte erhält.

Stichprobenanforderungen Die Stichprobe muss aus denjenigen Personen gezogen werden, die die Anordnung typischerweise nutzen und die typische Vorerfahrungen mit den zu untersuchenden Objekten haben. Die Versuchspersonen werden an typischen Nutzungsorten aus der Menge der dort befindlichen NutzerInnen ausgewählt. Ein Beispiel soll dies verdeutlichen: zur Entwicklung einer nutzerorientierten Artikelanordnung im Supermarkt werden Supermarktkundinnen für die Untersuchungsstichprobe angesprochen und für die Untersuchung ausgewählt. Damit ist sichergestellt, dass es sich um Personen handelt, die die gleiche Aufgabe bezüglich der Objekte haben. Die Auswahl wird zufällig oder aufs Geratewohl getroffen. Eine zufällige Auswahl ist aus Gründen der internen Validität vorzuziehen. Aus forschungsökonomischen Gründen ist eine Auswahl aufs Geratewohl auch vertretbar.

In der kognitiven Psychologie und Wissenspsychologie wurde das Paradigma der *Expertise* entwickelt (vgl. z.B. Larkin 1979, Chi & Glaser 1980, Anderson 1993), dessen wichtigster methodischer Ansatz der „Experten-Novizen-Vergleich“ ist. Dieses Paradigma, das durch die Entwicklung von computergestützten Expertensystemen entstand, hat zu sehr fruchtbaren Auseinandersetzungen mit Wissensinhalten und ihrer Bedeutung für die Wahrnehmung, die Problemlösung sowie der allgemeinen kognitiven Entwicklung geführt (vgl. Opwis 1992). Bereiche der Anwendungsforschung, wie der Fachbereich Software-Ergonomie oder Nutzerforschung, haben den Begriff der Expertise und des Experten-Novizen-Vergleichs übernommen, leider nicht immer richtig und daher nicht immer zielführend.

Ein Beispiel aus dem Handel soll dies verdeutlichen: FachverkäuferInnen mit ExpertInnen und die KundInnen mit NovizInnen gleichzusetzen, wäre falsch und nicht

zielführend. Oder um es salopp zu sagen: Es wäre, als würde man Äpfel mit Birnen vergleichen! Die KundInnen sind in dem Aufgabenbereich „Einkaufen“ noch lange keine AnfängerInnen und sie sollen keine ExpertInnen im Sinne von FachverkäuferInnen werden! Die Handlungsziele, Handlungskontexte und Handlungen beider Nutzergruppen sind grundlegend andere! Es wäre allerdings notwendig, die KundInnen nach Einkaufs-ExpertInnen und Einkaufs-NovizInnen zu unterscheiden, wenn entsprechend eine ExpertInnen-Anordnung und NovizInnen-Anordnung entwickelt werden soll. Dies ist aber in den folgenden Studien nicht der Fall.

Die Bezeichnung „Novize“ verleitet meiner Erfahrung nach zu einer Abqualifizierung von eigentlich wichtigen Nutzergruppen, im Sinne „Die sind einfacher gestrickt“! „Sie sind ja keine Profis!“ und damit einhergehend werden auch ihre Bedürfnisse weniger ernst genommen und wertgeschätzt. Meiner Erfahrung nach ist es deshalb zielführender zu sagen: jede Nutzergruppe hat ihr spezifisches Vorwissen, ihre spezifischen Handlungsziele, Handlungsanforderungen, Kontextbedingungen, Motive etc. und benötigt somit auch ihre spezifischen Lösungen. Und: sofern man für sie spezifische Lösungen anbieten will, sollte man sie auch untersuchen.

6.2 Bestimmung der Erhebungsmethoden

Zunächst werden die Anforderungen formuliert, denen die Erhebungs- und Auswertungsmethoden genügen müssen. Dann folgt eine Übersicht über gängige Methodenklassen der Begriffsforschung, anschließend werden diese Methodenklassen einzeln dargestellt und hinsichtlich einer Eignung bzw. Nichteignung begründet. Zuerst werden die Methodenklassen, die sich nicht eignen, dargestellt, zuletzt die geeignete.

6.2.1 Anforderungen an die Methoden

Erhebungs- und Auswertungsmethoden sind nicht bloße Werkzeuge – eine Tatsache, die nach meiner Kenntnis leider in der angewandten Psychologie und Nutzerforschung

häufig nicht beachtet wird. In der Nutzerforschung dominieren Fragen, wie nach der Reliabilität, Validität, Eichung von Methoden (z.B. von Fragebögen, Beobachtungsbögen etc.) und der Repräsentativität der Stichproben (vgl. Kanis 2000). Die Qualität der Untersuchungsergebnisse wird dann entsprechend der Kennwerte der Methoden und der Stichprobencharakteristika beurteilt. Die Ergebnisse werden augenscheinlich interpretiert und nicht in Bezug auf Hypothesen und die zugrunde liegende Theorie. Die Methoden müssen jedoch zur Theorie, zum Untersuchungsgegenstand, den Versuchspersonen passend ausgewählt werden. Im Folgenden werden die Anforderungen an die Methoden spezifiziert.

Anforderungen von Seiten des Untersuchungsgegenstandes Die Methoden müssen geeignet sein, semantische Ähnlichkeit von Begriffen über unterschiedliche Gegenstandsbereiche, wie Lebensmittelartikel, Zimmerpflanzen, Geschäftstypen (Bäcker, Metzger, Boutique etc.) und Flughafenserviceeinrichtungen (vom Check-In über Reisebüro bis zum Meeting-Point), zu erheben. Diese Gegenstandsbereiche unterscheiden sich a) phänomenologisch (z.B. in Größe, Reizvielfalt und -komplexität, Anfassbarkeit etc.); b) im Vorhandensein und in der Verwendung von unterschiedlichen Wortmarken; c) in der assoziativen Verfügbarkeit des Begriffes: Es wird vermutet, dass den Versuchspersonen z.B. weitaus mehr Objekte bekannt sind, wenn man sie ihnen sprachlich oder bildhaftanschaulich vorlegt, als wenn man sie frei assoziieren lässt. (So wird die Meringe oder die Monstera wohl assoziativ schwer verfügbar sein, aber anschaulich vorgelegt sofort erkannt werden. Die Meringe ist ein weißes Zuckergebäck, die Monstera ist eine Zimmerpflanze mit großen, dunkelgrünen Blättern.)

Für die Auswahl von Methoden ist daher zu berücksichtigen: a) Die Erhebungsmethode sollte sowohl mit realen Reizvorlagen, mit Abbildungen als auch mit Wortmarken arbeiten können. Wenn Wortmarken verwendet werden, muss sichergestellt werden können, dass die Versuchspersonen diese eindeutig verstehen. b) Reizgebundene (recognition) Verfahren sind assoziativen (recall) Verfahren vorzuziehen, weil dadurch auch zu Objekten ohne Wortmarke und mit geringer assoziativer Verfügbarkeit Reaktionen erhalten werden.

Anforderungen von Seiten der Begriffstheorie Die Methode sollte zu den Basisannahmen der Theorie über Begriffsbildung von Klix et al. passen, wie a) der Merkmalsabstraktion; b) es erlauben, die begrifflich semantische Nähe bzw. Ähnlichkeit von Begriffen zu erheben und zwar sowohl die ereignisbezogenen als auch die hierarchischen Begriffsbeziehungen; c) wenn möglich, kategoriale Mehrfachzuordnungen (Kreuzklassifikationen) zulassen; d) es erlauben mit realen Objekten, anschauliche Abbildungen oder den Versuchspersonen bekannte Wortmarken zu arbeiten; e) sofern sie zeitgebunden bzw. assoziativ ist, Eigenschaften von prototypischen Begriffen berücksichtigen.

Anforderungen von Seiten der Versuchspersonen Die Methoden müssen es erlauben, typische NutzerInnen für die Untersuchung zu rekrutieren. Um NutzerInnen zur Teilnahme zu motivieren, müssen die Methoden nutzernah sein, da die meisten Nutzergruppen (z.B. einkaufende Hausfrauen, viel fliegende Manager etc.) sich in engen Zeit- und Aufgabenplänen befinden. Methoden, die Laboruntersuchungen erfordern, beeinträchtigen die Bereitschaft zur Teilnahme und somit die Repräsentativität der Stichprobe. Die Erhebungsmethode sollte daher a) für einen Einsatz vor Ort geeignet sein; und b) die Versuchsperson annehmbar belasten hinsichtlich zeitlicher Dauer, psychischer Sättigung, kognitive Belastung etc.

6.2.2 Übersicht über Methoden der Begriffsforschung

Tabelle 6.1 gibt eine Übersicht über die in der Begriffsforschung verwendeten Methodenklassen (vgl. u.a. Velickovskij 1982, Kluwe 1995, Anderson 1996, Bredenkamp und Erdfelder 1996). Die Methodenklassen unterscheiden sich hinsichtlich des Untersuchungsziels, der erhobenen Daten und der theoretischen Herkunft, wobei zu bedenken ist, dass jede Methodenklasse zwar meist innerhalb eines speziellen Modells entwickelt und paradigmatisch für dessen Hypothesenprüfung wurde. Doch sie wurde meist später in anderen, auch für die Prüfung konkurrierender Theorien verwendet. Auffallend ist, dass bei den meisten Methoden nur die inhaltlichen Reaktionen und die Reaktionszeiten der Versuchspersonen ausgewertet werden. Angesichts der Komplexität und Leistungskraft unseres begrifflichen Wissens erscheint dies etwas dürftig.

Tabelle 6.1. Erhebungsmethoden der Begriffsforschung

Methoden- klasse	Untersuchungsziel	Erhobene Daten	Theoretische Herkunft
Assoziations- aufgaben	Erheben semantisch naher Begriffe, deskriptive Wortfeld- und Wissensbereichsanalysen, Hypothesenprüfung zu Repräsentationsmodellen	Anzahl und Inhalt der Nennungen	Assoziations- theorien
Erzählaufgaben	Ermitteln von ereignisbezogenen Begriffsstrukturen wie Schemata, Skripts, Frames	Propositionen: Inhalt und Reihenfolge	Schemata-Theorien
Satzverifikations- aufgaben	Überprüfen von Hypothesen zu Repräsentationsmodellen	Inhalt Reaktionszeit	Semantische Netzwerkmodelle
Priming- aufgaben	Überprüfen von Hypothesen durch Voraktivierung semantisch naher Begriffe	Inhalt Reaktionszeit	Assoziationstheorien Semantische Netzwerkmodelle
Begriffsfindungs- aufgaben	Klären der Oberbegriffe, Überprüfen von Hypothesen zu Repräsentationsmodellen und begrifflichem Identifizieren	Inhalt Reaktionszeit	Merkmals- repräsentations- modelle
Relations- bestimmungs- aufgaben	Bestimmen der Relationstypen zwischen Begriffen, Ereignisbegriffen und Ereignisfolgen	Inhalt Reaktionszeit	Merkmals- repräsentations- modelle
Ähnlichkeits- urteile	Bestimmen der Nähe bzw. Distanz, der Gruppierung und der Klassifizierungskriterien von Begriffen	Ähnlichkeits- urteile Nachbarschaft Gruppierungen	Merkmals- repräsentations- modelle

6.2.3 Nicht geeignete Methoden

Assoziationsaufgaben

Assoziationsaufgaben sind Methoden zur Erhebung der semantischen Nähe von Begriffen, zur deskriptiven Untersuchung von Wortfeldern und Wissensbereichen, aber auch, um Hypothesen z.B. über die innere Organisation des Gedächtnisses zu prüfen (vgl.

Strube 1984) oder begriffliche Repräsentationsmodelle (z.B. Schemata). Assoziationsaufgaben wurden im Rahmen des Assoziationismus entwickelt und gelten als dessen experimentelles Paradigma. Der Assoziationismus ist in der westlichen Literatur über die Empiristen (Locke, Hume) bis auf Aristoteles zurückverfolgbar. Er liegt auch neueren „neoassoziationistischen“ Begriffsrepräsentationsmodellen (vgl. Semantische Netzwerkmodelle) sowie Konnektionistischen Modellen (vgl. Anderson 1996) zugrunde. (Hier allerdings werden grundlegende Annahmen des „klassischen“ Assoziationismus verworfen, z.B. dass jeder Begriff in elementare Empfindungen/ Wahrnehmungen zerlegbar ist, was anzeigt, dass die klassische Position nicht beizubehalten ist.)

Beschreibung von Assoziationsaufgaben Der Versuchsperson wird eine Serie von Reizworten dargeboten und sie soll die ihr dazu einfallenden Begriffe nennen. Es wird unterschieden zwischen einer freien Assoziation, bei der die Versuchsperson keine Vorgaben für mögliche Nennungen bekommt; einer gebundenen (restringierten) Assoziation, bei der die Versuchsperson Vorgaben für die Art der Assoziationen erhält und einer Nachbarschaftsuntersuchung: bei der die Versuchsperson auf ein Reizwort nur ein Wort assoziieren soll. Gemessen wird der Inhalt der Assoziation(en), eventuell zusätzlich die Reaktionszeit (vgl. Strube 1984, Marx 1988). Begriffe, die zeitlich früh auf das Reizwort genannt werden, werden durch eine höhere Übergangswahrscheinlichkeit erklärt.

Bewertung hinsichtlich einer Eignung als Erhebungsmethode Assoziationsaufgaben werden als wenig geeignet bewertet, um – entsprechend den Hypothesen – semantische Nähe-Distanz-Beziehungen für eine gegebene Begriffsmenge personenübergreifend abzutesten. Assoziative Nennungen sind belastet von Abrufeffekten: die Versuchspersonen nennen bei freien Assoziationen vorzugsweise zunächst Begriffe, die nahe um den Schlüsselreiz liegen, um nach einer Pause Begriffe aus situationalen Kontexten (vergleichbar mit Schemata bzw. Ereignisbegriffen) zu assoziieren (vgl. Strube 1984, Walker & Kintsch 1985 zitiert nach Klix 1992). Zwischen den ersten spontanen Nennungen und den situationalen Kontexten folgen diverse Phasen, bei denen nur Unterbegriffe, Oberbegriffe, Gegensatzbegriffe (Antonyme) oder nur Reimworte assoziiert werden, sodass zwischen „automatischem“ und „strategischem“ Abruf unterschieden werden muss (Clark 1972; zitiert nach Klix 1992). Sie liefern außerdem in methodenvergleichenden Untersuchungen undifferenziertere Ergebnisse als z.B. Sortieraufgaben (vgl.

Henley 1969, Storm 1980, Hejj & Strube 1988). Es ist auch davon auszugehen, dass die zu untersuchenden Begriffe nicht alle die gleiche assoziative Verfügbarkeit der Wortmarke haben. Daher sind reizabhängige Methoden wie Sortierverfahren assoziativen Verfahren vorzuziehen. Untersuchungstechnisch erfordern Assoziationsaufgaben nur dann Laborbedingungen, wenn auch Reaktionszeiten erhoben werden. Ansonsten können sie vor Ort eingesetzt werden. Problematisch ist die Auswertung. Für personenübergreifende Auswertungen müssen die Nennungen inhaltsanalytisch ausgewertet werden, da die Versuchspersonen meist unterschiedliche Worte für gleiche Begriffe verwenden. Dadurch wird die Auswertung aufwendig und subjektiv belastet. Inhaltlich eignet sich die Methode besser für die deskriptive Analyse individueller Assoziationsstrukturen als für den personenübergreifenden Vergleich von Ähnlichkeitsbeurteilungen, wie er hier angestrebt wird.

Erzählaufgaben

Erzählaufgaben werden z.B. zur inhaltlichen Analyse von ereignisbezogenen Begriffsstrukturen wie Schemata (Skripts und Frames) verwendet. Anderson oder Bower verwendeten Erzählaufgaben mit einer Propositionalen Inhaltsanalyse (bei der die Aussagen in Propositionen zerlegt und interindividuell verglichen wurden), um nationale Normen für Skripts und Frames zu entwickeln (vgl. Anderson 1996).

Beschreibung der Erzählaufgabe Die Versuchsperson wird aufgefordert, eine Situation bzw. ein Ereignis (z.B. einen Restaurantbesuch) zu beschreiben. Die Beschreibung wird inhaltlich, d.h. qualitativ ausgewertet.

Bewertung hinsichtlich einer Eignung als Erhebungsmethode Diese Methodenklasse ist als Erhebungsmethode nicht geeignet, weil sie ein zu eingeschränktes Untersuchungsziel hat. Erzählaufgaben können zwar vor Ort durchgeführt werden. In der Auswertung sind sie aber aufwendig und subjektiv belastet. Sie bevorzugen tendenziell ereignisbezogene Begriffsnennungen, was ein Überprüfen der Hypothese „ereignisbezogener versus hierarchisch-taxonomischer“ Begriffsstrukturen erschweren würde. Sie sind ähnlich problematisch einzustufen wie assoziative Verfahren.

Satzverifikationsaufgaben

Satzverifikationsaufgaben werden zur Überprüfung von Hypothesen bezüglich der Repräsentation von Begriffen und Begriffsstrukturen verwendet. Collins und Quillians (1972) verwendeten diese Methode beim Modell der semantischen Netzwerke, um die Speicher-Ökonomie-Hypothese zu prüfen. Kintsch (1974) verwendete sie, um sein Modell der propositionalen Repräsentation zu testen. Smith, Shoben & Rips (1974) untersuchten die Zusammenhänge von Ähnlichkeitsurteilen und Reaktionszeiten in Satzverifikationsaufgaben und verwendeten die Ergebnisse, um ihr Merkmalsrepräsentationsmodell für Begriffe (charakterisierende und definatorische Merkmale) zu prüfen. Rosch (1975) verwendete sie zur Prüfung der Prototypenrepräsentation.

Beschreibung der Satzverifikationsmethode Der Versuchsperson (Vp) werden Aussagen wie „Der Kanarienvogel ist gelb.“ zeitkontrolliert per Projektor vorgelegt, die sie per Knopfdruck mit „ja“ oder „nein“ auf Richtigkeit beurteilen soll. Dabei werden auch falsche Aussagen (Distraktoren) vorgelegt. Gemessen wird die Reaktionszeit und die Richtigkeit der Aussage. Aus der Dauer der Darbietungszeit und Reaktionszeit werden Rückschlüsse gezogen auf die Art der Repräsentationen.

Beschreibung der Faktenabrufmethode (als Variante von Satzverifikationsaufgaben) Die Vp soll eine Liste von Sätzen (einfache Fakten) lernen. In der weiteren Phase werden weitere Fakten dargeboten und es wird gefragt, ob diese in den ursprünglichen Sätzen enthalten waren. Gemessen werden Reaktionszeit und richtige Antworten.

Bewertung hinsichtlich einer Eignung als Erhebungsmethode Satzverifikationsaufgaben eignen sich für diese Arbeit nicht. Sie erfordern Laborbedingungen, um die Darbietungszeiten zu kontrollieren und die Antwortzeiten zu erheben und sind daher nicht nutzernah einsetzbar. Inhaltlich stellt sich das Problem durch die Vorgabe des Reizmaterials: Die Sätze müssen formuliert werden. Es stellen sich Probleme bei der Parallelisierung der Satz- und Wortlängen oder der Erzeugung einer gleichbleibend hohen Aufmerksamkeitszuwendung der Versuchspersonen (vgl. Bredenkamp und Erdfelder 1996). Sie könnte als Methode in einem Nachexperiment dienen, um Gesetzmäßigkeiten der begrifflichen Nähe von alltäglichen Begriffen zu überprüfen.

Primingaufgaben

Primingaufgaben nutzen das Phänomen der begrifflich-semanticen Nähe und sie werden zur Überprüfung unterschiedlichster Theorien und Hypothesen in der kognitiven Grundlagenforschung sowie der Wahrnehmungs- und Handlungsforschung eingesetzt.

Beschreibung von Primingaufgaben Der Versuchsperson wird zuerst ein Prime-Reiz dargeboten, darauf wird nach einer definierten Pause (SOA Stimulus Onset Asymmetrie) von 200ms, 500ms etc. ein Target-Reiz bzw. ein Distraktor dargeboten. Variiert werden der Prime- und der Target-Reiz sowie die Dauer der SOA-Pause. Die Pausendauer, sowie die Schnelligkeit, mit der der Target-Reiz bzw. der Distraktor identifiziert wird, sind dabei relevante Maße. Der Priming-Effekt geht davon aus, dass ein Wort schneller erkannt wird, wenn kurz vorher ein begrifflich nahes Wort vorgegeben worden ist. Durch die Identifikation des Prime-Reizes wird das begriffliche Umfeld in eine erhöhte Aktivität versetzt, sodass der Target-Reiz schneller erkannt wird.

Bewertung hinsichtlich einer Eignung als Erhebungsmethode Primingaufgaben wären hinsichtlich des Untersuchungsziels und der theoretischen Grundlagen geeignet und sie würden die präzisesten Ergebnisse liefern. Problematisch ist die Wirkung von prototypischen Begriffen (=Primärbegriffen). Sie werden schneller identifiziert als ihre Ober- oder Unterbegriffe. Bei der Interpretation der Reaktionszeiten müsste diese Wirkung berücksichtigt werden. Als Erhebungsmethode eignen sie sich dennoch nicht, da sie Laborbedingungen mit einem hohen, technischen Aufwand erfordern. Sie sind nicht nutzernah einsetzbar, was Probleme in der Versuchspersonenrekrutierung und somit Probleme für eine Erhebung von nutzerspezifischen Begriffsstrukturen mit sich bringen würde. Sie könnten eventuell als Methode in einem Nachexperiment dienen, um Gesetzmäßigkeiten über begriffliche Nähe von alltäglichen Gegenständen an kleinen Personengruppen zu überprüfen.

Begriffsbestimmungsaufgaben

Untersuchungsziel ist die Überprüfung von Hypothesen zu Repräsentationsannahmen in semantischen Netzwerk-, Propositions-, Prototypen- und Merkmalsrepräsentationsmo-

dellen. Zimmer (1984) verwendete z.B. Begriffszuordnungsaufgaben, um Klassifikationsleistungen in Abhängigkeit von den begrifflichen Wissensstrukturen von Experten und Laien zu prüfen. Begriffsbestimmungsaufgaben werden auch verwendet, um Hypothesen über den Einfluss von begrifflichem Wissen bei der Objektidentifikation (begriffliche Identifikation vgl. Hoffmann 1983) zu prüfen. Bei Begriffsbestimmungsaufgaben wird im Labor die zeitabhängige Findung bzw. Zuordnung von Begriffen untersucht. Von den Antwortzeiten wird auf die Art der Begriffsrepräsentation geschlossen.

Beschreibung der Begriffsbestimmungsaufgaben Die Vp erhält die Beschreibung eines Gegenstandes und muss die dazu passende Wortmarke finden. Ausgewertet werden die verschiedenen Äußerungen in Abhängigkeit von der Zeit. Die Methode wird auch Produktionsmethode genannt.

Beschreibung der Begriffszuordnungsaufgabe Der VP wird ein Bild vorgelegt und daran anschließend schriftlich das Wort. Die Vp muss entscheiden, ob Wort und Bild zusammenpassen. Gemessen wird die Reaktionszeit und die Richtigkeit der Antwort.

Bewertung hinsichtlich einer Eignung als Erhebungsmethode Begriffsbestimmungsaufgaben eignen sich hinsichtlich des Untersuchungsziels nicht, da sie hierarchisch-taxonomische Begriffsbeziehungen abtesten. Sie erfordern überdies Laborbedingungen, was einen nutzernahen Einsatz nicht möglich macht. Sie stellen Probleme bei der Parallelisierung des sprachlichen bzw. bildlichen Untersuchungsmaterials. Und es ist nicht geklärt, welchen Einfluss die linke bzw. die rechte Hemisphäre des Gehirns auf die Antwortzeiten hat.

Relationsbestimmungsaufgaben

Relationsbestimmungsaufgaben wurden entwickelt, um die unterschiedlichen Beziehungsarten zwischen Begriffen zu untersuchen (vgl. Klix 1984). Van der Meer (1991, 1996) verwendete darüber hinaus Analogieerkennungsaufgaben, um Relationstypen zwischen ereignisbezogenen Begriffsstrukturen (Folgen von Ereignisbegriffen) zu untersuchen.

Beschreibung der Semantischen Kategorisierungsaufgabe Der Vp wird eine Serie von Begriffen dargeboten, dann wird der Oberbegriff genannt und die VP muss beurteilen, ob die vorgegebene Oberbegriffsrelation zutreffend ist (Klimesch 1989).

Beschreibung von Analogieerkennungsaufgaben Der Vp werden zwei schriftliche oder sprachliche Analogiepaare vorgelegt; z.B. „Fluss : Bach“ wie „See : Tümpel“ Die Vp muss entscheiden, ob die Analogiepaare einander entsprechen. Gemessen wird die Zeit und die inhaltliche Richtigkeit der Antwort (Kukla, zitiert nach Klix 1984).

Bewertung hinsichtlich einer Eignung als Erhebungsmethode Relationsbestimmungsaufgaben würden über das Untersuchungsziel hinausschießen, da sie nicht nur die begriffliche Nähe untersuchen, sondern auch die Relationstypen zwischen Begriffen unterscheiden würden. Entsprechend der Hypothesen reicht es aber aus, die begriffliche Nähe zwischen den Begriffen zu ermitteln. Eine inhaltliche Differenzierung hinsichtlich des Relationstyps ist gemessen am Aufwand nicht nötig. Diese Methodenklasse erfordert überdies Laborbedingungen, was Probleme in der Versuchspersonenrekrutierung mit sich bringen würde.

Zusammenfassung. Assoziations- und Erzählaufgaben sind als Methoden für den Hauptversuch aus unterschiedlichen Gründen wenig geeignet. Sie sind hinsichtlich der Theorie und des Untersuchungsziels wenig geeignet. Sie bringen Probleme hinsichtlich des Untersuchungsgegenstandes (assoziative Verfügbarkeit, Vorhandensein einer Wortmarke etc.). Sie sind in der Auswertung subjektiv belastet und die Ergebnisse sind undifferenzierter als die von Sortieraufgaben. Assoziationsaufgaben eignen sich allerdings als Methode für den Vorversuch, um die Objekte für den Hauptversuch auszuwählen. Begriffsbestimmungsaufgaben eignen sich hinsichtlich des Untersuchungsziels nicht, weil sie nur hierarchische Begriffsbeziehungen abtesten. Priming-, Satzverifikations- und Relationsbestimmungsaufgaben würden präzise Ergebnisse liefern, sind jedoch aufgrund der Laboranforderungen nicht geeignet (Rekrutierungsprobleme). Sie könnten in Nachuntersuchungen eingesetzt werden, um die Gesetzmäßigkeiten über semantische Nähe von Alltagsbegriffen an kleinen Stichproben im Experiment zu überprüfen.

6.2.4 Aufgaben mit Ähnlichkeitsurteilen als geeignete Methodenklasse

Im Folgenden werden Aufgaben mit Ähnlichkeitsurteilen vorgestellt. Ziel dieser Methodenklasse ist es, die Versuchspersonen die Ähnlichkeit des Reizmaterials (Worte, Bilder, Objekte) beurteilen zu lassen und anhand der Ähnlichkeitsbeurteilungen auf die begriffliche Nähe bzw. Distanz der Begriffe und die zugrunde liegenden Klassifizierungsdimensionen zu schließen. Es gibt unterschiedliche Aufgabentypen, die im Folgenden vorgestellt werden:

Beschreibung von Ähnlichkeits-Paarvergleich bzw. Triadenvergleich Die Versuchsperson soll beim Paarvergleich zwei Objekte bzw. beim Triadenvergleich drei Objekte auf einer 5-stufigen Skala von „äußerst ähnlich“ bis „äußerst unähnlich“ beurteilen. Der Person werden keine Kriterien/Merkmale vorgegeben. Die Auswertung erfolgt üblicherweise mittels der (nonmetrischen) Multidimensionalen Skalierung und/oder der Clusteranalyse bzw. Hierarchischen Clusteranalyse (vgl. Borg 1981, Gigerenzer 1981, Bortz 1984, Bortz & Döring 1995).

Beschreibung von Sortieraufgaben Der Versuchsperson (Vp) werden Gegenstände oder Kärtchen mit Begriffen vorgelegt, die sie sortieren soll. Es werden keine Sortierkriterien vorgegeben. Es gibt unterschiedliche Sortiervarianten, z.B.:

Freies Sortieren Die Vp soll diejenigen Begriffe, die sie für ähnlich hält, nebeneinander legen, wie bei einem Dominospiel, bis alle Begriffe in einer Reihe liegen. Doppelnennungen sind nicht möglich. Abbildung 6.1 veranschaulicht den Sortiervorgang.

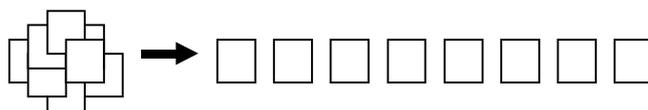


Abbildung 6.1. Veranschaulichung des Freien Sortierens: Die Vp soll die ungeordneten Karten mit Begriffen in die für sie „richtige“ Nachbarschaftsreihenfolge legen.

Free Card Sorting hier soll die Vp die Begriffe zu Gruppen sortieren, die ihr vernünftig erscheinen. Sie kann Begriffe auch mehrfach zuordnen (vgl. Zimmer 1973). (Abb. 6.2).

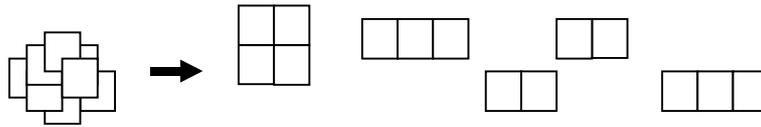


Abbildung 6.2. Veranschaulichung des Free Card Sorting - Die ungeordneten Karten mit Begriffen sollen in die „richtige“ Gruppen sortiert werden.

Hierarchisches Sortieren hier soll die Vp die Begriffe in zwei Gruppen teilen, wobei die Gruppen nicht gleich groß sein müssen. Die Gruppen müssen dann wieder in zwei Gruppen geteilt werden, usw. Dieser Prozess setzt sich so lange fort, bis in den Gruppen nur noch jeweils ein Begriff ist (Abb. 6.3).

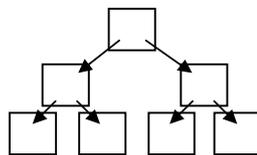


Abbildung 6.3. Veranschaulichung der Sortierschritte beim Hierarchischen Sortieren. Die Objekte bzw. Begriffe auf Kärtchen werden so lange in zwei Gruppen geteilt, bis sie alleine stehen.

Die Versuchsperson wird im Sinne einer "forced choice" gezwungen, sich bei jedem Sortierschritt auf ein Sortierkriterium zu beschränken. Sie kann aber bei jedem Sortierschritt ein anderes wählen. Dies hat den Vorteil, dass die Versuchsperson jeweils das ihr wichtigste Kriterium wählen kann (vgl. Marx & Hejj 1989).

Ausgangspunkt dieser Methoden ist die Annahme der klassischen Begriffsforschung, dass Begriffe und Begriffsstrukturen aufgrund von Klassifizierungsprozessen entstehen, denen Ähnlichkeitsbeurteilungen und Abstraktionsprozesse zugrunde liegen. Aufgaben, die Ähnlichkeitsurteile abverlangen, wurden bereits in den ersten Abstraktionsexperimenten von Hull (1920) zur Überprüfung der Abstraktionstheorie eingesetzt, später dann in unterschiedlichsten psychologischen Fachbereichen, z.B. in der Psychophysik, in der Psycholinguistik zur Analyse der Semantischen Struktur von Wortfeldern, in der Sozialpsychologie zur Analyse von Einstellungsstrukturen und schließlich auch in der Wissenspsychologie zur Analyse von begrifflichen Wissensstrukturen.

In kognitionspsychologischen Grundlagenexperimenten verwendeten z.B. Kukla und Sydow (1970) Ähnlichkeitsurteile, um die Lernleistungen für begriffsanaloge Klassifi-

zierungsprozesse zu überprüfen. Die Versuchspersonen mussten nicht bekannte Objekte nach Merkmalen klassifizieren lernen, die der Versuchsleiter vorgegeben hatte und die den Versuchspersonen zunächst wenig eingängig waren. Vor und nach dem Lernen wurden Ähnlichkeitsskalierungen durchgeführt. Die erlernte Klassenbildung spiegelte sich in der Veränderung der Ähnlichkeitsbeurteilung der phänomenalen Metrik wieder.

Untersuchungstechnisch eignen sich diese Methoden für den Einsatz im Feld. Sie setzen voraus, dass die Versuchspersonen die zu beurteilenden Begriffe kennen, was vor der Ähnlichkeitsbeurteilung abgeklärt werden muss. *Paar- und Triadenvergleich* liefern besonders differenzierte Ergebnisse (vgl. Bortz 1984, Bortz & Döring 1995). Sie sind jedoch in der Durchführung für die Versuchspersonen zeitintensiv und kognitiv belastend, sodass sie nur eine geringe Anzahl von zu untersuchenden Begriffen erlauben. Nur 12 zu beurteilende Begriffe erfordern 66 Paarvergleiche bzw. 220 Triadenvergleiche. Als Erhebungsmethoden für die geplanten Untersuchungen eignen sie sich daher nicht, da in jeder Studie mehr als 12 Begriffe untersucht werden sollen.

Sortiertechniken erlauben im Hinblick auf die zeitliche und kognitive Belastung der Versuchspersonen – je nach Sortier- und Auswertungsmethode – bis zu 100 zu untersuchende Begriffe. In methodenvergleichenden Untersuchungen (vgl. Untersuchungen zur Begriffsstruktur von Säugetieren von Henley 1969, Storm 1980, zitiert nach Hejj & Strube 1989) konnte gezeigt werden, dass Sortiertechniken zwar weniger differenzierte Ergebnisse als Triaden- bzw. Paarvergleiche liefern, aber weitaus differenziertere Strukturen als Assoziationsaufgaben.

Hierarchisches Sortieren und *Free Card Sorting* entsprechen inhaltlich Erkenntnissen der modernen Begriffsforschung über natürliche Begriffe, weil sie Mehrfachzuordnungen (Kreuzklassifikationen) erlauben: Wurden in der traditionellen Begriffsbildungsforschung noch eindeutige Zuordnungen entsprechend den Regeln der Booleschen Algebra gefordert („Alles-oder-Nichts“-Regel, disjunkte Klassen etc.), so setzte sich später die Erkenntnis durch, dass „natürliche“ Begriffe nicht als klar umgrenzte Boolesche Mengen von Objekten mit eindeutigen Zuordnungskriterien zu beschreiben sind, sondern vielmehr als „unscharfe“ Mengen mit Objekten unterschiedlicher Zugehörigkeit. (vgl. Zadeh 1965, Bierwisch 1974 zitiert nach Hoffmann 1993 S. 124) Insofern kommen diese Sortiermethoden Eigenschaften natürlicher Begriffe entgegen. Das Hier-

archische Sortieren erlaubt der Vp, bei jedem Sortierschritt ein anderes Sortierkriterium zu wählen. Zu beachten ist bei der Interpretation der Ergebnisse, dass die Versuchsperson bei jedem Sortierschritt zu einer Differenzierung der Beurteilungskriterien „gezwungen“ wird, die sie normalerweise sonst nicht machen würde.

Das *Free-Card-Sorting* erlaubt eine Mehrfachzuordnung und fordert von den Versuchspersonen nicht eine künstliche Differenzierung. Diese Sortiertechnik wurde ursprünglich von Shipstone (1960) entwickelt. Allerdings gab Shipstone seinen Versuchspersonen noch die Kriterien vor, nach denen sie die Gruppen bilden sollten. Miller (1969) entwickelte die Methode weiter, indem er die Versuchspersonen selbst das Klassifizierungskriterium wählen ließ, ließ aber einen Begriff nur einmal zuordnen. Er forderte disjunkte Klassen. Zimmer (1973) entwickelte die Methode wiederum weiter und forderte von seinen Versuchspersonen keine disjunkte Klassenbildung mehr: sie konnten einen Begriff mehreren Gruppen zuordnen (vgl. Wecker 1994). Diese ”zunehmende” Freiheit beim Sortieren entspricht auch zeitlich der sich verändernden Erkenntnis über Begriffsbildungsprozesse, dass sich natürliche Begriffe von Booleschen Klassifikationen unterscheiden. Die Methode des Freien Sortierens ermöglicht keine Mehrfachzuordnung, ist dafür aber schneller und einfacher in der Durchführung.

Bewertung hinsichtlich einer Eignung als Erhebungsmethode Sortierverfahren eignen sich dazu die Hypothesen aus Kapitel 5 zu testen. Sie fordern Ähnlichkeitsurteile, durch die sich begriffliche Nähe bzw. Distanzen zwischen Begriffen erheben lassen. Sie passen auch zur Theorie von Klix, Hoffmann et al. Sortierverfahren wurden ursprünglich in Begriffstheorien (Hull 1920) eingesetzt, die Merkmalsabstraktionsprozesse als Grundlage der Begriffsbildung annehmen. Die verhaltensorientierten Merkmalsrepräsentationstheorien von Klix (1971,1992), Hoffmann (1983,1986) et al. legen der Begriffsbildung auch Abstraktionsprozesse zugrunde. Die Sortiermethoden Free Card Sorting und Hierarchisches Sortieren kommen Anforderungen der natürlichen Begriffe (Mehrfachzuordnungen) besonders entgegen und geben der Versuchsperson die Freiheit, sowohl nach hierarchisch-taxonomischen als auch ereignisbezogenen Klassifizierungskriterien zu sortieren. Sortierverfahren liefern zwar hinsichtlich der Art und des Inhalts der begrifflichen Relationen (z.B. Instrumentalrelation, Ortsrelation, Zeitrelation etc.) keine differenzierten Aussagen. Sortierverfahren sind einfacher in der Durchführung und weniger

anfällig in der Auswertung für Einflüsse durch Versuchsleiter als z.B. Assoziationsverfahren.

Sortieraufgaben erfüllen die Anforderungen hinsichtlich des Untersuchungsgegenstandes: sie können sowohl mit realen Objekten, anschaulichen Abbildungen als auch bekannten, sprachlichen Wortmarken durchgeführt werden. Sortieraufgaben eignen sich auch hinsichtlich der Versuchspersonenauswahl: Sie sind nutzernah, vor Ort einsetzbar, sodass auch rekrutierungssensible Nutzergruppen (z.B. Hausfrauen, Manager etc.) als Versuchspersonen gewonnen werden können. Sie belasten die Versuchspersonen zeitlich und kognitiv nicht übermäßig. Paar- und Triadenvergleich würden zwar differenziertere Ergebnisse als Sortierverfahren liefern. Man könnte mit ihnen aber nur eine kleine Anzahl (<8) von Begriffen abtesten, um die Versuchspersonen nicht zu überfordern. Insgesamt erscheinen Sortieraufgaben als geeignete Erhebungsmethoden. Im Folgenden werden die Auswertungsmethoden vorgestellt, die sich zur statistischen Analyse der Sortierergebnisse eignen.

6.3 Bestimmung der Auswertungsmethoden

Die Auswertung von Ähnlichkeitsurteilen aus Sortieraufgaben erfolgt üblicherweise mit beschreibenden, dimensionsanalytischen Verfahren, wie der *metrischen* bzw. *nonmetrischen multidimensionalen Skalierung* (MDS bzw. NMDS) und der *Hierarchischen Clusteranalyse* (HCA) (nach Johnson 1967), (vgl. Borg 1981, Bortz 1984, Gigerenzer 1981, Bortz & Döring 1995, Städler 1998). Die NMDS ist eine Weiterentwicklung der metrischen Multidimensionalen Skalierung MDS, um ordinal skalierte Daten verarbeiten zu können. Sie wurde von 1962 Shepard in den Bell Telephone Laboratories entwickelt, unweit der Princeton University, an der Gullikson und Torgerson, die 1938 von Young & Householder entwickelte klassische oder metrische MDS vervollständigten. Mit der ersten Anwendung der NMDS untersuchte Shepard die subjektive Organisation der Farbwahrnehmung. Kruskal und Guttman erarbeiteten dann die entscheidenden konzeptuellen und algorithmischen Verbesserungen des ersten Entwurfes von Shepard.

MDS bzw. NMDS sind statistische Verfahren, die Ähnlichkeitsurteile der Untersuchungspersonen auf Ordinal- bzw. Intervallskalenniveau in räumlichen Distanzmaßen abbilden. Die HCA ist ein statistisches Verfahren, um Ähnlichkeitsurteile auf Ordinalskalenniveau in Cluster abzubilden. Es sind Verfahren, die für die Berechnung von Urteilsdaten („ist ähnlich“ versus „ist „unähnlich“ bzw. ”gehört zusammen“ versus ”gehört nicht zusammen“) geeignet sind. Der MDS bzw. NMDS und der HCA liegen unterschiedliche Modellannahmen und Metriken zugrunde, sodass beide Verfahren unterschiedliche Informationen über die gleichen empirischen Daten bringen. Da beide Verfahren beschreibende, dimensionsanalytische und somit qualitative und keine quantitativen Verfahren sind, ist die Interpretation der Ergebnisse zu beachten.

Solche flachen, datennahen Modellierungen gelten als gut geeignet für den Vergleich von Gruppen (Strube 1984). Es liegen in der Literatur zahlreiche Untersuchungen zur Analyse von Begriffsstrukturen vor, in denen Ähnlichkeitsurteile mittels MDS bzw. NMDS und HCA ausgewertet werden: Henley (1969), Rips, Shoben & Smith 1973), Marx (1988), Marx & Hejj (1989) u.a. verwendeten die Methoden für Begriffsfeldstudien (z.B. von Tieren, Berufen, Verwandtschaftsbegriffen etc.). Downs & Stea (1982) untersuchten damit mentale Landkarten. Kukla & Sydow (1970) verwendeten sie, um die Abhängigkeit der Begriffsbildung von Lernzielen zu zeigen. Klix (1971) verweist auf die funktionelle Bedeutung von MDS und NMDS zur Untersuchung von Begriffsstrukturen. Hoffmann (1980) untersuchte mit der Methode der freien Reproduktion an 20 Begriffen und der hierarchischen Clusteranalyse von Johnson (1967), ob die Reproduktionsleistung von zwischenbegrifflichen (= ereignisbezogenen) Relationen abhängt. Im Folgenden werden die Auswertungsverfahren einführend dargestellt. Für eine vertiefende Darstellung der Verfahren wird verwiesen auf Borg (1981) und Gigerenzer (1981).

6.3.1 Metrische und Nonmetrische Multidimensionale Skalierung

Das Ziel der metrischen bzw. nonmetrischen multidimensionalen Skalierung besteht darin, die Ähnlichkeitsurteile in räumlichen Distanzrelationen abzubilden und, in ihre zugrunde liegenden Dimensionen zu zerlegen (Gigerenzer 1981 S.321). Die Ähnlich-

keitsurteile der Versuchspersonen werden dabei in eine Rangreihe gebracht und dann in einer Punktekonfiguration verarbeitet.

Modellannahmen Bei dieser Suche nach den Distanzrelationen und Dimensionen werden Verknüpfungsregeln vorausgesetzt. Als Modell für die Zerlegung der Ähnlichkeitsrelationen wird die Minkowski r -Metrik gewählt (zur inhaltlichen Begründung der Wahl der Minkowski r -Metrik siehe: Gigerenzer 1981), deren bekanntester Sonderfall die euklidische Metrik ist (Gigerenzer 1981 S. 324). Die Wahl einer euklidischen Metrik als Modellvoraussetzung für Prozesse der kognitiven Informationsverarbeitung mag für andere psychologische Fragestellungen kritisch und diskussionswürdig sein. Für diese Studie sind *MDS bzw. NMDS* hingegen *inhaltlich* in Bezug auf das Untersuchungsproblem deshalb *gut geeignet*, weil sie der Fragestellung: „In welchen Distanzrelationen muss was im Raum platziert sein?“ *intuitiv entsprechen*. *MDS bzw. NMDS wandeln psychologische Ähnlichkeitsurteile in räumliche Distanzmaße* um und bilden somit einen begrifflich psychologischen Raum in einen euklidischen Raum ab. Die Modellannahme der *MDS bzw. NMDS*, dass semantische Ähnlichkeitsurteile in euklidische, räumliche Distanzrelationen übersetzt werden, entspricht hier genau der Aufgabenstellung. Auch hinsichtlich der Implikationen auf Begriffstheorien sind sie theoriekonsistent: sie passen zur Basisannahme der Merkmalsabstraktion; sie erlauben, die begrifflich semantische Nähe bzw. Ähnlichkeit von Begriffen zu erheben und sie berücksichtigen kategoriale Mehrfachzuordnungen bzw. Kreuzklassifikationen. Die Ergebnisse sind jedoch nicht mit realen Begriffsstrukturen gleichzusetzen. Sie stellen ein methodenabhängiges Abbild dar.

Prozedere *MDS bzw. NMDS* berechnen die Abstände (= Distanzen) zwischen zwei Ähnlichkeitsurteilen (Gigerenzer 1981 S. 324). Sie setzen bei jedem Rechenschritt alle Ähnlichkeitsurteile zueinander in Beziehung. Dabei werden schrittweise 2-, 3-, 4- und mehrdimensionale Punktekonfigurationen berechnet und so lange verändert, bis die euklidischen Distanzen zwischen den Punkten in der Punktekonfiguration mit der Rangreihe der empirisch gefundenen Ähnlich- bzw. Unähnlichkeiten möglichst gut übereinstimmen (vgl. Gigerenzer 1981, Bortz 1984). Ergebnis einer *MDS bzw. NMDS* ist eine Punktekonfiguration.

Aus den berechneten 2-, 3-, 4-, 5-dimensionalen Punktekonfigurationen muss dann diejenige Punktekonfiguration ausgewählt werden, die bei möglichst geringer Anzahl von Dimensionen die Ähnlichkeitsurteile mit dem geringsten Anpassungsdruck einpasst. Hierfür werden zwei Kriterien angelegt: Einmal soll die räumliche Punktekonfiguration der MDS bzw. NMDS so wenig Dimensionen wie möglich aufweisen (*Ökonomiebedingung*). Optimal sind 2- oder 3-dimensionale Darstellungen, weil sie einfacher zu verstehen sind als 4- oder 5-dimensionale Punktekonfigurationen. In Bezug auf das Anordnungsproblem sind die 2- bzw. 3-dimensionalen Konfigurationen auch deshalb von Vorteil, weil sie sich einfacher in räumliche Anordnungen übersetzen lassen. Zum anderen sollen aber auch die Rangordnungen der Ähnlichkeitsurteile den Rangordnungen der Distanzen im räumlichen Modell entsprechen (*Monotoniebedingung*) oder einfacher formuliert: die räumlichen Distanzen in der NMDS-Konfiguration sollen den empirischen Distanzen der Ähnlichkeitsurteile entsprechen.

$$\delta_{ab}^m = d_{ab}$$

wobei: δ_{ab} = Unähnlichkeiten zwischen den Objekten a und b
 d_{ab} = Distanz zwischen a und b

δ_{ab}^m = ist gleich monotone Distanzfunktion

Die Güte zwischen Anpassung der Ähnlichkeitsurteile in räumliche Distanzen und der Anzahl der Dimensionen wird mittels Scree-test ermittelt. Das Gütemaß ist der "Stresswert" (Alienation). Der Stresswert gilt nach Sherman (1972; zitiert aus Gigerenzer 1981) als die best geeignete Maßzahl für die Anpassungsgüte. Kruskal (1964; zitiert nach Gigerenzer 1981) definierte ihn im Sinne der traditionellen Definition von Fehlermaßen als eine Funktion der Summe der quadrierten Abweichungen.

$$\text{Stress } S = \sqrt{\frac{\sum_{a < b} (d_{ab} - \hat{d}_{ab})^2}{\sum_{a < b} d_{ab}^2}}$$

Der Stresswert (= Alienation), der zwischen 0 und 1 liegt, ist somit ein Maß für die Anpassungsgüte des berechneten Modells an die Ausgangsdaten. Ein niedriger Stresswert zeigt eine hohe Anpassungsgüte an, ein hoher Stresswert eine schlechte Anpassungsgüte. Über Simulationsstudien (Monte-Carlo-Studien) wurden kritische Stress-

werte ermittelt, die anzeigen, ob eine Rangordnung der Unähnlichkeiten im empirischen System zufällig zustande gekommen ist oder nicht. Gigerenzer gibt eine Liste mit diesen kritischen Stresswerten an (1981 S. 353). Ist der berechnete Stresswert größer als der kritische Stresswert, so ist die Nullhypothese beizubehalten, die behauptet, dass die Rangordnungen zufällig zustande gekommen sind. Ist der berechnete Stresswert kleiner als der kritische Stresswert, zeigt dies eine zufallskritisch abgesicherte Konfiguration an; d.h. die Distanzrelationen bzw. die Ähnlichkeitsurteile sind nicht zufällig zustande gekommen, sondern sie sind bedeutsam und stabil. Die Nullhypothese darf verworfen werden. Aus den 2-, 3-, 4-, 5-dimensionalen Punktekonfigurationen wird somit diejenige ausgewählt, die bei möglichst geringer Anzahl von Dimensionen den geringsten, signifikanten Stresswert aufweist. Sie gilt als die beste Repräsentation der untersuchten Begriffsbeziehungen (vgl. Gigerenzer 1981; Bortz 1984, S. 112).

Zur Interpretation der MDS bzw. NMDS-Konfigurationen

Die MDS- bzw. NMDS-Konfiguration zeigt eine Anordnung von Punkten, wie im Beispiel Abbildung 6.4 zu sehen ist.

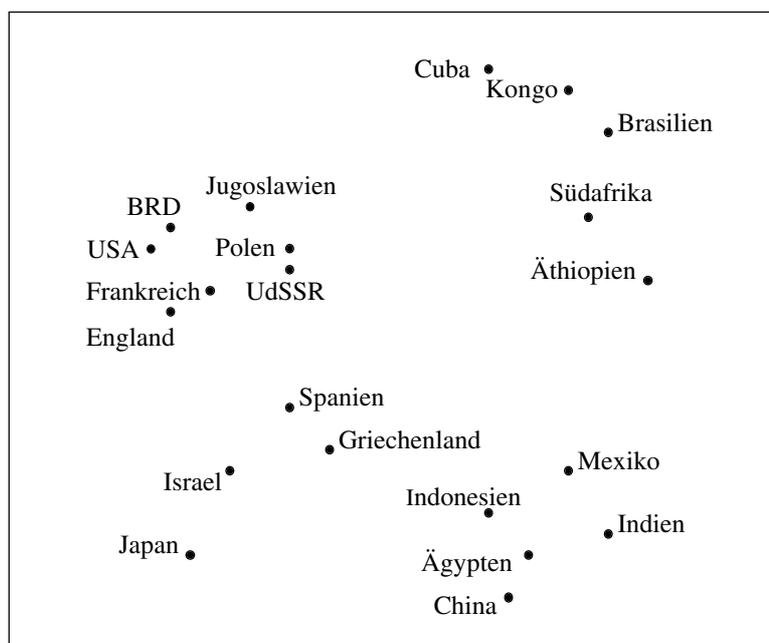


Abbildung 6.4. Beispiel einer subjektiven Landkarte als NMDS-Konfiguration (entnommen aus Gigerenzer 1981)

Bei den Untersuchungen dieser Arbeit repräsentieren sie die auf Ähnlichkeit beurteilten Objekte. Für die Interpretation der Distanzrelationen zwischen den Punkten der MDS- bzw. NMDS-Konfigurationen ist Folgendes zu beachten: Die räumliche Nähe bzw. Distanz zwischen den Punkten stellt die Ähnlichkeits- bzw. Unähnlichkeitsrelation dar; was als ähnlich beurteilt wird, wird räumlich nah beieinander dargestellt bzw. umgekehrt. Die räumliche Ausrichtung, das Koordinatenkreuz hat dabei keine inhaltliche Bedeutung. Die MDS bzw. NMDS-Konfiguration kann beliebig rotiert, translatiert, dilatiert bzw. gespiegelt werden, solange die Relationen invariant bleiben, d.h. nicht verändert werden (vgl. Borg 1981). Bei der Interpretation der räumlichen Distanzrelationen muss außerdem berücksichtigt werden, dass einzelne Begriffe auch unter einem hohen Stress eingepaßt worden sind. Da man nicht weiß, welche Begriffe dies sind, kann man zur Absicherung die Originaldaten noch einer zweiten statischen Analyse der HCA (Hierarchischen Clusteranalyse) unterziehen und deren Ergebnisse mit der MDS bzw. NMDS-Konfiguration vergleichen.

6.3.2 Hierarchische Clusteranalyse

Im Folgenden wird ein Überblick über die HCA gegeben. Eine ausführliche Darstellung findet sich bei Gigerenzer (1981) oder Bortz (1984).

Modellannahmen Die Hierarchische Clusteranalyse (nach Johnson 1967) ist ein statisches Verfahren zur systematischen Klassifizierung einer Objektmenge – ähnlich einer Faktorenanalyse (vgl. Bortz 1984). Bei der Hierarchischen Clusteranalyse (nach Johnson 1967) werden die Originaldaten, d.h. die empirisch ermittelten Ähnlichkeitsurteile erneut ausgewertet. Durch einen anderen statistischen Ansatz mit anderen Modellannahmen – als ihn die NMDS verwendet – werden die Ähnlichkeitsurteile ebenso in Distanzmaße übersetzt, aber nicht in euklidische Distanzen. Die hierarchische Clusteranalyse ist eine monotone Transformation der Unähnlichkeiten in Distanzen, welche der ultrametrischen Ungleichung genügen (vgl. Gigerenzer 1981). Die ultrametrische Ungleichung besagt, dass die Distanz zweier Objekte nie größer ist als deren Distanzen zu einem beliebigen dritten Objekt. Die Ungleichung präzisiert die beschriebene hierarchische Zusammenfassung von Objekten zu umfangreicheren Clustern. Die HCA kann

somit zusätzliche „Varianz“ erklären und Informationen liefern, die mit der NMDS alleine nicht sichtbar werden. Auch die HCA ist inhaltlich in Bezug auf die Untersuchungshypothesen geeignet, weil sie der Fragestellung „Was gehört zusammen? Was muss zusammen platziert sein?“ intuitiv entspricht. Sie wandelt die Ähnlichkeitsurteile in Gruppierungen, d.h. Cluster um.

Prozedere Bei der Berechnung werden als Erstes diejenigen Begriffe schrittweise zusammengeslossen, deren Unähnlichkeit den kleinsten Wert hat. Die nächstähnlichen Begriffe werden bereits bestehenden Clustern zugeordnet oder wiederum zu neuen Clustern zusammengefasst. Diese Prozedur dauert so lange, bis alle Begriffe in hierarchisch geordneten, disjunkten Clustern erfasst sind. Ergebnis der HCA ist eine hierarchische Baumstruktur. Abbildung 6,5 zeigt einen Ausschnitt aus einer HCA. Die beiden Subcluster USA-BRD und England-Frankreich sind in der hierarchischen Baumstruktur als Cluster zu erkennen. Sie sind durch hinterlegtes Grau noch einmal visualisiert.

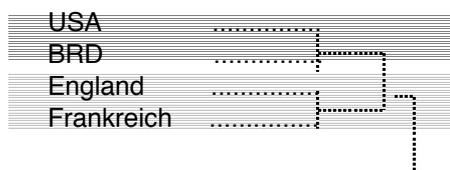


Abbildung 6.5. Beispiel Ausschnitt aus der Hierarchischen Clusteranalyse „Subjektive Landkarte“ (entnommen aus Gigerenzer 1981)

Integration der HCA- und NMDS-Ergebnisse Die Ergebnisse der HCA werden dann in den NMDS-Ausdruck eingezeichnet: es werden diejenigen Punkte (=Begriffe) mit Kreisen zusammengefasst, die auch in der HCA-Berechnung zusammengehören. Wurde nun ein Begriff bei der NMDS-Berechnung unter einem hohen Anpassungsdruck in die Punktekonfiguration eingepasst, so kann das nun – in der Kombination mit den HCA-Ergebnissen – zur Folge haben, dass der Kreis Beulen bekommt, um einen im NMDS-Ausdruck entfernten Begriff mit einzufassen. Dies bedeutet dann lediglich, dass der entfernte Begriff bereits unter großem Stress in die NMDS-Konfiguration eingepasst wird. Oder eben, dass sich die Ergebnisse von NMDS und HCA nicht ganz einig sind! Abbildung 6.6 zeigt an einem Beispiel die integrierten Ergebnisse von HCA und NMDS.

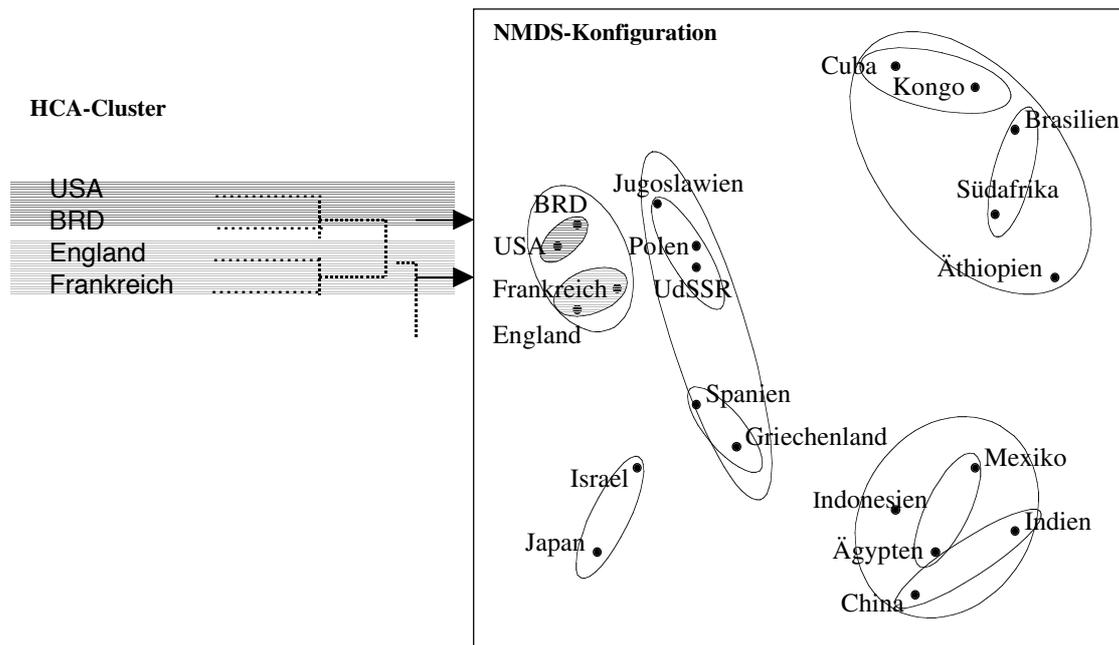


Abbildung 6.6. NMDS-Konfiguration mit integrierten HCA-Ergebnissen „Subjektive Landkarte“ (entnommen aus Gigerenzer 1981)

In einem nächsten Schritt können nun die Dimensionen von MDS bzw. NMDS inhaltlich interpretiert und benannt werden. Für die Interpretation werden mehrere Herangehensweisen empfohlen, da die Interpretation, wie bei allen deskriptiven, dimensionsanalytischen Verfahren, problematisch ist, weil Versuchsleiterinteressen die Interpretation beeinflussen können (vgl. Gigerenzer 1981). Traditionell erfolgt zur Interpretation der Dimensionen eine Rotation auf Einfachstruktur mit anschließender phänomenologischer Deutung der rotierten Dimensionen. Gigerenzer (1981) weist auf das Problem des forschungsleitenden Interesses und der impliziten Vorannahmen hin, wenn der Forscher quasi "selbständig" versucht, die Dimensionen zu benennen und zu erklären. Zur Objektivierung empfiehlt er eine empirische Interpretationsstrategie, z.B. durch eine Befragung der Versuchspersonen oder durch eine Property Fitting Analyse (vgl. Gigerenzer 1981).

Gigerenzer schlägt einmal eine *Befragung der Versuchspersonen* vor, d.h. die Versuchsperson wird nach jedem Sortierschritt nach ihrem Beurteilungskriterium gefragt. Diese Nennungen werden festgehalten und bei der Benennung der Dimensionen unterstützend herangezogen. Als weitere Methode schlägt er eine *Property Fitting Analyse*

vor, die in einem separaten Nachversuch gemacht werden muss. Hierbei müssen Versuchspersonen die beim Sortieren genannten Sortierkriterien in eine Rangreihe legen. Die Property Fitting Analyse rotiert dann die einzelnen Rangreihen so lange in der NMDS-Konfiguration, bis eine maximale Korrelation mit den Skalenwerten der in der NMDS-Konfiguration abgebildeten Begriffe erreicht ist. Diese Dimensionen werden dann in die Punktekonfiguration als Linien (Pole) eingepasst. Bei diesen beiden empirisch gestützten Verfahren wird davon ausgegangen, dass die Versuchspersonen das Sortierkriterium nennen, nach dem sie auch tatsächlich sortieren.

Es wird also eine Übereinstimmung von verbalen Äußerungen und faktischem Sortierverhalten unterstellt. Es kann aber auch sein, dass die Ähnlichkeitsbeurteilung der Versuchsperson vielschichtigeren, „impliziten“ Kriterien folgt und die verbalen Benennung nicht das wiedergeben, was die Versuchsperson tatsächlich tut. (In der Studie über Zimmerpflanzen wurde eine Property Fitting-Analyse durchgeführt, um hiermit Erfahrungen zu sammeln.)

Hypothesengeleitete Interpretationsstrategie Zur Interpretation wird ein hypothesengeleitetes Vorgehen favorisiert. Die Ergebnisse werden in Bezug auf die Prognosen (im Sinne von Hypothesen), die auf der Basis der Begriffstheorie von Klix, Hoffmann et al. (vgl. Kapitel 5) entwickelt wurden, interpretiert. Die Ergebnisse der NMDS und HCA werden also inhaltlich dahingehend interpretiert, ob die Ähnlichkeitsbeurteilungen kategorialen Kriterien (im Sinne von Ober-, Unter-, Nebenbegriffen) folgen oder handlungsbezogenen Kriterien.

6.4. Übersicht über Methoden für den Vorversuch

Der Vorversuch hat das Ziel, diejenigen Objekte auszuwählen, die die Versuchspersonen im Hauptversuch sortieren sollen. Die Auswahl sollte folgende Aspekte berücksichtigen: 1) Die Anzahl der zu untersuchenden Objekte muss begrenzt werden, um die Versuchspersonen und die Methoden nicht zu überlasten. 2) Die Auswahl sollte inhaltlich das gesamte Begriffsfeld repräsentieren. 3) Die Versuchspersonen müssen mit den

zu sortierenden Objekten typische Handlungserfahrung gemacht haben. 4) Wenn Wortmarken oder Abbildungen verwendet werden, müssen diese eindeutig verstanden werden.

In der Literatur werden häufig assoziative Verfahren mit Inhaltsanalyse verwendet, um Objekte für den Ähnlichkeitsvergleich auszuwählen (vgl. Marx & Hejj 1989). Assoziative Verfahren haben aber die hier nachteilige Eigenschaft, dass die Versuchspersonen das nennen, was begrifflich nahe mit dem Reizwort verankert ist. Hier sollten aber Begriffe untersucht werden, die zusätzlich eine *repräsentative* Teilmenge aus der Grundgesamtheit der Begriffe darstellen, um Erkenntnisse für das gesamte Begriffsfeld zu gewinnen. Geeignete Verfahren sind:

Auswahl durch die Liste mit Zufallszahlen Bei dieser Auswahlmethode werden die Elemente der Grundgesamtheit nummeriert und zufällig anhand einer Liste an Zufallszahlen ausgewählt. Mit Hilfe von Zufallszahlen (siehe Bortz 1984 S. 543) wird aus dieser Liste die entsprechende Anzahl von zu untersuchenden Elementen ausgewählt. Die Methode ist einfach und verspricht eine zufällige Auswahl.

Befragung von Nutzern und Experten Zunächst werden die Nutzer nach ihren Nutzungskriterien befragt. Diese Kriterien dienen dann zur Auswahl einer Teilmenge aus der Grundgesamtheit der Begriffe. Um zu klären, ob diese Teilmenge tatsächlich repräsentativ in Hinblick auf die Grundgesamtheit ist, werden Experten zur Beurteilung der Auswahl befragt. Im Anschluss daran werden noch einmal Nutzer hinsichtlich der Bekanntheit der ausgewählten Begriffe befragt. Dieses Vorgehen wird gewählt, damit die Untersuchungsobjekte den Versuchspersonen bekannt sind und damit die Auswahl für das Begriffsfeld repräsentativ ist.

Restringierte Assoziation Bei der restringierten Assoziation (vgl. Strube 1984, S. 22) wird der Versuchsperson ein Schlüsselreiz vorgegeben. Im Unterschied zur freien Assoziation, bei der alle Antworten ausgewertet werden, werden hier nur Nennungen ausgewertet, die zur gesuchten Kategorie gehören. Diese Methode wird verwendet, um Geschäftstypen zu untersuchen, die typischerweise auf das Schlüsselwort „Einkaufszentrum“ genannt werden.

6.5 Zusammenfassung

In diesem Kapitel wurde das Untersuchungsdesign für die Erhebung einer nutzerspezifischen, erwartungsgerechten Anordnung entwickelt. Es spezifiziert eine quasi-experimentelle Untersuchung mit zwei abhängigen (Ähnlichkeitsbeurteilung, Klassifizierungskriterien) und zwei unabhängigen Variablen (Handlungsvorerfahrung, aktuelle Zielorientierung). In einem Vorversuch ist (mittels Zufallsliste, restringierter Assoziation und/ oder Befragung) eine für die zu untersuchende Objektmenge repräsentative Objektauswahl zu bestimmen und auf Bekanntheit zu überprüfen. Im Hauptversuch werden die Objekte mittels Sortiertechniken auf semantische Ähnlichkeit beurteilt. Als Versuchspersonen werden nur echte NutzerInnen untersucht. Die Ergebnisse werden mittels Nonmetrischer Skalierung und Hierarchischer Clusteranalyse ausgewertet. Die Interpretation wird hypothesengeleitet durchgeführt und um Erfahrungen mit dieser Methode zu sammeln mit einer Property Fitting Analyse. Das Untersuchungsdesign hat eine hohe ökologische Validität mit einer dementsprechend geringen internen Validität.

7. Darstellung von vier empirischen Studien

Im folgenden Kapitel werden die Ergebnisse von vier empirischen Studien dargestellt. Die *erste* Studie beschäftigt sich mit einer architektur-psychologischen Aufgabenstellung und untersucht die nutzerorientierte Anordnung von Geschäften für die Planung von Einkaufszentren. Die *zweite* Studie beschäftigt sich ebenfalls mit einer architektur-psychologischen Aufgabenstellung und untersucht die nutzerorientierte Anordnung von Serviceeinrichtungen in Flughafengebäuden für die beiden Nutzergruppen Vielflieger und Wenigflieger. Die *dritte* Studie beschäftigt sich mit einer psychologischen Aufgabenstellung aus der Verkaufsflächengestaltung und untersucht die KäuferInnen-orientierte Anordnung von Zimmerpflanzen für Gartencenter. Hierfür wird die begrifflich semantische Ähnlichkeit von 24 Zimmerpflanzen an 45 KundInnen und an 11 Fachkräften untersucht und verglichen. Die *vierte* Studie beschäftigt sich ebenfalls mit einer psychologischen Aufgabenstellung aus der Verkaufsflächengestaltung und untersucht die KäuferInnen-orientierte Anordnung von Süßgebäck-Artikeln für Süßgebäck-Regale. In einem Feldexperiment wird an 1047 Kundinnen untersucht, ob die ermittelte Begriffsstruktur-orientierte Anordnung tatsächlich das Such- und Findeverhalten verbessert. Abbildung 7.1 zeigt die Themen der vier Studien im Überblick.

Studie	Thema der Studie: Erhebung der Begriffsstruktur...
Einkaufszentrum	... von 31 Geschäftstypen an 60 KundInnen für die Entwicklung einer kundInnenorientierten Anordnung in Einkaufszentren
Flughafen	... von 84 Flughafen-Serviceeinrichtungen an 20 Vielfliegern und 20 Wenigfliegern für die Entwicklung einer nutzerorientierten Anordnung in Flughäfen .
Zimmerpflanze	... von 24 Zimmerpflanzen an 45 KundInnen und 11 Fachkräften für die Entwicklung einer kundInnenorientierten Anordnung in Gartencentern
Süßgebäck	... von 25 Süßgebäckartikeln an 51 KundInnen für die Entwicklung einer kundInnenorientierten Anordnung von Süßgebäck in Lebensmittelgeschäften und Überprüfung der Begriffsstruktur-orientierten Regalanordnung an 1047 KundInnen

Abbildung 7.1: Themen der Studien

7.1 Ermittlung einer nutzerorientierten Anordnung von Einzelhandelsgeschäften für Einkaufszentren

7.1.1 Einleitung

Problemstellung Wo müssen Funktionen/Handlungsangebote in einem Gebäude platziert sein, damit sie die NutzerInnen schnell und einfach finden? Gibt es eine Möglichkeit, dem Architekten oder Bauherren bereits in der Entwurfsphase empirische Erkenntnisse darüber zu liefern, wo die NutzerInnen welche Funktionen erwarten werden? Folgende Studie untersucht die erwartungsgerechte Anordnung von Geschäften in einem Einkaufszentrum. Also: Wo erwarten die KundInnen die Bäckerei, wo die Boutique und in welcher Nachbarschaft das Blumengeschäft? Wie müssen die verschiedenen Geschäfte angeordnet sein, damit die KundInnen sie später im Einkaufszentrum intuitiv finden?

Anlass der Studie Orientierungsprobleme in Gebäuden versucht man üblicherweise über Leitsysteme und Orientierungspläne zu lösen, zu einem Zeitpunkt, wenn das Gebäude schon steht. Die Anordnung von Funktionen und Nutzungen legt der Architekt jedoch bereits in der Entwurfsphase eines Gebäudes fest, sodass spätere Änderungen meist nicht mehr möglich sind oder nur durch hohe bauliche Kosten. Diese Studie will ein Vorgehen aufzeigen, wie man Architekten und Bauherren bereits in der Entwurfsphase eines Gebäudes darin unterstützen kann, dass das Gebäude später orientierungsfreundlich ist. Konkreter Anlass der Studie war eine Akquisitionsvorbereitung.¹ Erfahrungen mit Gebäuden und architektonischen Anordnungsproblemen lagen bisher nicht vor.

Ziele der Studie Die Untersuchung verfolgt zwei Ziele:

1) Die Untersuchung sollte klären, ob KundInnen eine personenübergreifende Begriffsstruktur für Einzelhandelsgeschäfte haben und wenn ja, wie diese aussieht: ob sie handlungsbezogenen oder anderen Kriterien folgt. In der Literatur der angewandten Psychologie fanden sich keine vergleichbaren Studien. Es gab sehr wohl Untersuchungen, die die

¹. Die empirische Untersuchung wurde von Christiane Berker im Rahmen einer Diplomarbeit durchgeführt.

Mentale Landkarte bezüglich bestehender Gebäudestrukturen abtesteten und somit individuelles, semantisches Wissen erhoben. Hier wurden Personen z.B. gebeten, die Grundrisse ihnen bekannter Gebäude nachzuzeichnen. Aber es fanden sich keine Studien, die explizit das Ziel hatten, begrifflich-semantische Wissensstrukturen über Einzelhandelsgeschäfte oder andere räumliche Handlungsangebote *vor* der Erstellung eines Gebäudes zu ermitteln, um damit eine „nutzerorientierte“ Grundrissplanung dieser Handlungsangebote zu entwickeln, die nach der Erstellung des Gebäudes die Erwartungen der NutzerInnen treffen soll. Die Begriffsstruktur wurde entsprechend den Ausführungen in Kapitel 5 als semantische Ähnlichkeit operationalisiert.

2) Mit dieser Untersuchung sollten überdies Erfahrungen mit schnell durchführbaren und kostengünstigen Erhebungsmethoden gewonnen werden. Zum einen rentieren sich hohe Kosten für empirische Studien bei manchen Produkten nicht, zum anderen besteht ein Bedarf nach kostengünstigen Methoden, um Auftraggeber erst einmal mit Methoden und empirischen Problemlösungswegen der angewandten Psychologie vertraut zu machen.

In dieser Studie wurden die folgenden *Hypothesen* untersucht:

Hypothese 1: Personenübergreifende, stabile Ähnlichkeitsbeurteilung KundInnen haben eine personenübergreifende Ähnlichkeitsbeurteilung bzgl. Einzelhandelsgeschäften.

Hypothese 2: Handlungsbezogene Kriterien Der Ähnlichkeitsbeurteilung von Einzelhandelsgeschäften liegen handlungsbezogene Kriterien der KundInnen zugrunde; d.h. es werden solche Einzelhandelsgeschäfte als semantisch ähnlich beurteilt, deren Produkte die KundInnen in einem gemeinsamen Handlungskontext verwenden (z.B. zum Essen und Kochen, zum Anziehen, zum Schenken, etc.). Die Ähnlichkeitsbeurteilung folgt nicht willkürlichen oder anderen Kriterien (z.B. Essbares-Nichtessbares, Frischware-Trockenware, Materialbeschaffenheit, Größe, Preis etc.).

Untersuchungsdesign Die Studie besteht aus 1 Vorversuch und 1 Hauptversuch: Der Vorversuch hatte das Ziel, ca. 30 Einzelhandelsgeschäfte auszuwählen, die typisch für ein Einkaufszentrum sind. Im Hauptversuch wurde dann für diese 30 Einzelhandelsgeschäfte eine Ähnlichkeitsbeurteilung durchgeführt. Abbildung 7.1-1 gibt Überblick.

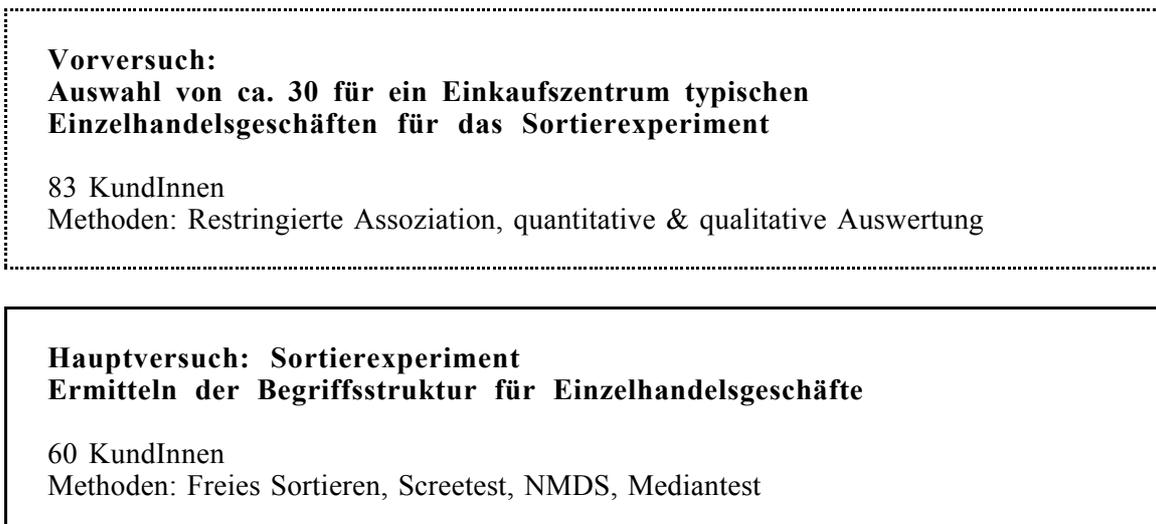


Abbildung 7.1-1. Untersuchungsdesign zur Ermittlung der Begriffsstruktur von Geschäften

7.1.2 Vorversuch zur Auswahl typischer Einzelhandelsgeschäfte

Einleitung

Ziel des Vorversuches war es, ca. 30 typische Geschäfte für die Hauptuntersuchung auszuwählen. Typische Geschäfte wurden als solche definiert, die die Personen beim Assoziieren am häufigsten nennen.

Methode

Erhebungsmethoden 83 Personen wurden mittels der *Methode der restringierten Assoziation* untersucht. Assoziative Verfahren eignen sich nur bedingt für die Erhebung der semantischen Ähnlichkeit von Begriffen im Hauptversuch (vgl. Kapitel 6), aber es ist ein übliches Verfahren, um die Auswahl der Begriffe zu treffen, die dann bezüglich ihrer Begriffsstruktur, d.h. ihrer begrifflichen Ähnlichkeit untersucht werden (vgl. Strube 1984). Bei dieser Methode soll die Versuchsperson auf das Schlüsselwort so viele Begriffe nennen, wie ihr einfallen. Es werden dann diejenigen Assoziationen gewertet, die derselben Kategorie wie das Schlüsselwort angehören. Die restringierte Assoziation ist eine Abänderung der freien Assoziation. Bei der freien Assoziation wird jede Antwort gewertet. Auf das Schlüsselwort „Einkaufszentrum“ wäre die Antwort „Supermarkt“

genauso gültig wie „Hintergrundmusik“ oder „Kindergeschrei“. Bei der restringierten Assoziation sind die Antworten dagegen eingeschränkt und auf das Schlüsselwort „Einkaufszentrum“ werden nur Antworten gewertet, die der gleichen Kategorie entsprechen, z.B. „Schuhgeschäft“, „Fotoladen“ oder „Zoohandlung“.

Versuchspersonen, Durchführung und Dokumentation Es wurden 83 KundInnen untersucht: 7 Männer, 76 Frauen. Der hohe Frauenanteil ist gerechtfertigt, da mehr Frauen KundInnen von Einkaufszentren sind als Männer. Tabelle 7.1-1 zeigt die Altersverteilung. Die Stichproben wurden an drei Orten in München zufällig gezogen: am Hauptbahnhof (17 Personen), im Untergeschoss des U-Bahnhofs Münchner Freiheit (19 Personen) und am Rotkreuzplatz (47 Personen). Die Befragungen wurden von drei Versuchsleitern durchgeführt. Pro Person dauerte die Befragung ca. fünf Minuten. Alle genannten Assoziationen wurden auf einem Protokollbogen dokumentiert. Die Instruktion lautete: *„Bitte stellen Sie sich ein Einkaufszentrum vor. Welche Einzelhandelsgeschäfte sollten Ihrer Meinung nach darin vorkommen? Zählen Sie bitte alle auf, die Ihnen einfallen.“*

Tabelle 7.1-1. Versuchspersonen des Vorversuchs „Einzelhandelsgeschäfte“

Alter	Frauen	Männer	Gesamt
< 20	4	–	4
20 – 29	22	1	23
30 – 39	19	3	22
40 – 49	7	2	9
50 – 59	13	1	14
> 59	11	–	11
Summe	76	7	83

Ergebnisse des Vorversuchs

Insgesamt nannten die 83 Personen 720 Begriffe, davon 109 unterschiedliche. Die durchschnittliche Anzahl der Nennungen betrug pro Person 8,7 Begriffe, wobei die geringste Zahl 2 Nennungen und die höchste 24 war. Die Hälfte derer, die 13 oder mehr Nennungen nannten, waren jünger als 30 Jahre. 64 Begriffe wurden nur ein oder zweimal ge-

nannt. Die 109 Nennungen wurden wie folgt auf 31 Begriffe reduziert: 40 Begriffe wurden gestrichen, weil es sich nicht um Einzelhandelsgeschäfte, sondern um Dienstleistungseinrichtungen handelte (z.B. Infostand, Kindergarten, Café, Friseur etc.). Inhaltlich gleiche, aber sprachlich unterschiedliche Nennungen wurden zu einem Wort zusammengefasst (z.B. Boutique, Strumpfgeschäft etc. wurden zu „Damenbekleidung“). Einige Geschäfte wurden aus ethischen Gründen gestrichen (z.B. Erotikgeschäft, Waffenhandlung). Einige Geschäfte wurden gestrichen, weil sie nicht typisch für ein Einkaufszentrum sind (z.B. Möbelhaus, Getränkemarkt) und z.B. eine Autovorfahrt benötigen. Tabelle 7.1-2 zeigt 31 Einzelhandelsgeschäfte, die in die Hauptuntersuchung aufgenommen wurden. Die Geschäfte sind in der Tabelle thematisch geordnet, um dem Leser den Überblick zu erleichtern. Die Versuchspersonen erhielten eine zufällige Reihenfolge.

Tabelle 7.1-2. Im Vorversuch ermittelte 31 Einzelhandelsgeschäfte (thematisch geordnet)

Ausgewählte 31 Einzelhandelsgeschäfte für den Hauptversuch			
Apotheke	Fisch/Wild/Geflügel	Sportartikel	Haushaltswaren
Drogerie	Weinhandlung	Schuhgeschäft	Buchhandlung
Blumenladen	Teegeschäft	Kinderkleidung	Geschenkladen
Obst / Gemüse	Schreibwaren	Damenbekleidung	Juwelier/Schmuck
Bäckerei	Handarbeiten	Herrenbekleidung	Fotogeschäft
Metzgerei	Zoohandlung	Lederwaren	Optiker
Supermarkt	Zeitschriften	Spielwaren	Schallplatten
Naturkost	Parfümerie	Heimwerken	

7.1.3 Hauptversuch

Methode

Erhebungsmethode Die semantische Ähnlichkeit wurde über die *Methode des freien Sortierens* an 60 Personen erhoben. Bei dieser Methode wurde jede Versuchsperson aufgefordert, 31 Kärtchen, auf die die Einzelhandelsgeschäfte des Vorversuches geschrieben waren, in eine Reihe zu legen, wie sie es für stimmig hält. Es wurde kein Sortierkriterium vorgegeben. Die Kärtchen hatten eine zufällige Reihenfolge (Abb. 7.1-2).

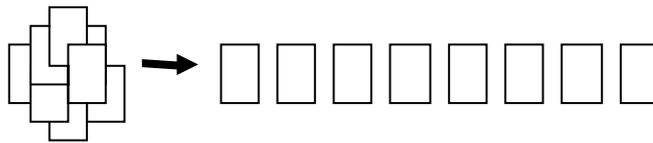


Abbildung 7.1-2. Veranschaulichung des Freien Sortierens: Die Versuchsperson soll die ungeordneten Karten mit Begriffen in die für sie „richtige“ Nachbarschaftsreihenfolge legen.

Die Methode des freien Sortierens basiert auf Ähnlichkeitsurteilen und ist eine modifizierte Methode der Nachbarschaftsuntersuchung. (Bei der Nachbarschaftsuntersuchung soll die Versuchsperson zu jedem genannten Item zwei „Nachbarn“ nennen. Diese Methode ist jedoch aufwändiger als das Freie Sortieren.) Auswahlkriterium für das Freie Sortieren war der geringe Aufwand in der Durchführung und in der Auswertung. Aus theoretischer Sicht hat die Methode den Nachteil, dass sie keine Mehrfachzuordnungen erlaubt. Dies ist bei der Interpretation der Ergebnisse zu berücksichtigen.

Untersuchungsmaterial Die im Vorversuch ermittelten 31 Geschäfte wurden in schwarzer Schrift auf weiße Kärtchen geschrieben. Da für die Auswahl die in der restringierten Assoziation häufig genannten Geschäfte gewählt wurden, darf davon ausgegangen werden, dass auch die Versuchspersonen im Hauptversuch Gedächtnisrepräsentationen über diese Geschäfte haben. Die Kärtchen wurden vor jedem Versuch gemischt.

Versuchspersonen, Durchführung und Dokumentation Es wurden insgesamt 60 Personen, davon 17 Männer, 43 Frauen untersucht. Der hohe Frauenanteil ist gerechtfertigt, da mehr Frauen KundInnen von Einkaufszentren sind als Männer. Tabelle 7.1-3 zeigt die Altersverteilung der Versuchspersonen. Die Stichproben wurden in drei Münchner Cafés zufällig gezogen. Die Untersuchung wurde von 3 Versuchsleitern durchgeführt. Pro Person dauerte die Befragung ca. 15 Minuten. Die Instruktion lautete: „Hier habe ich 31 Karten. Auf jeder Karte steht ein Einzelhandelsgeschäft. Suchen Sie sich eine Karte aus, mit der Sie beginnen möchten. An diese Karte legen Sie das Geschäft, das Ihrer Meinung nach am besten daneben passt. Daneben dann in gleicher Weise das Nächste, bis alle Kärtchen in einer Linie liegen. Es ist so ähnlich wie bei einem Dominospiel. Sie können nichts richtig oder falsch machen, legen Sie die Karten so hin, wie es für Sie stimmt.“ Die Versuchsleiterin dokumentierte die gelegten Kärtchen pro Versuchsperson auf einem Protokollbogen.

Tabelle 7.1-3. Versuchspersonen für den Hauptversuch „Einkaufszentrum“

Alter in Jahren	Frauen	Männer	Gesamt
20 – 29	16	6	22
30 – 39	14	4	18
40 – 49	8	5	13
50 – 59	5	1	6
> 59	–	1	1
Summe	43	17	60

Ergebnisse des Hauptversuches

Auswertungsmethoden Für die Sortierdaten wurde zunächst eine Distanzmatrix erstellt. *Erstens* wurden die Daten mit der Nonmetrischen Multidimensionalen Skalierung (NMDS) ausgewertet, um die Ähnlichkeitsurteile in eine räumliche Darstellung abzubilden (vgl. Gigerenzer 1981, Borg 1981). Hierbei wurde das Programm MINISSA von Lingoes & Roskam (1975) verwendet. *Zweitens* wurde mittels Sreetest der Stresswert ermittelt. Ein signifikanter Stresswert besagt, dass die räumliche Konfiguration die Ähnlichkeitsurteile der Versuchspersonen gut wiedergibt, dass die Ähnlichkeitsbeurteilungen eine personenübergreifende stabile Struktur ergeben, und dass es überzufällig konsistente Kriterien für die begrifflich-semantische Ähnlichkeitsbeurteilung gibt. Als Signifikanzniveau wurde das 1% Niveau festgelegt. Dieses Signifikanzniveau ist angesichts der Stichprobe (60 Vp) als hoch zu bewerten um sicherzugehen, dass die Nullhypothese irrtümlich abgelehnt wird. Ein irrtümliches Beibehalten der Nullhypothese (β -Fehler), d.h. dass es keine Übereinstimmung der KundInnen in den Ähnlichkeitsbeurteilungen für Einzelhandelsgeschäfte gibt, würde den Betreibern mehr Nachteile (bzgl. Kundenzufriedenheit, Orientierung, eventuellem Abverkauf, Umsatz) bringen als ein irrtümliches Ablehnen. Es wird angenommen, dass eine Anordnung, die den Überlegungen der Architekten entspricht, nicht kundenorientierten Kriterien folgt. *Drittens* wurde geprüft, ob die Einzelhandelsgeschäfte beim Sortieren in einer systematischen Reihenfolge gelegt wurden, oder ob die Einzelhandelsgeschäfte gleich häufig jede Position einnahmen. Hierfür wurden die Medianwerte errechnet und mittels des Mediantests geprüft.

Ergebnisse des Scree-tests Der Scree-test liefert für die Stichprobe (60 Vpn) für die 2-dimensionalen NMDS-Konfigurationen einen Stresswert von 0.1975. Er ist kleiner als der kritische Stresswert von $\sigma = 0.327$ für 31 Objekte auf dem 1% Signifikanzniveau (vgl. Gigerenzer 1981, S. 353).

Ergebnisse der Zufallsverteilung beim Freien Sortieren Mittels des Mediantests (vgl. Siegel, 1976) wurde geprüft, ob die Einzelhandelsgeschäfte gleich wichtig waren beim Sortieren, ob die Versuchspersonen strategisch vorgehen und mit bestimmten Geschäften das Sortieren anfangen, oder ob bestimmte Geschäfte in Kombination gelegt wurden etc. Hierfür wurden die Medianwerte berechnet. Sie geben Auskunft darüber, vor der wievielten Stelle 50% der Versuchspersonen ein Geschäft legen. Dann wurde geprüft, ob eine Zufallsverteilung vorliegt. Bei 30 Freiheitsgraden beträgt der kritische Wert für das 5% Niveau 43,77. (Bortz 1985 S. 837). Der empirisch beobachtete Wert beträgt 9,9. Das Ergebnis weist darauf hin, dass eine Zufallsverteilung vorliegt, d.h. dass alle Geschäfte gleich wichtig waren.

Darstellung der NMDS-Konfiguration Abbildung 7.1-3 zeigt die 2-dimensionale NMDS-Konfiguration für die Gesamtstichgruppe. Geschäfte, die als begrifflich ähnlich beurteilt wurden, sind hier räumlich nahe abgebildet. Die NMDS-Konfiguration zeigt folgende begrifflich-semantiche Anordnung: Im linken Bereich der Abbildung sind Lebensmittelgeschäfte wie Fisch, Wild, Wein, Metzgerei, Tee, Bäcker, Gemüse, Supermarkt, Naturkost. An diese Lebensmittelgeschäfte schließen oben Haushaltgeschäfte an mit Haushaltswaren, Heimwerken, Handarbeiten und Spielwaren. Die Spielwaren leiten begrifflich-semantiche über zu den Bekleidungsgeschäften: Kinderbekleidung, Herrenbekleidung, Sportbekleidung, Damenbekleidung, Schuhe, Lederartikel. In der begrifflichen Nähe Lederartikel wird das Schmuckgeschäft beurteilt und daran anschließend die Geschäfte: Foto, Optiker, Schallplatten. In der semantiche-begrifflichen Nähe folgen dann Parfümerie, Apotheke, und Drogerie sowie Buch, Schreibwaren, Tabak. Im begrifflich semantiche Zentrum liegen Geschäfte mit Geschenken und Blumen.

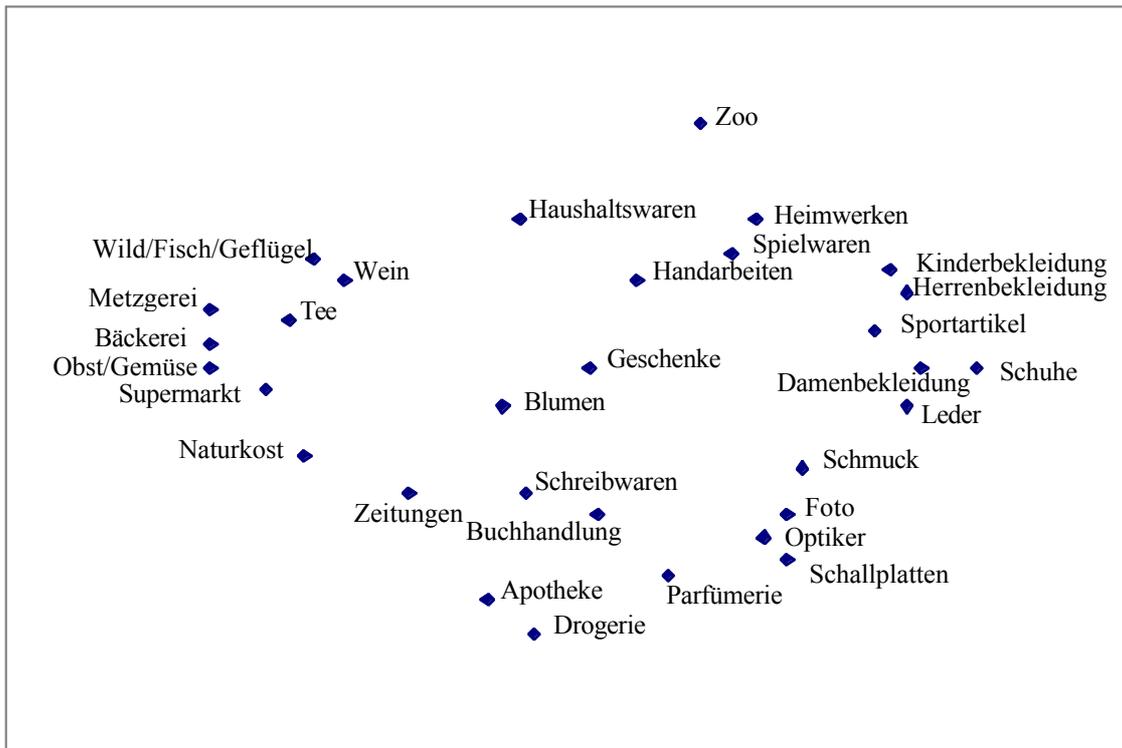


Abbildung 7.1-3. 2-dimensionale NMDS Konfiguration für 31 Einzelhandelsgeschäfte

7.1.4 Interpretation der Ergebnisse

Interpretation in Hinblick auf die Hypothese 1: Personenübergreifende Ähnlichkeitsbeurteilung Die Ergebnisse weisen darauf hin, dass es eine ausgeprägte, personenübergreifend ähnliche Begriffsstruktur für Einzelhandelsgeschäfte gibt: Die Ergebnisse des Scree-test zeigen für die 2-dimensionalen NMDS-Konfigurationen ($= 0.1975$) signifikant niedrigere Stresswerte als der kritische Wert ($= 0.327$). Der Stresswert ist ein Maß für die Anpassungsgüte zwischen den Ausgangsdaten und der berechneten NMDS-Konfiguration. Die Ergebnisse bestätigen die Hypothese 1, dass die Personen die Einzelhandelsgeschäfte nicht nach individuell unterschiedlichen Kriterien sondern vielmehr – überzufällig – nach gemeinsamen Kriterien beurteilen. Die Hypothese wird daher beibehalten. Da die Versuchspersonen an unterschiedlichen Orten in München rekrutiert wurden, ist davon auszugehen, dass die Personen unterschiedliche Einkaufszentren und Geschäftsstraßen mit unterschiedlichen Geschäftsanordnungen kennen. Die hohe Übereinstimmung der Ergebnisse lässt daher die Schlussfolgerung zu, dass die Ähnlichkeitsbeur-

teilungen nicht die Mentale Landkarte eines konkreten Einkaufszentrums oder einer konkreten Einkaufsstraße repräsentieren.

Interpretation in Hinblick auf die Hypothese 2: Handlungsbezogene Kriterien Untersucht man die Ähnlichkeitsbeurteilung in Hinblick auf die zugrunde liegenden Kriterien, so bestätigt sich die Hypothese „Handlungsbezogene Kriterien“. Den Beurteilungen liegen definierte Handlungsfelder zugrunde, wie „Essen & Kochen“, „Sich Bekleiden“, „Schreiben und Lesen“, „Den Körper pflegen und gesunden“, „Schenken“. Tabelle 7.1-4 zeigt die Geschäfte, die in der NMDS nahe beieinander liegen und nennt die zugrunde liegenden Handlungsfelder.

Tabelle 7.1-4. Ergebnisse und Interpretation der NMDS „Einkaufszentrum“

Ergebnisse: Begriffliche Nähe der Geschäfte entsprechend der NMDS-Konfiguration	Interpretation: mögliche zugrunde liegende Handlungsfelder
Reformhaus/Naturkost, Supermarkt, Obst und Gemüsegeschäft, Bäckerei, Metzgerei, Teegeschäft, Weingeschäft, Fisch-Wild-Geflügelgeschäft	essen und kochen
Haushaltswarengeschäft	den Haushalt machen
Zoohandlung	sich ums Tier kümmern
Spielwarengeschäft, Heimwerkergeschäft, Handarbeitsgeschäft	spielen und werken
Kinderbekleidungsgeschäft, Sportgeschäft, Herren-, Damenbekleidungsgeschäft, Lederwarengeschäft, Schuhgeschäft	sich bekleiden
Schmuckgeschäft	sich schmücken
Fotogeschäft, Schallplattengeschäft, Optiker	sehen und hören
Parfümerie, Drogerie, Apotheke	den Körper pflegen und gesunden
Zeitschriftengeschäft, Buchhandlung, Schreibwarengeschäft	schreiben und lesen
Blumengeschäft, Geschenkartikelgeschäft	sich und andere beschenken

Die Ähnlichkeitsbeurteilungen geben Handlungsfelder wieder, wie die Personen die Gegenstände, die sie in den Geschäften kaufen, im Alltag verwenden. Es wurden entsprechend der Hypothese nicht alle Geschäfte als ähnlich beurteilt, die z.B. etwas Essbares verkaufen. So ist die Apotheke – obwohl sie ja Essbares verkauft – nicht den Nahrungsmittelgeschäften und dem Handlungsfeld „Essen & Kochen“ zugeordnet, sondern eindeutig der Drogerie und Parfümerie und somit dem Bereich „Den Körper pflegen und gesunden“. Leder- und Schuhgeschäft wurden eindeutig dem Damenober-, Herrenober- und Kinderbekleidungsgeschäft zugeordnet und nicht separat. Auch das Material oder die Frische sind keine bestimmenden Kriterien: Das Handarbeitsgeschäft wurde eher dem Bereich Haushaltswaren als dem Bereich Bekleidung zugeordnet, obwohl es dort auch Wolle und Nähmaterialien für Bekleidungszwecke gibt. Das Blumengeschäft wurde – obwohl es ja auch ein Frischeartikel aus dem Garten oder vom Feld anbietet – nicht dem Obst- und Gemüsegeschäft, sondern dem Geschenkartikelgeschäft zugeordnet zum Bereich „Sich und andere beschenken“. Abbildung 7.1-4 zeigt die semantischen Ähnlichkeitsbeziehungen und Handlungsfelder noch deutlicher.

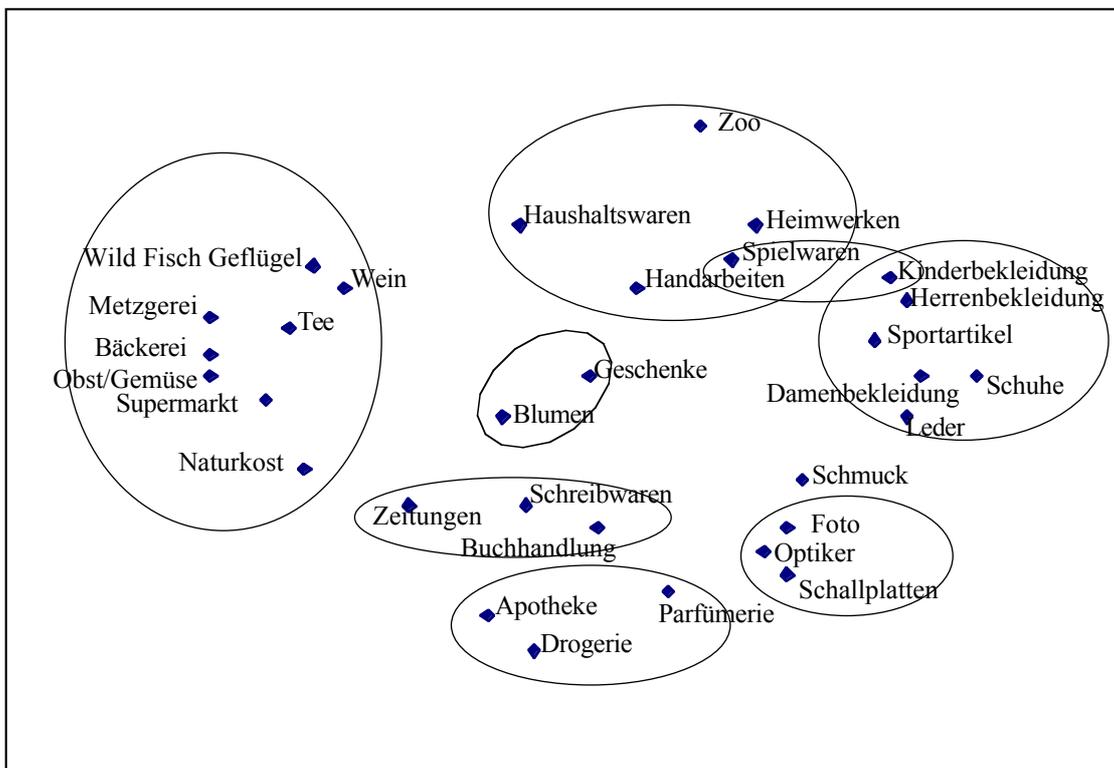


Abbildung 7.1-4. Zweidimensionale NMDS Konfiguration mit 31 Einzelhandelsgeschäften und eingezeichneten Handlungsfeldern

Es wird auch deutlich, dass Geschäfte nicht nur einem Handlungsfeld zuzuordnen sind, sondern im Sinne von Kreuzklassifikationen mehreren Gruppen angehören können oder Brücken bilden. Betrachtet man z.B. das Spielwarengeschäft, so liegt es in der NMDS-Konfiguration räumlich nahe der Kinderbekleidung, dem Heimwerken und dem Handarbeiten. Bedenkt man, dass Heimwerker- und Handarbeitsgeschäft auch Sortimente für Kinder haben, lässt sich so auch eine neue Gruppe „Geschäfte fürs Kind“ definieren.

Die Analyse macht deutlich, dass die Versuchspersonen die Einzelhandelsgeschäfte nach handlungsbezogenen Kriterien beurteilen. Die Ergebnisse lassen sogar den Schluss zu, dass die Versuchspersonen die Einzelhandelsgeschäfte hinsichtlich des Handlungsziels und des Verwendungszweck für die gekauften Artikel beurteilen. Geschäfte, deren Artikel dem gleichen Handlungsziel dienen oder im gleichen Handlungskontext erlebt werden, – wie Geschenkartikel und Blumen – werden als ähnlich beurteilt. Beide Geschäfte oder aber nur eins können aufgesucht werden, um ein Geschenk einzukaufen. Dies verweist einmal auf eine Gleichwertigkeit bzw. Austauschbarkeit der Geschäfte hinsichtlich des Handlungsziels „Schenken“, aber auch darauf, dass sie in einem gemeinsamen Handlungsablauf bzw. -kontext erlebt werden können. Das Handlungsziel, was man mit dem gekauften Artikel macht, scheint eine Ordnung stiftende Funktion für die Begriffsrepräsentation zu haben. Diese Annahme geht noch über das hinaus, was Klix et al. in ihrer Theorie formulieren.

7.1.5 Beurteilung der verwendeten Methoden

Die Methode der restringierten Assoziation im Vorversuch zeigte sich als aufwändig und subjektiv belastet in der Auswertung: so musste z.B. der Versuchsleiter entscheiden, inwieweit Nennungen zusammengefasst werden können. Auch blieb letztlich unklar, ob die ausgewählten 31 Einzelhandelsgeschäfte repräsentativ für die Geschäfte von Einkaufszentren sind. Es ist ein in der Literatur bekanntes Verfahren für die Auswahl von Elementen für einen Sortiersuch. Doch es bleibt zu bedenken, ob es geeignetere gibt, die auch die Repräsentativität gewährleisten.

Die Methode des Freien Sortierens im Hauptversuch zeigte sich als einfach und schnell in der Durchführung. Obwohl die Versuchspersonen die Kärtchen eindeutig zuzuordnen mussten und begriffliche Kreuzklassifikationen nicht möglich waren, zeigen sich in der NMDS-Konfiguration Kreuzklassifikationen. Dennoch wird daran erinnert, dass die Methode des Hierarchischen Sortierens in der Literatur differenziertere Ergebnisse liefert. Daher gilt es abzuwägen, ob die Methode des hierarchischen Sortierens nicht vorzuziehen ist.

7.1.6 Zusammenfassung

KundInnen haben eine ausgeprägte, überindividuelle Ähnlichkeitsbeurteilung für Einzelhandelsgeschäfte, die dem Verwendungszweck folgt. Dies zeigen die Ergebnisse des Sortierexperimentes, bei dem 60 Personen (17 Männer, 43 Frauen) mittels der Methode des freien Sortierens 31 Einzelhandelsgeschäfte auf Ähnlichkeit beurteilten. Die Sortierergebnisse wurden mit dem Screenshot und der NMDS ausgewertet. Die 2-dimensionale NMDS-Konfiguration hat mit einem signifikanten Stresswert (1%) eine stabile, überzufällige Struktur und zeigt deutlich begrifflich-semantiche Nachbarschaften. Der Ähnlichkeitsbeurteilung liegen dabei handlungsbezogene Kriterien zugrunde. Zum Beispiel die Gruppierung Obst/Gemüse-Geschäft, Bäckerei, Metzgerei, Wild/Fisch/Geflügel-Geschäft, Naturkostgeschäft gibt das Handlungsfeld „Essen & Kochen“ wieder. Die Apotheke, die zwar auch Essbares verkauft, ist der Drogerie und Parfümerie zugeordnet und verweist auf das Handlungsfeld „den Körper pflegen und gesunden“. Oder die Gruppierung von Zeitungsgeschäft, Buchhandlung und Schreibwarengeschäft gibt das Handlungsfeld „Schreiben und Lesen“ wieder. Die Gruppierung Blumen- und Geschenkegeschäft gibt das Handlungsfeld „Sich und andere beschenken“ wieder, obwohl Blumen ja auch Frischeartikel wie Obst und Gemüse sind. Die handlungsbezogenen Kriterien zeigen somit das Handlungsziel bzw. den Verwendungszweck an. Diese empirisch erhobene Begriffsstruktur lässt sich übersetzen in eine erwartungsgerechte Anordnung von Einzelhandelsgeschäften, die als Vorgabe für die Grundrissplanung eines Einkaufszentrums dienen kann.

7.2 Ermittlung einer nutzerorientierten Anordnung von Serviceeinrichtungen für Flughäfen

7.2.1 Einleitung

Problemstellung Die folgende Untersuchung beschäftigt sich mit dem architekturpsychologischen Problem: Wie müssen die Einrichtungen in einem Flughafen angeordnet sein, damit sich die Flughafengäste schnell und einfach zurechtfinden? Wo erwarten sie den Zeitungskiosk und den Meeting-Point? Wo die Bank und den Ticket-Schalter?

Anlass der Studie Diese Studie sollte empirische Daten liefern über eine nutzerorientierte Anordnung von Serviceangeboten für ein Flughafen-Terminal.¹ Sie sollte klären, ob sich für Flughafengäste Gemeinsamkeiten in der Ähnlichkeitsbeurteilung für Flughafeneinrichtungen finden lassen. Existieren überindividuelle Ähnlichkeitsbeurteilungen, dann kann auf deren Basis eine Vorgabe für Architekten entwickelt werden. Diese Vorgabe würde das Architektenteam bereits im Briefing (bzw. in der Ausschreibung) erhalten, damit sie sie bereits in den Gestaltungsentwürfen berücksichtigen können. Dieses Vorgehen ist zur Zeit nicht üblich. Zur Zeit wird erst nach der Erstellung des Gebäudes mit Hilfe von Übersichtsplänen und Leitsystemen versucht, die Flughafengäste in ihrer Orientierung zu unterstützen, was aber Probleme mit sich bringt (s. Kapitel 3).

Ziele der Studie Die Studie verfolgte zwei Ziele: 1) Es sollte die Begriffsstruktur von Flughafen-NutzerInnen bezüglich Flughafeneinrichtungen abklärt werden. Es wurde angenommen, dass sie aufgrund ähnlicher Handlungsabläufe und -erfahrungen eine hohe Übereinstimmung in der Begriffsstruktur haben. 2) Die Studie sollte abklären, ob die Begriffsstrukturen von Vielfliegern und Wenigfliegern Unterschiede aufweisen. Es wird vermutet, dass sich die unterschiedlichen Handlungsziele und -erfahrungen auch in den Begriffsstrukturen äußern. Da Vielflieger häufiger Linienflüge und Inlandsflüge für Geschäftszwecke tätigen, müssten sie Inlands-Fluglinien einen anderen Stellenwert als z.B. Chartergesellschaften geben. Außerdem benötigen sie nach der Ankunft einen reibungslosen und schnellen Transport in die Stadt, um ihre Termine wahrzunehmen, was

¹ Die empirische Untersuchung wurde von Bettina Wecker im Rahmen einer Diplomarbeit durchgeführt.

sich in einer anderen semantischen Beurteilung der Transport- und Informationsbereiche ausdrücken müsste. Und sie reisen nicht mit Kindern. Es wird vermutet, dass sie den entsprechenden Serviceeinrichtungen eine andere semantische Bedeutung geben als Wenigflieger, die, wenn sie fliegen, dann meist mit Charterlinien in den Urlaub fliegen und, soweit vorhanden, mit Kindern fliegen.

In dieser Studie wurden die folgenden Hypothesen untersucht:

Hypothese 1: „Personenübergreifende, stabile Begriffsstruktur“ Es lässt sich für Flughafengäste eine personenübergreifende Ähnlichkeitsbeurteilung von Flughafeneinrichtungen finden. Die Ähnlichkeitsbeurteilung von Flughafeneinrichtungen in einem Sortierexperiment erfolgt nach überindividuellen Kriterien.

Hypothese 2: „Bisherige Handlungserfahrungen“ Die Ähnlichkeitsbeurteilungen von Vielfliegern für Flughafeneinrichtungen unterscheiden sich von denen der Wenigflieger. Die Unterschiede entsprechen dabei den unterschiedlichen Handlungszielen und -abläufen (berufsbezogene versus urlaubsbezogene Handlungsziele und -abläufe).

Hypothese 3: „Handlungsbezogene Kriterien“ Den Ähnlichkeitsbeurteilungen von Serviceeinrichtungen am Flughafen liegen handlungsbezogene (z.B. Abfliegen versus Ankommen) Kriterien zugrunde und nicht kategoriale (z.B. Flugbezogene Schalter versus Konsumbezogene Einrichtungen; oder Fluggesellschaften, d.h. alle Einrichtungen der Fluglinie A versus alle Einrichtungen der Fluglinie B).

Historische Bedingtheiten Die Untersuchung wurde 1994 durchgeführt, zu einer Zeit, als die Lufthansa noch uneingeschränkte Herrin im inländischen Personenflugverkehr war. Duty-Free-Einkäufe innerhalb der EU waren möglich. Das Kerosin war teurer und die Zahl der Fernreisen gering. Die First-Class-Einrichtungen zur Markendifferenzierung hatten eine andere Bedeutung, Kooperationen wie die Staralliance gab es noch nicht. Die Pass- und Zollkontrollen waren noch nicht nach Schengen- und Nicht-Schengenländern organisiert. Die Ergebnisse sind eine historische Momentaufnahme. Heute würden wir mehr Serviceeinrichtungen und auch mehr Nutzergruppen untersuchen, um die z. Zt. aktuellen Fragen zu beantworten.

Untersuchungsdesign Die Untersuchung besteht aus einem Sortierexperiment (Abb. 7.2-1). Die Auswahl der zu untersuchenden Flughafeneinrichtungen wurde in das Sortierexperiment integriert und nicht in einem Vorversuch ermittelt.

Hauptversuch:	
Ermitteln der Begriffsstruktur Flughafeneinrichtungen	
40 Personen:	20 Vielflieger – 20 Wenigflieger
Methode:	Free Card Sorting
Auswertung:	Häufigkeiten, MDS, Scree-test, HCA

Abbildung 7.2-1. Untersuchungsdesign zum Ermitteln der Begriffsstruktur von Flughafeneinrichtungen

7.2.2 Hauptversuch

Erhebungsmethode Als Methode wurde das *Free Card Sorting* gewählt, bei der die Versuchspersonen die auf Kärtchen geschriebenen Flughafen-Einrichtungen in Gruppen ordnen, die ihnen vernünftig erscheinen (Abb. 7.2-2). Es werden keine Kriterien vorgegeben. Die Methode des Free-Card-Sorting erlaubt einen Ähnlichkeitsvergleich mit sehr vielen Objekten. (Zur ausführlichen Begründung siehe Kapitel 6: Entwicklung des Untersuchungsdesigns.)

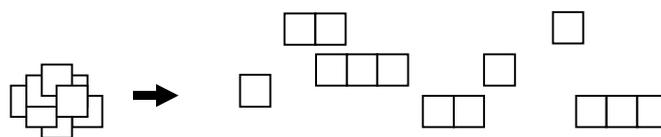


Abbildung 7.2-2. Veranschaulichung des Free Card Sorting. Die Vp soll die ungeordneten Karten mit Begriffen in für sie vernünftige Gruppen sortieren.

Untersuchungsmaterial Es wurden 84 verschiedene Flughafeneinrichtungen auf weiße Kärtchen geschrieben (Arial 16 pt), die vor jedem Versuch gemischt wurden. Tabelle 7.2-1 zeigt die untersuchten Flughafeneinrichtungen.

Tabelle 7.2-1. Liste der 84 untersuchten Flughafen-Einrichtungen thematisch geordnet

Ticket LH *	Ticket BA**	Zoll	Café
Check-In LH	Check-In BA	Passkontrolle	Restaurant
Gate LH	Gate BA	Duty-Free	Snack
Warteraum LH	Warteraum BA	Post	Bar
Ankunft LH	Ankunft BA	Bank	Lounge
Gepäckausg. LH	Gepäckausg. BA		
Ticket TWA	Ticket Malaysia	Zeitungen	Sozialdienst
Check-In TWA	Check-In Malaysia	Buchhandlung	Kapelle
Gate TWA	Gate Malaysia	Reisebedarf	Seelsorge
Warteraum TWA	Warteraum Malay.	Tabak	Fundbüro
Ankunft TWA	Ankunft Malaysia	Foto	Polizei
Gepäckausg. TWA	Gepäckausg. Malay.	Souvenir	Medizin. Zentrum
Ticket AA***	Ticket Air Conti	Delikatessen	Gepäckaufbew.
Check-In AA	Check-In Air Conti	Lotto-Toto	Parkleitzentrale
Gate AA	Gate Air Conti	Friseur	Transferdienst
Warteraum AA	Warteraum Air Conti	Modeboutique	Stadtinformation
Ankunft AA	Ankunft Air Conti	Elektroartikel	Hotelinformation
Gepäckausg. AA	Gepäckausg. Air C.	Friseur	Meeting Point
Check-In LTU	Karstadt-Reisen	Blumen	Toilette
Gate LTU	Kreutzer-Reisen	Apotheke	Babywickelraum
Warteraum LTU	TUI-Reisen	Infostand	Dusche
Ankunft LTU	LastMinute Huckepack	Telefonzellen	
Gepäckausg. LTU			

* LH = Lufthansa, ** BA = British Airways, *** AA = American Airways

Versuchspersonen Es wurden 40 Personen untersucht: je 20 Männer und Frauen, davon je 20 Viel- und Wenigflieger (Tab. 7.2-2). Als Vielflieger wurden Personen definiert, die im Jahr mindestens 6 mal fliegen. Sie arbeiten in unterschiedlichen Branchen. Die Vielflieger wurden nicht nach First-Class und Business-Fliegern unterschieden.

Durchführung, Instruktion und Dokumentation Die Sortierexperimente fanden bei den Versuchspersonen zu Hause oder im Büro statt. Die Instruktion lautete: „*Ich habe hier Kärtchen, auf denen verschiedene Einrichtungen, die man am Flughafen hat, stehen. Sie sollen nun die Kärtchen in verschiedene Gruppen sortieren. Dabei gehören in eine Gruppe immer die Kärtchen, die Ihrer Meinung nach zusammenpassen, d.h. die Sie an einem Flughafen gerne in räumlicher Nähe hätten. Überlegen Sie bitte, was Sie wann, wo brauchen könnten, was Sie an einem Flughafen machen. Wichtig ist, dass Sie die*

Kärtchen nicht so sortieren, wie Sie das schon von einem Flughafen kennen, sondern so, wie es Ihnen sinnvoll erscheint. Wenn Ihnen etwas völlig überflüssig erscheint, können Sie diese Einrichtung aussortieren.“ ... „Würden Sie die einzelnen Häufchen jetzt noch so platzieren, wie sie zueinander in Beziehung stehen, also wie sie an einem Flughafen räumlich angeordnet sein sollen.“ Alle Sortiererergebnisse und Versuchspersonen-Angaben wurden auf einem Protokollbogen dokumentiert.

Tabelle 7.2-2. Versuchspersonen für das Sortierexperiment von Flughafeneinrichtungen

	Männer	Frauen
Vielflieger ≥ 6 Flüge pro Jahr	10 davon fliegen - 7 geschäftlich - 3 privat	10 davon fliegen: - 5 geschäftlich - 5 privat
Wenigflieger < 6 Flüge pro Jahr	10 davon fliegen: - 1 geschäftlich - 9 privat	10 davon fliegen: - 1 geschäftlich - 9 privat
Alter		
< 29	7	7
30 – 39	4	5
40 – 49	3	5
50 – 59	6	2
> 59	–	1
Summe	20	20

Ergebnisse

Auswertungsmethoden Die Daten wurden in zwei Stufen ausgewertet: Als *Erstes* wurden die *quantitativen Häufigkeiten*, mit denen die Einrichtungen sortiert wurden, ermittelt. Als *Zweites* wurden dann die Sortiererergebnisse für die häufig sortierten Einrichtungen mittels Screenshot, Multidimensionaler Skalierung MDS und Hierarchischer Clusteranalyse HCA nach Johnson (1967) mit Hilfe des Programmes CSS.STATISTICA ausgewertet (siehe Kapitel 6). MDS und HCA wurden jeweils auf die Originaldaten angewendet. Die HCA wurde nach dem Modell der complete linkage, das dem Maximum-Modell entspricht, berechnet (vgl. Gigerenzer 1981, Borg 1981). Als Signifikanzniveau wurde das 1% Niveau festgelegt. Dieses Signifikanzniveau ist angesichts der

kleinen Stichprobe (40 Vp) als hoch zu bewerten um sicherzugehen, dass die Nullhypothese irrtümlich abgelehnt wird.

Analyse der Sortierhäufigkeiten Ziel der quantitativen Analyse war es, diejenigen Einrichtungen auszuwählen, für die dann die MDS und HCA berechnet werden sollte. Es ergaben sich 5 Häufigkeitsgruppen: *Häufiger als 1,5-mal* wurden WC, Telefon, Bar, Snacks, Infostand, Café, Restaurant, Passkontrolle sortiert. Tabelle 7.2-3 zeigt die Einrichtungen, die *zwischen 1- bis 1,5-mal* sortiert wurden. Tabelle 7.2-4 zeigt die Einrichtungen, die *genau 1-mal* sortiert wurden. Sie wurden in der MDS und HCA ausgewertet. Tabelle 7.2-5 zeigt die Einrichtungen, die *zwischen 0,5- und 1-mal* sortiert wurden. Diese Einrichtungen wurden in MDS und HCA ausgewertet. Tabelle 7.2-6 zeigt die Einrichtungen, die *seltener als 0,5-mal* sortiert wurden.

Tabelle 7.2-3. Einrichtungen mit Sortierhäufigkeit 1,0 -1,5

Bank	Duty Free	Lounge	Zeitungen	Apotheke
Zoll	Gepäckaufbew.	Meeting-Point	Reisebedarf	Buchhandlung
Babywickelraum				

Tabelle 7.2-4. Gruppe 3. Einrichtungen mit Sortierhäufigkeit = 1

Ticket LH*	Ticket AA	Ticket BA*	Ticket TWA	Ticket Malaysia
Check-In LH	Check-In AA	Check-In BA	Check-In TWA	Check-In Malay
Gate LH	Gate AA	Gate BA	Gate TWA	Gate Malaysia
Warteraum LH	Warter. AA	Warteraum BA	Warteraum TWA	Warteraum Mala.
Ankunft LH	Ankunft AA	Ankunft BA	Ankunft TWA	Ankunft Malaysia
Gepäckaus. LH	Gepäckausg. Am	Gepäckaus. BA	Gepäckaus. TWA	Gepäckaus. Mala.
Check-In Air Con	Ankunft Air C.	Check-In LTU	Stadt-Information	Transfer
GateAir Conti	Warteraum LTU	Gate LTU	Hotel-Information	Last-Minute
Warteraum Air C*.	Ankunft LTU	Post	Polizei	Med. Zentrum
Gepäckauf. Air C.	Gepäckausg. LTU			

* LH = Lufthansa, ** BA = British Airways, *** AA = American Airways

Tabelle 7.2-5. Gruppe 4: Einrichtungen mit Sortierhäufigkeit 1 - 0,5

Karstadt-Reisen	Duschen	Tabak	Friseur	Elektro
Kreuzer-Reisen	Blumen	Foto	Delikatessen	Kapelle
TUI-Reisen	Souvenir	Mode	Sozialdienst	Lotto
Fundbüro	Parkleitzentrale	Seelsorge		

Tabelle 7.2-6. Einrichtungen mit der Sortierhäufigkeit $< 0,5$

Mietwagen	Frachtaufgabe	Reklamationen	Schuhe	Kinderbetreuung
Konferenzräume	Spielsalon	Schlaf/Ruheraum	Frauenraum	Notiztafel
Drogerie	Flugangstzentr.	Wäscherei	Behindertenserv.	Öffent. Ausrufer
Schließfächer	Kino			

Interpretation und Konsequenzen der Häufigkeitsanalyse Insgesamt zeigen die Ergebnisse, dass die Versuchspersonen die Einrichtungen differenziert beurteilen: es gibt Einrichtungen, die unterschiedlichen Kontexten zugeordnet werden (z.B. Toiletten, Telefone, Restaurationen) und deshalb mehrfach gruppiert werden. Andere Einrichtungen werden nicht von allen Personen gruppiert bzw. auch nicht am Flughafen erwartet (wie Spielsalon, Kino, Kinderbetreuung). Für die statistische Auswertung mittels MDS und HCA wurde die Anzahl begrenzt: Es wurden nur die Einrichtungen berechnet, die öfter als 0,5- und seltener als 1,5-mal sortiert wurden. Die Einrichtungen, die öfter als 1,5-mal sortiert werden (WC, Telefon, Bar, Snacks, Infostand, Cafe', Restaurant, Passkontrolle), werden über den ganzen Flughafen verteilt erwartet. Die Einrichtungen, die seltener als 0,5-mal sortiert wurden, werden nur von sehr wenigen am Flughafen erwartet. Es ist zu vermuten, dass sie nur von spezifischen Nutzergruppen erwartet werden. Da sie so häufig bzw. selten sortiert wurden, wurden sie aussortiert. Ihre Ergebnisse wären nicht aussagekräftig.

Analyse der Ähnlichkeitsurteile durch Scree-test, MDS- und HCA Die Einrichtungen mit einer Sortierhäufigkeit größer 0,5 und kleiner 1,5 wurden mittels Scree-test, MDS und HCA ausgewertet. Es wurden zunächst Distanzmatrixen erstellt. Da die meisten Versuchspersonen mehrere Unterteilungen vornahmen, ergaben sich für diese Versuchspersonen Matrizen auf unterschiedlichen Hierarchieebenen. Dabei wurde immer von der untersten (= feinsten) Gruppenbildung ausgegangen.

Ausschluss von Versuchspersonen Zwei von 40 Versuchspersonen wurden aus der Auswertung ausgeschlossen. Beide Personen waren weiblich und hatten nur wenig Flugerfahrung. Sie sortierten nach einem anderen, vollständig eigenständigen Modell: Sie bauten jeder Fluggesellschaft quasi einen kleinen Terminal mit eigenen Läden, Restaurations- und Serviceeinrichtungen.

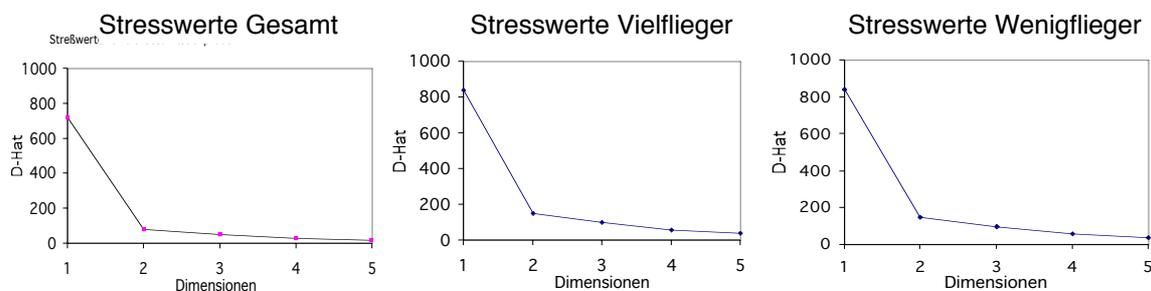
Ergebnisse der Stresswerte für die MDS-Konfigurationen Mittels des Scree-tests wurden für euklidische MDS Lösungen die Stresswerte berechnet. Tabelle 7.2-7 zeigt die Stresswerte für die 2- und 3-dimensionalen Konfigurationen diverser Stichproben:

Tabelle 7.2-7. Stresswerte der 2- und 3-dimensionalen MDS-Konfigurationen

Stichproben	Stresswerte für die 2-dimensionale MDS-Konfiguration	Stresswerte für 3-dimensionale MDS-Konfiguration
Gesamtstichprobe	0.1162 *	0.0744 *
Vielflieger	0.1522 *	0.1020 *
Wenigflieger	0.1172 *	0.0735 *
Frauen	0.1375 *	0.0951 *
Männer	0.1296 *	0.0809 *
weibl. Vielflieger	0.1799 *	0.1213 *
weibl. Wenigflieger	0.1317 *	0.0888 *
männl. Vielflieger	0.1750 *	0.1172 *
männl. Wenigflieger	0.1284 *	0.0830 *

* Stresswerte = < kritische Stresswerte $\alpha = 0.01$ nach Spence/Ogilvie 1973, in: Gigerenzer 1981

Die Ergebnisse liegen in jeder Stichprobe für die 2-dimensionalen Lösungen unterhalb des kritischen Stresswertes bei einer Irrtumswahrscheinlichkeit von $\alpha = 0.01$. (vgl. Gigerenzer 1981 S. 353). Abbildung 7.2-3 zeigt die Stresswerte für die drei wichtigen Untersuchungsstichgruppen in Diagrammform. Die Diagramme zeigen den Ellenbogenknick für die 2-dimensionale Konfiguration, was heißt, dass sie die am besten angepassten sind.



Abbildungen 7.2-3. Stresswert-Diagramme für die Stichproben Gesamt, Viel- und Wenigflieger

Es wird daher die 2-dimensionale Lösung gewählt. Ein signifikanter Stresswert zeigt, dass die Ähnlichkeitsurteile und somit die räumlichen Distanzen in der MDS-Konfiguration nicht zufällig zustande gekommen sind. Die Anordnung der Serviceeinrichtungen in der 2-dimensionalen MDS Konfiguration ist als überzufällige, stabile Struktur anzusehen.

Sortierergebnisse in der MDS und HCA-Auswertung für die Gesamtstichprobe MDS und HCA basieren auf unterschiedlichen Distanzbegriffen: die MDS auf der euklidischen Metrik, die HCA auf ultrametrischen Distanzen). Die MDS stellt ein einfaches Näherungsmodell dar, das einen komplexen Sachverhalt räumlich abbildet. Die Clusteranalyse trägt zur Erklärung weiterer Varianzen bei. Bei den erhobenen Daten stimmen die Ergebnisse der beiden Methoden weitgehend überein, sodass die Clusteranalyse weitgehend die MDS-Konfiguration bestätigt. Ausnahmen sind durch lang gestreckte Umkreisungen zu erkennen (z.B. beim Babywickelraum). Abbildung 7.2-4 zeigt die Ergebnisse für die Gesamtstichprobe.

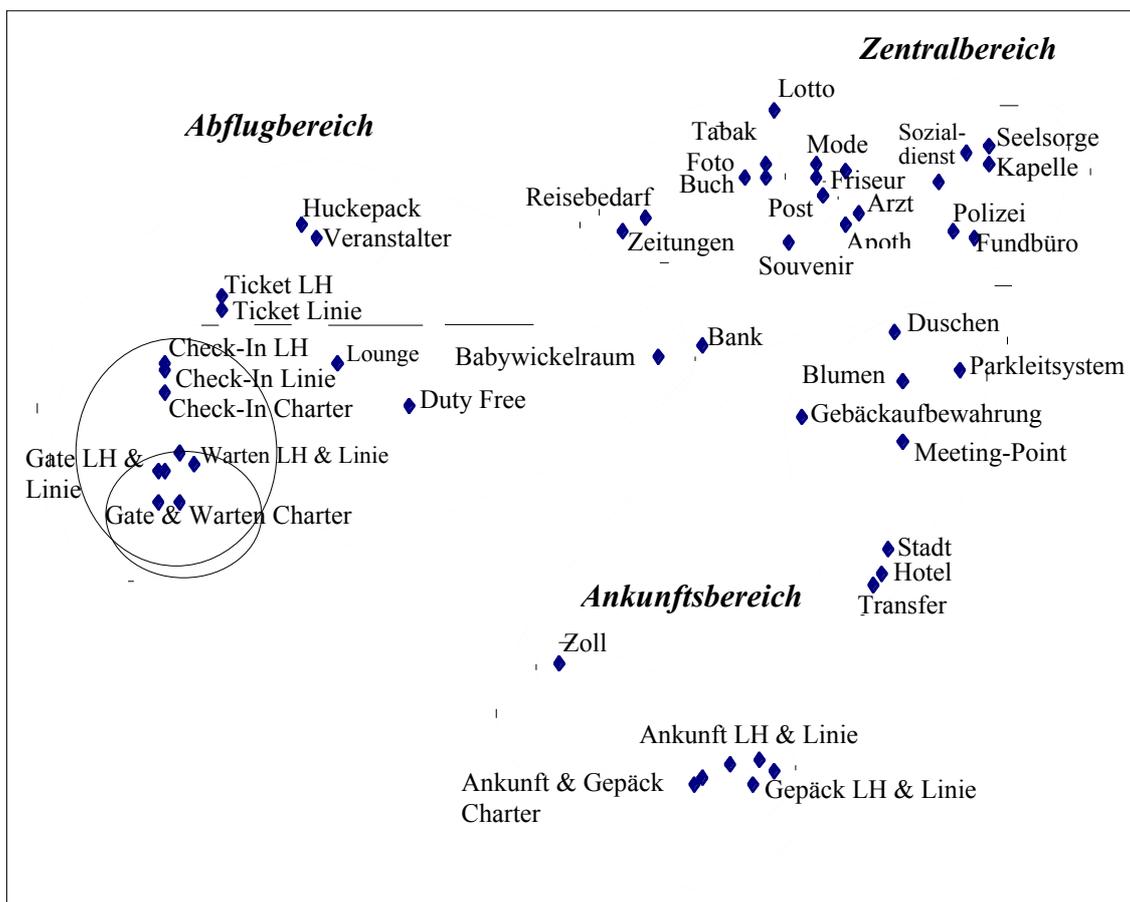


Abbildung 7.2-4. Integrierte Ergebnisse von MDS und HCA für die **Gesamtstichprobe**. Es sind die drei Cluster Ankunfts-, Abflug-, Zentralbereich zu erkennen

Es wurden in die 2-dimensionalen MDS-Konfigurationen (Punkte) die Ergebnisse der Clusteranalyse (Umkreisungen) eingetragen. In der MDS-Konfiguration zeigen sich zwei orthogonale Dimensionen, denen entlang die Serviceeinrichtungen gruppiert sind (ankunfts- versus die abflugbezogene Serviceeinrichtungen sowie flugnahe versus öffentliche Serviceeinrichtungen und Geschäfte). In der HCA zeigen sich drei Cluster mit Serviceeinrichtungen für die Ankunft, den Abflug und zentrale Geschäfts- und Serviceeinrichtungen interpretiert werden können. Inhaltlich wird hierauf in der Interpretation eingegangen.

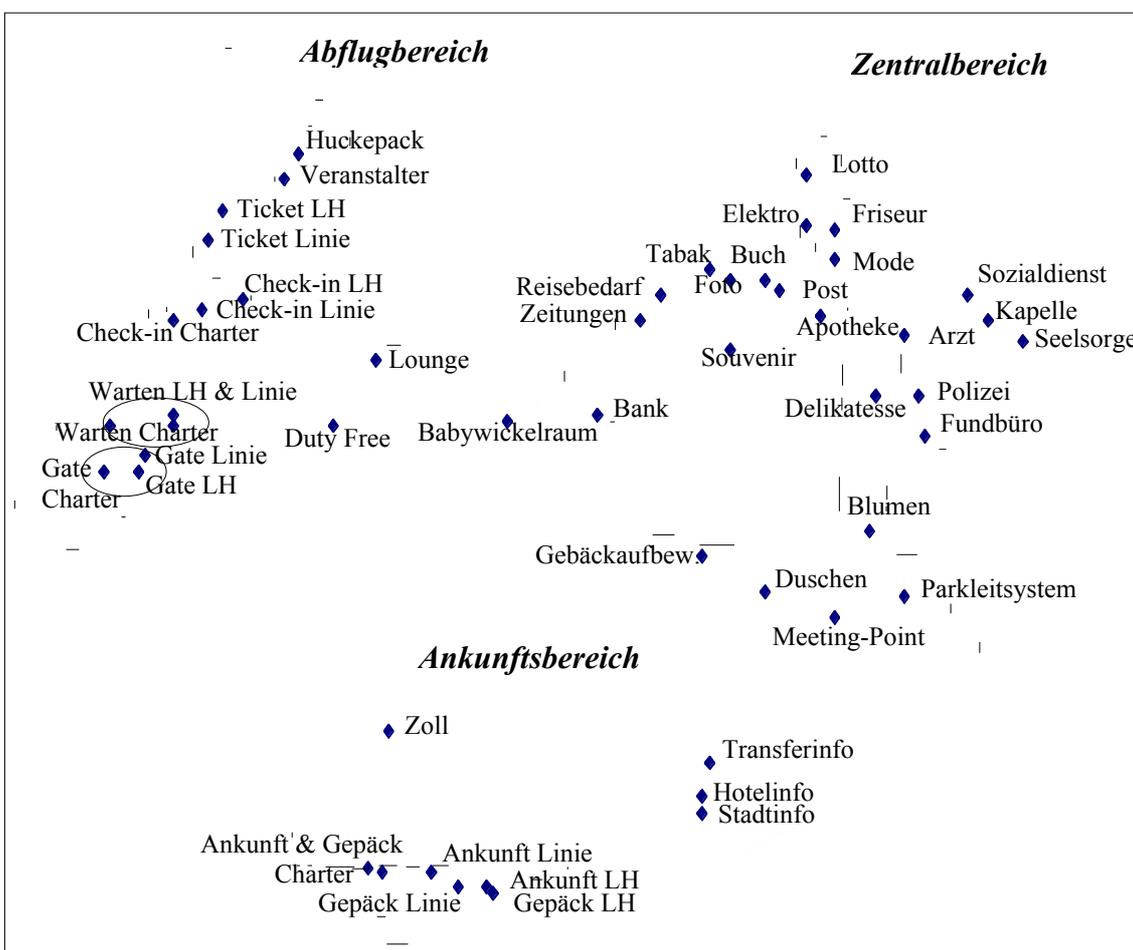


Abbildung 7.2-5. Integrierte MDS und HCA Ergebnisse für die **Vielflieger** (≥ 6 Flüge pro Jahr) Es sind die drei Cluster Ankunfts-, Abflug-, Zentralbereich zu erkennen, die zur Leseunterstützung bereits interpretiert sind.

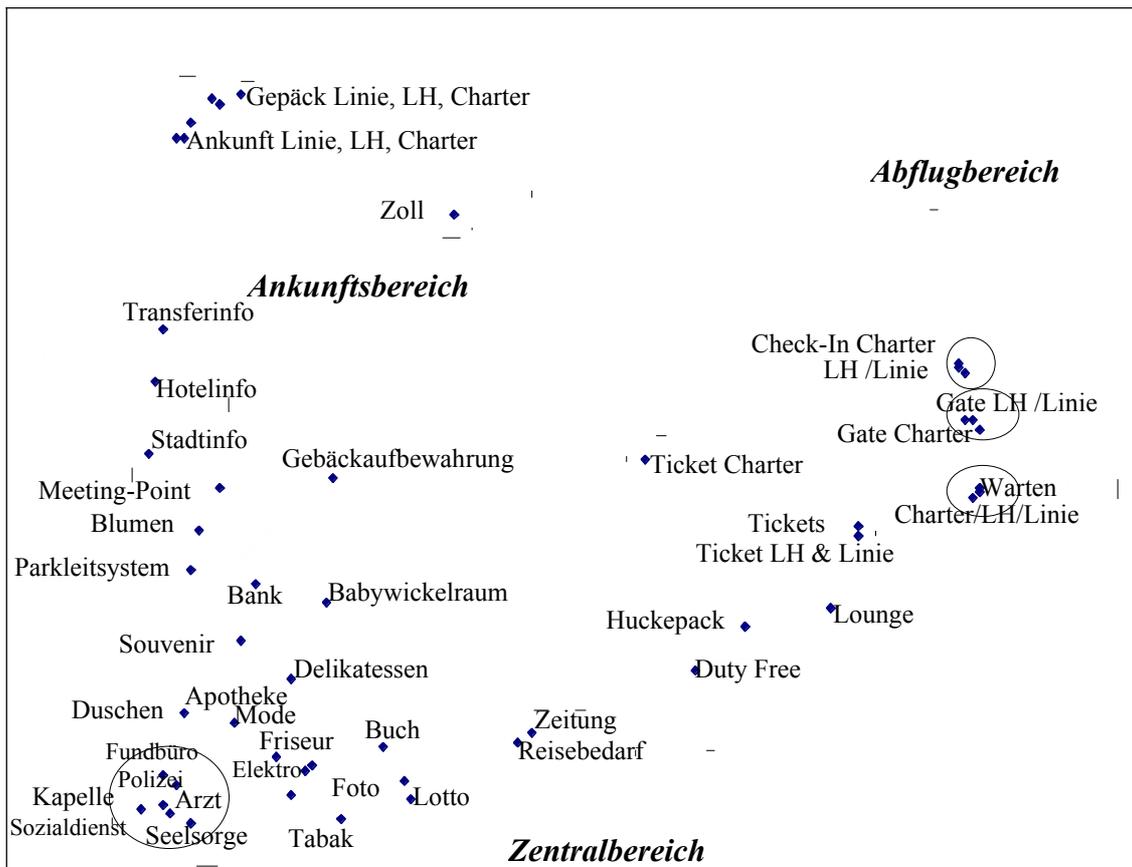


Abbildung 7.2-6. Integrierte MDS und HCA Ergebnisse für die **Wenigflieger** (< 6 Flüge pro Jahr) Es sind die drei Cluster Ankunfts-, Abflug-, Zentralbereich zu erkennen, die zur Leseunterstützung bereits interpretiert sind.

Ergebnisse der Stichproben Vielflieger und Wenigflieger Die Abbildungen 7.2-5 und 7.2-6 zeigen die integrierten Ergebnisse von MDS und HCA für die beiden Stichproben Vielflieger (6 oder mehr Flüge pro Jahr) und Wenigflieger (weniger als 6 Flüge pro Jahr): Die Unterschiede zwischen den beiden Stichproben sind geringfügig. Die verschiedenen Stresswerte (Vielflieger 0.1522 und Wenigflieger 0.1172) zeigen, dass die Komplexität der Sichtweise zwischen den Versuchspersonengruppen etwas variiert. Die Ergebnisse der Wenigflieger zeigen aufgrund des niedrigeren Stresswertes eine größere Stabilität, was auf eine größere Übereinstimmung der Ähnlichkeitsurteile hinweist. Die Stresswerte für die Konfigurationen beider Stichproben sind signifikant, was darauf schließen lässt, dass die Ähnlichkeitsurteile bzw. Distanzen nicht zufällig zustande gekommen sind. Im Folgenden werden die Unterschiede für die Viel- und Wenigflieger in den drei Bereichen Abflug, Ankunft und Zentraler Bereich dargestellt.

Unterschiede der Stichproben im Abflugbereich Auffallend ist hier die unterschiedliche Beurteilung der Lufthansa zwischen Viel- und Wenigfliegern. (Die Lufthansa war zu Zeit der Erhebung die einzig innerdeutsch fliegenden Fluggesellschaft.) Die *Vielflieger* betonen die Lufthansa, und unterscheiden sie deutlich von den internationalen Linien-Fluggesellschaften und den Charter-Fluggesellschaften – sowohl bei den Ticketschaltern, bei den Warteräumen als auch bei den Gates. Die *Wenigflieger* unterscheiden nur zwischen Linien- und Charter-Fluggesellschaften und beurteilen die Lufthansa ähnlich wie die damals nur international fliegenden Linien-Fluggesellschaften British Airways (BA), Malaysia Airways etc..

Unterschiede der Stichproben im Ankunftsbereich Die *Vielflieger* trennen auch hier zuerst deutlich die zwischen den Fluggesellschaften: der Lufthansa, den Internationalen Linienfluggesellschaften und den Charter-Fluggesellschaften. Sie ordnen jeder Gesellschaftsart Ankunftsterminals und Gepäckbereich zu. Die *Wenigflieger* sortieren einen Ankunftsterminal mit von Linien- und Charter-Fluggesellschaften und davon getrennt einen Gepäckbereich für Linien- und Charter-Fluggesellschaften. Die Lufthansa sortieren sie zu den Linien-Fluggesellschaften.

Die *Vielflieger* ordnen dem Ankunftsbereich außerdem sieben Serviceeinrichtungen zu, Stadtinformationen, Hotelinformationen, Transferinformationen und davon etwas distanzierter Meeting Point und Parkleitsystem, Gepäckaufbewahrung und die Duschen. Die *Wenigflieger* hingegen ordnen dem Ankunftsbereich nur fünf Serviceeinrichtungen zu. Sie sortieren die Gepäckaufbewahrung und die Duschen zum Zentralbereich.

Unterschiede und Gemeinsamkeiten der Stichproben im Zentralen Bereich Die *Vielflieger* ordnen die Bank in die Nähe des Abflugbereiches, die *Wenigflieger* ordnen die Bank in die Nähe des Ankunftsbereiches. Vielflieger beurteilen Zeitungen, Reisebedarf, Foto, Buch und Tabak als weitaus ähnlicher dem Abflug als die Wenigflieger. Deutlich stimmen beide Gruppen in der Beurteilung der folgende Einrichtungen überein: Blumen (ankunftsnah), Mode, Friseur, Delikatessen, Elektro, Lotto, Post, Apotheke (Zentralbereich Mitte) und Fundbüro, Polizei, Medizinisches Zentrum, Seelsorge, Kapelle und Sozialdienste (flugfern, öffentlich).

7.2.3 Interpretation der Ergebnisse

Interpretation im Hinblick auf die Hypothese 1: Personenübergreifende, stabile Ähnlichkeitsbeurteilung Die Untersuchungsergebnisse bestätigen die Annahme, dass es personenübergreifende Ähnlichkeitsbeurteilungen von Fluggästen für Flughafeneinrichtungen gibt. Hierfür sprechen die signifikanten Stresswerte und die eindeutigen Konfigurationen und Cluster der Multidimensionalen Skalierung und der Hierarchischen Clusteranalyse. Die Dimensionen, die sich in der MDS zeigten, können als *ankommend* \leftrightarrow *abfliegend*, *öffentlich* \leftrightarrow *flugnah* bezeichnet werden (Abb. 7.2-7).

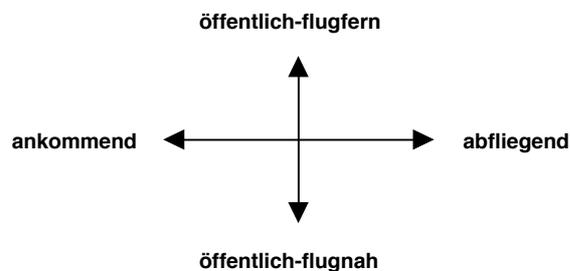


Abbildung 7.2-7. Dimensionen der 2-dimensionalen MDS-Konfiguration Flughafeneinrichtungen

Die drei HCA-Cluster, die mit den Dimensionen in der MDS-Konfiguration übereinstimmen, können beschrieben werden als Abflugbereich (Ticketschalter, Gates, Wartebereiche etc.), Ankunftsbereich (Ankunftsterminals, Gepäckausgabe, Zoll etc.) und öffentlicher Bereich (Zeitungen, Souvenir, Apotheke, etc.) (Abb. 7.2-8). Die drei Cluster und die beiden MDS-Dimensionen widersprechen sich nicht sondern korrespondieren.



Abbildung 7.2-8. Cluster der HCA Flughafeneinrichtungen

Die Hypothese 1 wird aufgrund der Ergebnisse angenommen. Da die MDS-Konfiguration und HCA-Cluster inhaltlich bereits offensichtlich sind, kann auf eine empirische Nachuntersuchung zur Benennung der Dimensionen verzichtet werden.

Interpretation in Hinblick auf die Hypothese 2: Bisherige Handlungserfahrungen und Hypothese 3: Handlungsbezogene Kriterien Die Unterschiede zwischen den beiden Stichgruppen sind geringfügig. Die Stresswerte zeigen, dass die Komplexität der Sichtweise zwischen den einzelnen Stichgruppen variiert. Doch bei 2-dimensionalen Konfigurationen werden zwischen 82% und 89% der Varianz der Stichgruppensortierung erklärt.

Es treten dennoch *Unterschiede* auf, die sich auf den Faktor Flugerfahrung zurückführen lassen: die Vielflieger unterscheiden eindeutig im *Abflugbereich und im Ankunftsbereich* zwischen der Lufthansa als damals einziger Inlandsfluggesellschaft, internationalen Liniengesellschaften und Chartergesellschaften, während die Wenigflieger nur zwischen Linien- und Charterfluggesellschaften unterscheiden. Dies lässt sich durch unterschiedliche Handlungsziele erklären: Unter den Vielfliegern sind viele geschäftlich fliegende Personen (12 von 20). Sie nutzen die Lufthansa für geschäftliche Inlandsflüge. Die Wenigflieger (18 von 20 Wenigfliegern fliegen nur privat) fliegen in der Regel mit Charter in den Urlaub. Für sie hat die Lufthansa keine besondere Handlungsrelevanz und wohl deshalb auch keine besondere Bedeutung, sondern – wie die Ergebnisse schließen lassen – die gleiche wie die internationalen Fluggesellschaften. Die Vielflieger ordnen im Ankunftsbereich Ankunftsterminal und Gepäckbereich für jede Fluglinie zusammen. Wenigflieger bilden einen Ankunftsterminal für alle Fluggesellschaften und davon getrennt einen Gepäckbereich für alle Fluggesellschaften. Auch dies verweist auf das Handlungsziel der Vielflieger: nach der Landung schnell das Gepäck zu erhalten.

Auch die Unterschiede im *Zentralbereich*, wie z.B. bei der Beurteilung der Gepäckaufbewahrung und den Duschen im Ankunftsbereich können auf unterschiedliche Handlungsziele für die Stichproben zurückgeführt werden: Für Urlaubsreisende (mit Charter) besteht in der Regel kein Bedarf, Gepäck am Flughafen zu deponieren, da die Fahrt ins Hotel unmittelbar bevorsteht. Dort kann der Urlaubsreisende dann auch duschen. Für Geschäftsreisende kann es sehr wichtig sein, Gepäck am Flughafen zu deponieren, da es beim Kundenbesuch stören kann. In der Regel kehrt der geschäftlich fliegende Fluggast an den Flughafen zurück, um nach Hauser bzw. weiter zu fliegen. Auch die Dusche kann für den geschäftlich fliegenden Gast von Interesse sein, um sich nach einem langen Flug vor dem Kundentermin frisch zu machen.

Auch die unterschiedliche Beurteilung der Bank weist auf unterschiedliche Handlungsabläufe der Stichgruppe hin. Geschäftliche Vielflieger beurteilen die Bank ähnlich dem Abflugbereich, was darauf schließen lässt, dass sie vor dem Abflug Geld holen bzw. wechseln. Wenig fliegende Urlauber beurteilen die Bank ähnlich dem Ankunftsbereich, was darauf schließen lässt, dass sie nach der Ankunft Geld holen bzw. wechseln. Dies kann durch unterschiedliche Handlungskontexte erklärt werden: der geschäftlich fliegende Vielflieger hat eher vor dem Flug noch Zeit, Geldgeschäfte zu erledigen und startet nach der Landung gleich zu seinem Termin. Der urlaubsreisende Wenigflieger scheint hingegen nach der Landung eher Zeit zu haben, Geld zu wechseln bzw. hofft auf einen günstigeren Wechselkurs.

Die Vielflieger beurteilen Tabak, Buch, Foto, Reisebedarf und Zeitung als ähnlich und nahe dem Abflug. Dies kann als Gruppe von Geschäften interpretiert werden, die für die Vielflieger relevant sind. Geschäfte, die eine höhere Verweildauer benötigen (Elektro, Mode, Friseur), weil man sich mit dem Sortiment intensiver auseinandersetzen muss, sind hier eindeutiger als ähnlich beurteilt. Sie bilden die Mitte des Zentralbereiches für die Vielflieger. Bei den Wenigfliegern sind diese beiden Gruppierungen – „schnelle Geschäfte vor dem „take-off“ und „zeitintensive Geschäfte“ – nicht so deutlich zu erkennen.

Vielflieger und Wenigflieger stimmen in hohem Maße darin überein, dass die sechs Einrichtungen – Fundbüro, Polizei, Medizinische Station, Kapelle, Seelsorge und Sozialdienst – untereinander sehr ähnlich sind, aber weit vom Flugbetrieb entfernt sind. Sie haben nicht direkt etwas mit dem Flugbetrieb zu tun, aber sie scheinen dennoch zum Flughafen zu gehören: sie gehören zu den Einrichtungen, die von den Versuchspersonen aus- bzw. sehr selten sortiert wurden.

Nur 2 von 40 Versuchspersonen sortierten nach kategorialen Kriterien: sie bauten jeder Fluggesellschaft einen „eigenen“ Terminal und ordneten ihnen eigene Geschäfte, Serviceeinrichtungen, Restaurationsbetriebe etc. zu. Es zeigte sich aber, dass beide Versuchspersonen sehr wenig Flugerfahrung hatten. Es ist anzunehmen, dass sie auch wenig begriffliches Wissen über Flughäfen besitzen. Die Beurteilung nach dem kategorial-

len Kriterium könnte eventuell ein Indiz dafür sein, dass eine Person bisher wenig bis keine Handlungserfahrungen mit dem Untersuchungsgegenstand gemacht hat.

Die Ausführungen verweisen darauf, dass den Ähnlichkeitsbeurteilungen handlungsbezogene Kriterien zugrunde liegen. Die semantische Ähnlichkeit der Begriffe drückt zeitlich-räumliche Abfolgen von Handlungsabläufen aus. Definiert man Handlungsabläufe als Aktionen zur Erreichung von Handlungszielen, kann formuliert werden, dass den semantischen Ähnlichkeitsbeurteilungen Handlungsziele zugrunde liegen: die Serviceeinrichtungen gliedern sich entsprechend den Handlungszielen Ankommen, Abfliegen mit den für die Nutzergruppen spezifischen Handlungszielen bezüglich der Versorgung durch Geschäfte etc. Dieses Ergebnis geht noch über die Annahmen von Klix hinaus. Es weist den Handlungszielen eine grundlegende Bedeutung für die begriffliche Nähe bzw. Distanz zu: was hinsichtlich eines Handlungszieles gemeinsam verwendet wird, ist begrifflich nahe. Diese Annahme könnte verwendet werden, um die Theorie von Klix zu überarbeiten.

Zur Überprüfung einer nutzerorientierten Anordnungen Die Ergebnisse dieser Studie wären geeignet, daraus Vorgaben für Architekten und Bauherren zu entwickeln. Bleibt noch die Frage zu klären, ob eine Anordnung entsprechend den Ergebnissen dieser Studie tatsächlich das Such- und Findeverhalten der Flughafengäste verbessern kann. Eine Überprüfung dieser Hypothese wäre – ohne gleich einen neuen Flughafen zu bauen – auf folgendem Wege möglich: Anhand der Ergebnisse dieser Arbeit könnten Hypothesen über die Such- und Findefreundlichkeit von bereits existierenden Flughäfen gemacht werden: durch einen Vergleich der hier empirisch gefundenen Vorgabe und des Grundrissplans eines existierenden Flughafens könnten Hypothesen darüber formuliert werden, welche Einrichtungen einfach oder schwer zu finden sind. Dabei dürfen die Untersucher den Flughafen natürlich nur vom Grundrissplan kennen und nicht in Realität. Dann könnten vor Ort Flughafengäste mittels Laufprotokollen und Interviews beobachtet und befragt werden, um diese Hypothesen zu testen. Zeigen die Flughafengäste tatsächlich das prognostizierte Verhalten, wäre dies eine Bestätigung, dass eine Anordnung entsprechend der Begriffsstruktur benutzerfreundlicher wäre.

7.2.4 Beurteilung der verwendeten Methoden

Als nicht ideal erwies sich im Nachhinein die Reduktion der Flughafeneinrichtungen durch die Integration in den Sortiersversuch. Die Versuchspersonen hatten einen hohen Aufwand (zeitlich und mental), um 84 Einrichtungen zu sortieren. Die Reduktion anhand der Häufigkeiten ist als fragwürdig zu beurteilen, weil Einrichtungen wie Mietwagen, Konferenzzentrum oder Behindertenservice aufgrund der geringen Sortierhäufigkeit nicht weiter untersucht wurden. Diese Einrichtungen sind aber wichtig für definierte Kundengruppen (Geschäftspersonen, Behinderte etc.) am Flughafen. Es empfiehlt sich für weitere Studien die Reduktion in einem Vorversuch unter Einbezug der Versuchspersonen zu machen. Die Methode des Free Card Sorting erwies sich in der Durchführung für die Versuchspersonen als einfach, aber als aufwändig in der Auswertung. Da ein Flughafen sehr viele verschiedene Kundengruppen zufrieden stellen muss, scheint es ratsam zu sein, die unterschiedlichen Nutzergruppen zu ermitteln, mit ihnen gesonderte Untersuchungen zu machen und die Ergebnisse zusammenzutragen.

7.2.5 Zusammenfassung

Ausgangspunkt waren die theoriegeleiteten Prognosen, dass sich für Flughafengäste personenübergreifende, stabile Ähnlichkeitsbeurteilungen für Flughafeneinrichtungen finden lassen, und dass diese Unterschiede für Viel- und Wenigflieger aufweisen aufgrund unterschiedlicher Handlungsvorerfahrungen der beiden Nutzergruppen. 40 Personen (je 20 Viel- und Wenigflieger) sortierten 84 Flughafeneinrichtungen mittels der Methode des Free-Card-Sorting und ihre Ergebnisse wurden mittels MDS, Screeplot und HCA ausgewertet. Die Interpretation erfolgte hypothesengeleitet, dass sich für Viel- und Wenigflieger Unterschiede in den semantischen Ähnlichkeitsbeurteilungen zeigen, und dass die Ähnlichkeitsbeurteilung handlungsbezogenen Kriterien folgt

Die Ergebnisse der Studie zeigen, dass sich für die Versuchspersonen eine personenübergreifende Ähnlichkeitsbeurteilung finden lässt. In der HCA zeigen sich deutlich drei Cluster: Abflug-, Ankunft- und Zentralbereich mit Serviceeinrichtungen und Geschäften. In der MDS zeigen sich die beiden Dimensionen „ankommend/abfliegend“,

und „öffentlich/flugnah“. Im Zentralbereich spannen sich die Serviceeinrichtungen und Läden zwischen den Polen „flugnah: zum Flugbereich gehörend“ und „öffentlich: zum öffentlichen Bereich“ gehörend auf. So zeigen z.B. die Ergebnisse aller Stichproben, dass der Blumenverkauf im Zentralbereich nahe dem Ankunftsbereich sortiert wird, während das Fundbüro, die Seelsorge und die Polizei als sehr unähnlich dem Flugbereich beurteilt werden.

Für die Stichgruppen Vielflieger und Wenigflieger zeigen sich geringe Unterschiede, die sich aber eindeutig durch unterschiedliche Handlungsziele und -erfahrungen erklären lassen. So geben die Vielflieger (überwiegend geschäftlich fliegende Personen) der Lufthansa – als damals einziger Fluggesellschaft mit innerdeutschen Flügen – eine eigenständige Bedeutung, während die Wenigflieger (fast ausschließlich privat fliegende Personen) die Lufthansa den internationalen Linienfluggesellschaften zuordnen. Auch die unterschiedliche Beurteilung der Bank, der Gepäckaufbewahrung, der Duschen und der Geschäfte für den kurzen Einkauf (Zeitung, Tabak, Reisebedarf) lässt sich auf unterschiedliche Bedürfnisse und Handlungsziele der beiden Stichgruppen erklären. Die Ergebnisse lassen sich dahingehend interpretieren, dass die Handlungsziele eine Ordnung stiftende Funktion für die Begriffsstruktur haben.

7.3 Ermittlung einer käuferorientierten Anordnung von Zimmerpflanzen für Verkaufsflächen

7.3.1 Einleitung

Problemstellung Die folgende Untersuchung beschäftigt sich mit der kundenorientierten Platzierung von Zimmerpflanzen auf der Verkaufsfläche: Wie müssen Zimmerpflanzen – je nach Saison zwischen 40 und 80 Artikel – angeordnet sein, damit die KundInnen sich möglichst gut zurechtfinden? Wo sucht die Kundin Orchideen oder das Usambara-Weilchen? Was ist die erwartungsgerechte Anordnung von Efeu, Farn und Gummibaum?

Anlass der Studie Anlass der Studie war die Optimierung einer Zimmerpflanzen-Abteilung in einem Gartencenter¹. Die KundInnen fragten das Verkaufspersonal häufig, wo welche Zimmerpflanze zu finden sei. Durch das Fragen hat das Verkaufspersonal weniger Zeit für die Beratung, vor allem von verkaufsintensiveren, exotischen und hochpreisigen Zimmerpflanzen. Der Besitzer hatte daher ein großes Interesse, dies zu ändern. In der Untersuchung sollte experimentell abgesichertes Wissen für eine kundenorientierte Anordnung der Zimmerpflanzen ermittelt werden, damit die KundInnen sich beim Kauf von Zimmerpflanzen besser zurechtfinden und bei ihrer Suche von der Hilfe des Fachpersonals unabhängiger werden.

Ziele der Studie Die Studie verfolgt sechs Ziele, vier inhaltliche und zwei methodische:

1) *Inhaltlich* sollte die Untersuchung klären, ob die KundInnen eine gemeinsame Begriffsstruktur für Zimmerpflanzen haben und wenn ja, wie diese aussieht. Die Begriffsstruktur wird als semantische Ähnlichkeit operationalisiert. Es wird davon ausgegangen, dass die KundInnen im Verhalten mit Zimmerpflanzen Wissen ausbilden, dass in Begriffen und Begriffsstrukturen repräsentiert ist. KundInnen müssten bezüglich Zimmerpflanzen eine ähnliche Begriffsstruktur haben.

2) *Inhaltlich* sollte die Untersuchung des Weiteren den Einfluss unterschiedlicher Handlungserfahrungen abklären. Um dies zu untersuchen, wurden die Ähnlichkeitsbeurteilun-

¹ Barbara Schroeder-Köstler führte die Untersuchung im Rahmen einer Diplomarbeit durch.

gen für „Zimmerpflanzen“ an den beiden Nutzergruppen KundInnen und Fachkräfte (Gärtner- und FachverkäuferInnen) erhoben und verglichen. Beide Nutzergruppen unterscheiden sich in ihren Handlungserfahrungen, ihren Handlungszielen, Handlungsabfolgen, in ihren Interessen und Motiven im Umgang mit Zimmerpflanzen. Dies müsste sich folglich in unterschiedlichen Begriffsstrukturen ausdrücken (vgl. Klix 1971, Hoffmann 1983, 1986). Wie bereits in Kapitel 5 dargestellt, zeigte Zimmer (1984) in einem Laborexperiment mit Pflanzen die unterschiedliche Verfügbarkeit des begrifflichen Wissens bei Floristen und Laien. Auch die Arbeiten von Henley (1969), Storm (1980), Hejj & Strube (1988), Chi & Glaser 1980 sowie Maichle 1985 weisen eine Abhängigkeit des Begriffsfeldes von Alter, Ausbildung und Aufgabenbereich nach. Untersuchungen über Begriffsfelder bezüglich Gebrauchs- und Konsumgüter des Alltags lagen nicht vor. Untersuchungen mit Novizen und Experten sprechen von einer so genannten Expertise (vgl. z.B. Larkin 1979, Chi und Glaser 1980, Anderson 1993). Sie nennen sogar Kriterien für „typische“ Sortierergebnisse von Experten: Sie weisen klarere Strukturen auf; d.h. es gibt weniger Begriffscluster, die individuellen Variationen sind geringer und sie beurteilen eher nach Handlungsgesichtspunkten. Laien hingegen beurteilen eher nach sensorischen Kriterien, die individuelle Variation ist größer und die Ergebnisse zeigen mehr oder weniger klare Cluster. Wie bereits in Kapitel 6 dargestellt, wird jedoch in dieser Untersuchung nicht von einem Experten-Novizen-Vergleich gesprochen! KundInnen und Fachkräfte werden als eigenständige Nutzergruppen mit unterschiedlichen Interessen und Handlungserfahrungen betrachtet. Es wird dennoch angenommen, dass sich beide Gruppen in der Ähnlichkeitsbeurteilung von Zimmerpflanzen unterscheiden.

3) *Inhaltlich* sollte geklärt werden, welchen Kriterien die Begriffsstruktur folgt. Basierend auf der Theorie von Klix (1971), Hoffmann (1983,1986) et al. wird davon ausgegangen, dass die Begriffsstruktur der KundInnen *und* der Fachkräfte handlungsbezogenen Kriterien folgen und nicht kategorialen. Es wird angenommen, dass die Begriffsstruktur der KundInnen Kriterien folgt, die mit dem Pflegen und der häuslichen Verwendung zu tun haben, und dass die Begriffsstruktur der Fachkräfte Kriterien folgt, die mit der Pflege, der floristischen Verwendung und dem Verkauf zu tun haben und nicht einer kategoriale Pflanzensystematik folgt.

4) *Inhaltlich* sollte die Untersuchung außerdem den Einfluss der „aktuellen Zielorientierung“ abklären. Im Hauptversuch wurde den KundInnen die Zielorientierung über die Instruktion vorgegeben, nach denen sie die Zimmerpflanzen sortieren sollten: Eine KundInnen-Stichprobe erhielt die Zielorientierung „Kaufen“. Diese Versuchspersonen sollten sich beim Sortieren vorstellen, dass sie einen Einkauf tätigen. Die andere KundInnen-Stichprobe erhielt keine inhaltliche Vorgabe. Sie sollten sortieren, wie es ihnen beliebt.

5) *Methodisch* wurde versucht, die Auswahl der Objekte für das Sortierexperiment zu verbessern. Die Auswahl sollte nicht dem Zufall überlassen werden. Mit der Auswahl sollte das ganze Sortiment abgedeckt werden. Denn werden bestimmte Zimmerpflanzenarten (z.B. Orchideen) nicht untersucht, erhält man später auch keine Erkenntnisse darüber, wie diese in Bezug zu den anderen Zimmerpflanzen angeordnet werden sollen. Im Vorversuch sollte sichergestellt sein, dass die Auswahl der Elemente für das Sortierexperiment repräsentativ für das Gesamtsortiment ist. Darüber hinaus müssen die zu sortierenden Pflanzen den Versuchspersonen bekannt sein, damit sie sinnvoll beurteilt werden. Es wurde deshalb ein Vorgehen entwickelt, um eine für das Sortiment repräsentative *und* für die KundInnen bekannte Auswahl von Pflanzen zu treffen.

6) *Methodisch* wurde schließlich versucht die Interpretation der Sortiererergebnisse empirisch zu objektivieren. Im Nachversuch sollte daher Erfahrung mit dem empirischen Interpretationsverfahren, der Property Fitting Analyse, gesammelt werden.

In dieser Studie wurden die folgenden Hypothesen untersucht:

Hypothese: Personenübergreifende stabile Ähnlichkeitsbeurteilung Die Ähnlichkeitsbeurteilungen der Nutzergruppen für Zimmerpflanzen zeigen eine stabile, ausgeprägte Struktur; d.h. die Ähnlichkeitsbeurteilungen der Zimmerpflanzen erfolgen nicht nach individuellen Kriterien, sondern ihnen liegen gemeinsame, überindividuelle Kriterien zugrunde.

Hypothese: Bisherige Handlungserfahrung Die Ähnlichkeitsbeurteilungen von Zimmerpflanzen sind abhängig von den bisherigen Handlungsvorerfahrungen der Versuchspersonen: KundInnen beurteilen die Zimmerpflanzen entsprechend anders als die Fachkräfte.

Hypothese: Die Ähnlichkeitsbeurteilung von Objekten ist unabhängig von der aktuellen Zielorientierung der Personen. Ähnlichkeitsbeurteilungen von Zimmerpflanzen sind unabhängig von der aktuellen „Zielorientierung“. Die Ergebnisse der KundInnen mit der über die Instruktion gegebenen Zielorientierung: „Kaufen von Zimmerpflanzen“ unterscheiden sich nicht von denen einer neutralen Instruktion „ohne“ Zielorientierung.

Hypothese: Handlungsbezogene Kriterien Den Ähnlichkeitsbeurteilungen von Zimmerpflanzen liegen handlungsbezogene Kriterien (z.B. Kaufziel, Pflege, Warenpräsentation) zugrunde und nicht kategorial-hierarchische Kriterien (z.B. Pflanzenfamilie).

Untersuchungsdesign Die Untersuchung besteht aus einem Vorversuch, einem Hauptversuch, einem Nachversuch und einer Nachbefragung (Abb. 7.3-1).

Vorversuch:

Auswahl von Zimmerpflanzen für das Sortierexperiment

8 KundInnen und 1 Fachkraft, quantitative und qualitative Befragung

Hauptversuch: Sortierexperiment

Ermitteln der Begriffsstruktur für Zimmerpflanzen

Unabhängige Variablen: UV 1: Fachkräfte / KundInnen UV 2: aktuelle Zielorientierung

Abhängige Variablen: AV 1: semantische Ähnlichkeit, AV 2: Klassifizierungskriterien

Versuchspersonen: 45 KundInnen und 11 Fachkräfte

Methode: Hierarchisches Sortieren mit Befragung

Auswertung Verbaldaten: Häufigkeiten, Chi-Quadrat-Test

Auswertung Sortierdaten: Screenshot, Rangkorrelation, HCA, NMDS

Nachversuch: Property Fitting Analyse

Empirische Nachuntersuchung der Beurteilungsdimensionen der NMDS und HCA Ergebnisse

Versuchspersonen: 160 Personen

Methoden: Rangreihenbildung, Property Fitting

Nachbefragung:

Befragung eines Experten zur Interpretation der Sortierergebnisse

1 Fachkraft, offene Befragung, Pflanzenenzyklopädie

Abbildung 7.3-1. Untersuchungsplan Ermittlung der Begriffsstruktur von Zimmerpflanzen

Untersuchungsmaterial und Untersuchungsorte Für die Experimente wurden keine realen Pflanzen verwendet. Es wurden DIN A7 große Kärtchen mit naturgetreuen, farbigen Pflanzenfotos verwendet, Werbematerial der „Zentralen Marketing Organisation des Zierpflanzenanbaus“. An 5 Studenten wurde getestet, ob sie die Pflanzen wiedererkennen. Die Wiedererkennungslleistung war hundert Prozent. Die Untersuchungen fanden im Gartencenter in Grafing statt, einer Kleinstadt ca. 40 km von München entfernt, sowie im Dehner-Gartencenter in Waldtrudering, einem Vorort von München.

7.3.2 Vorversuch: Auswahl der Pflanzen für den Hauptversuch

Einleitung

Das durchschnittliche Zimmerpflanzen-Sortiment hat je nach Saison und je nach Größe der Verkaufsfläche ca. 40-80 Zimmerpflanzen-Artikel. In Hinblick auf das Hierarchische Sortieren und die NMDS-Auswertung sollten max. 24 Pflanzen ausgewählt werden. Die Auswahl sollte stellvertretend für das Gesamtsortiment sein, um aus den Ergebnissen Schlussfolgerungen für das gesamte Sortiment ziehen zu können. Überdies sollte sichergestellt sein, dass die zu sortierenden Zimmerpflanzen den Versuchspersonen bekannt sind. Die Auswahl sollte somit sortiments- und kundInnen-bezogene Kriterien berücksichtigen. Im Vorversuch wurden daher die KundInnen und die Fachkräfte befragt.

Methode

Erhebungsmethoden und Versuchsplan Zur Befragung der KundInnen dienten offene und geschlossene Fragen. Zunächst wurden mittels dreier offener Fragen KundInnen nach ihren Kriterien befragt, die sie beim Kauf einer Pflanze anstellen. Dann wählten die Versuchspersonen aus 65 häufig verkauften Zimmerpflanzen (definiert durch die „Zentrale Marketingorganisation des Zierpflanzenanbaus“) diejenigen aus, welche ihnen bekannt sind. Aus diesen bekannten Zimmerpflanzen wurden dann wiederum diejenigen 24 Pflanzen ausgewählt, die auch den KundInnen-Kriterien aus den Fragen 1-3 entsprachen. Schließlich wurde ein Gärtnermeister gebeten, die aufgrund der Befragung getroffene

Auswahl zu überprüfen und zu korrigieren, um sicherzustellen, dass die Pflanzen repräsentativ für das Gesamtsortiment sind.

Versuchspersonen, Durchführung, Dokumentation Es wurden 8 KundInnen (3 Männer, 5 Frauen) im Alter von 30 bis 69 Jahren im Gartencenter Grafing befragt, kurz nachdem sie einen Einkauf getätigt hatten. Die Personen wurden mündlich befragt. Die Antworten wurden schriftlich notiert. Die Stichprobe ist klein, aber ausreichend für eine Erkundungsstudie (vgl. Bortz 1984).

Ergebnisse des Vorversuchs

Die Antworten der KundInnen wurden qualitativ (Inhaltsanalyse) und quantitativ (Häufigkeitsanalyse) ausgewertet. Tabelle 7.3-1 zeigt die Ergebnisse der Fragen 1-3. Die Versuchspersonen konnten Mehrfachnennungen abgeben.

Die Antworten der drei Fragen lassen sich wie folgt zusammenfassen: Die Personen geben an mehr als die Hälfte der Zimmerpflanzen zum „Verschenken und Dankeschön sagen“ zu kaufen. Weniger als die Hälfte sagt, sie kaufe für das eigene Heim. Persönliches Gefallen und der Kaufpreis werden als besonders wichtig beim Kauf angegeben. Pflegerische Kriterien wie: Standortansprüche und Pflegeansprüche (Wasser, Düngen etc.) nannten die befragten KundInnen als unwichtigste Kriterien.

Bei der Frage: „Hier sehen Sie 65 Abbildungen von Pflanzen. Welche Pflanzen kennen Sie und welche sind Ihnen unbekannt?“ wurden die Pflanzen auf Bekanntheit beurteilt. Die befragten KundInnen mussten nur sagen, ob sie die Pflanzen kennen, sie mussten nicht den Namen nennen. Aus den bekannten Pflanzen wurden dann solche ausgewählt, die überdies den „wichtigen“ Kriterien der KundInnen entsprechen. Gemeinsam mit einem Gärtnermeister wurde diese Liste überprüft, damit sichergestellt war, dass die Auswahl auch vom botanischen Standpunkt (z.B. Pflanzenfamilie, Luftfeuchtigkeit etc.) repräsentativ für das Gesamtsortiment war.

Tabelle 7.3-1. Ergebnisse des Vorversuchs zur Auswahl der Zimmerpflanzen

Frage 1: „Zu welchem Anlass kaufen Sie Zimmerpflanzen?“	Absolute Nennungen		in %				
	Geschenk, Dankeschön	14	65%				
	Pers. Zimmerschmuck	6	26%				
	Neues ins Haus holen	2	9%				
	insg. Nennungen	22	100%				
Frage 2: „Welche Überlegungen haben Sie beim Kauf einer Pflanze?“	Absolute Nennungen		in %				
	Persönliches Gefallen	8	26%				
	Preis	7	23%				
	Lebensdauer der Pflanze	5	16%				
	Geplanter Standort	4	14%				
	Farbe	2	7%				
	Repräsentativität	2	7%				
	Pflegeansprüche	2	7%				
insg. Nennungen	30	100%					
Frage 3: „Welche der Kriterien sind Ihnen beim Pflanzenkauf am wichtigsten?“	Benotung: sehr wichtig --- sehr unwichtig						ungewichtete Summe der Nennungen
	Absolute Nennungen						
	1	2	3	4	5	6	
Aussehen	7	1					8
Persönl. Bekanntheitsgrad	6	1	1				8
Wuchsform	3	4	1				8
Kaufanlass	4	1	1	2			8
Standortansprüche	2	3	1	3			9
Pflegeansprüche	1		4	2	1		8

Tabelle 7.3-2 zeigt, welche Zimmerpflanzen ausgewählt wurden.

Tabelle 7.3-2. Liste der 24 repräsentativen und bekannten Pflanzen für den Hauptversuch

Nr Pflanze	Nr Pflanze	Nr Pflanze	Nr Pflanze
1 Araucaria	7 Bougainvillea	13 Solanum	19 Passiflora
2 Usambara	8 Yucca	14 Kaktus	20 Orchidee
3 Cyclamen	9 Farn	15 Gummibaum	21 Efeu
4 Stephanotis	10 Azalee	16 Dieffenbachia	22 Monstera
5 Kamelie	11 Phönix	17 Hibiskus	23 Philodendron
6 Aechmea	12 Kokospalme	18 Ficus Benjamin	24 Croton

Interpretation und Diskussion der Ergebnisse des Vorversuchs

Es erwies sich als aufwändig, die Auswahl der Pflanzen hinsichtlich der beiden Kriterien „Repräsentativität für das Sortiment“ und „Bekanntheit der Objekte für die KundInnen“ – zu kontrollieren. Der Vorversuch lieferte – neben einer für das Gesamtsortiment repräsentativen und für die KundInnen bekannten Auswahl von 24 Zimmerpflanzen – erste, vorsichtig zu interpretierende Erkenntnisse über Beurteilungskriterien von Zimmerpflanzen: Die befragten KundInnen geben an, Zimmerpflanzen für sich selbst zu Hause zu kaufen, oder um sie zu verschenken. Ihre Beurteilungskriterien beziehen sich auf Kriterien zum „Schenken“ und „Daheim aufstellen & genießen“ sowie auf die Kriterien Kaufpreis, persönliches Gefallen, Aussehen. Pflege- und Standortkriterien zeigen sich in der Befragung als die unwichtigsten Kriterien. Dies ist verwunderlich, da die Standortansprüche der Pflanze wesentlich sind, damit die Pflanze gut gedeiht und die Pflegeanforderungen die weiteren Aktivitäten mit der Pflanze bestimmen. Die Ergebnisse wurden nur an 8 Versuchspersonen gewonnen und sind daher nicht als repräsentativ anzusehen.

7.3.3 Hauptversuch

Methode

Erhebungsmethode Die semantisch-begriffliche Ähnlichkeit wurde mittels der Methode des Hierarchischen Sortierens (Marx & Hejj 1989) erhoben (siehe Abb. 7.3-2).

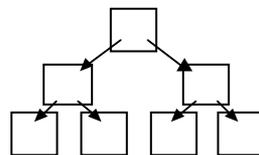


Abbildung 7.3-2. Veranschaulichung der Sortierschritte beim Hierarchischen Sortieren.

Das Hierarchische Sortieren eignet sich für die Erfassung von feinen Wissensunterschieden, wie es für den Vergleich der beiden Versuchspersonengruppen KundInnen und Fachkräfte notwendig ist (vgl. Marx & Hejj 1989, Hejj & Strube 1988). Die Versuchsperson wird gebeten, die Zimmerpflanzenfotos zunächst in zwei Gruppen zu teilen, diese wiederum in zwei Gruppen usw. bis jedes für sich alleine steht. Zusätzlich

wurde während des Hierarchischen Sortierens jede Versuchsperson nach jedem Sortierschritt nach ihrem Sortierkriterium gefragt.

Versuchsdurchführung und die Instruktion als eine unabhängige Variable Die KundInnen wurden im Gartencenter, nachdem sie ihren Einkauf getätigt hatten, angesprochen. Die Bereitschaft, an einem Experiment teilzunehmen, gilt als Verzerrung der Ergebnisse und somit als Beeinträchtigung der internen Validität. Doch die Ziehung einer echten Zufallsstichprobe wäre zu aufwändig gewesen. Wenn sich eine Kundin bereit erklärte, an der Untersuchung teilzunehmen, wurde sie in eine beruhigte Zone geführt, fernab der Zimmerpflanzen-Präsentation. Mit den Fachkräften wurden Termine an der Arbeitsstelle vereinbart. Die Versuchsperson wurden an einen Tisch geführt und erhielten folgende Instruktion:

Die **Instruktion ohne Zielorientierung** für die beiden Stichproben – Stichprobe 1 (KundInnen) und die Stichprobe 3 (Fachkräfte) – lautete: *„Herzlichen Dank, dass Sie mitmachen. Sie sehen einen Stoß Karten vor sich. Jede Karte zeigt die Abbildung einer Pflanze, wie Sie sie in jedem Blumengeschäft finden können. Ich möchte Sie nun bitten, diese Pflanzen so in zwei Gruppen zu unterteilen, wie Sie sie gerne in einem Gartencenter vorfinden würden. Es gibt keine richtigen und falschen Lösungen, mich interessiert nur, welche Pflanzen Sie zu einer Gruppe zusammenfassen.“*

Die **Instruktion mit Zielorientierung** für die Stichprobe 2 (KundInnen) war bis auf einen Satz identisch und lautete *„Herzlichen Dank, dass Sie mitmachen. Sie sehen einen Stoß Karten vor sich. Jede Karte zeigt die Abbildung einer Pflanze, wie Sie sie in jedem Blumengeschäft finden können. **Bitte stellen Sie sich nun vor, Sie stehen im Gartencenter und möchten eine Pflanze kaufen.** Ich möchte Sie nun bitten, diese Pflanzen so in zwei Gruppen zu unterteilen, wie Sie sie gerne in einem Gartencenter vorfinden würden. Es gibt keine richtigen und falschen Lösungen, mich interessiert nur, welche Pflanzen Sie zu einer Gruppe zusammenfassen.“*

Nach jedem Sortierschritt wurde die Versuchsperson gefragt, nach welchem Kriterium sie ihre Wahl getroffen hat. Nachdem die Versuchsperson dies gemacht hat, wurde sie gebeten: *“Bitte trennen Sie jede dieser zwei Gruppen wieder in zwei Gruppen, so*

dass Sie vier Gruppen haben.“ Die Instruktion zum weiteren Trennen der Gruppen und anschließendem Nennen des Sortierkriteriums erfolgte so lange, bis alle Produkte einzeln auf dem Tisch lagen. Auf Zwischenfragen der Versuchsperson, wonach sie die Artikel trennen soll, antwortete die Versuchsleiterin: *“Gehen Sie nach einem Ihnen sinnvoll erscheinenden Kriterium vor.”* oder *“Das bleibt Ihnen überlassen.”* Die Ergebnisse des Hierarchischen Sortierens und der Kriteriumsennung wurden in einem Protokollbogen notiert.

Auswahl der Versuchspersonen Insgesamt wurden 56 Personen untersucht: 45 KundInnen sowie 11 Fachkräfte, d.h. Gärtner und Zimmerpflanzen-VerkäuferInnen (Tab. 7.3-3). Der hohe Frauenanteil in der Stichprobe ergab sich zufällig und ist gerechtfertigt, da die KäuferInnen bzw. Fachkräfte in der Mehrzahl Frauen sind.

Tabelle 7.3-3. Versuchspersonen für den Zimmerpflanzen-Hauptversuch

Stichproben		Männer	Frauen	Gesamt
1. KundInnen <u>ohne</u> Zielorientierung		6	25	31
2. KundInnen <u>mit Kauf-</u> Zielorientierung		0	14	14
Alter	unter 20	0	2	2
	20 – 29	1	9	10
	30 – 39	2	7	9
	40 – 49	2	8	10
	50 – 59	0	8	8
	60 – 69	1	4	5
	70 und älter	0	1	1
3. Fachkräfte <u>ohne</u> Zielorientierung		4	7	11
Alter	unter 20	0	0	0
	20 – 29	2	2	4
	30 – 39	1	1	2
	40 – 49	0	1	1
	50 und älter	1	3	4
		0	0	0
insgesamt		10	46	56

Ergebnisse

Auswertungsmethoden Die Verbaldaten (=Nennungen der Sortierkriterien) der KundInnen und Fachkräfte wurden nach Häufigkeiten ausgewertet und mit dem Chi-Quadrat-Test (Siegel 1976) auf Unterschiede geprüft. Für die Verhaltensdaten (=Sortiererergebnisse) des Hierarchischen Sortierens wurde eine Distanzmatrix erstellt. Die Daten wurden mit der Nonmetrischen Multidimensionalen Skalierung (NMDS) ausgewertet (vgl. Gigerenzer 1981, Borg 1981). Zur Bestimmung der Güte der NMDS-Lösung wurde mittels des Scree-tests der Stresswert (=Alienation) berechnet. Zur Berechnung der NMDS wurde das Programm Minissa in PC-Version von Markus Gampferling (1989) verwendet. Die Daten der Distanzmatrix wurden auch mit der Hierarchischen Clusteranalyse nach Johnson (1967) mit dem Programm von Strube (1983; Hierarchical Clustering Schemes HCS) ausgewertet. Des Weiteren wurden anhand der Distanzmatrixen die Rangkorrelationskoeffizienten von Spearman berechnet.

Als Signifikanzniveau wurde bei allen Auswertungsverfahren das 5% Niveau festgelegt. Dieses Signifikanzniveau ist angesichts der Stichprobe (56 Vp) als genügend hoch zu bewerten, um sicherzugehen, dass die Nullhypothese irrtümlich abgelehnt wird. Ein irrtümliches Beibehalten der Nullhypothese (β -Fehler), d.h. dass es keinen Unterschied zwischen KundInnen und Fachkräften gibt und damit ein Beibehalten der vorhandenen Pflanzenanordnung, würde dem Unternehmer mehr Nachteile (bzgl. Kundenzufriedenheit, Orientierung, eventuell Abverkauf, Umsatz) bringen, als ein irrtümliches Ablehnen.

Ergebnisse der verbalen Nennungen Die Auswertung der Sortierkriterien, die nach jedem Sortierschritt erfragt wurden, erfolgte quantitativ und qualitativ. Dabei wurde jede Nennung pro Versuchsperson nur einmal gewertet, auch wenn eine Person ein Kriterium mehrmals nannte. Für jedes Sortierkriterium wurden die absoluten und relativen Häufigkeiten sowie prozentuale Anteile berechnet. Die Kriterien wurden dann in vier Kategorien unterteilt. Die Kategorisierung wurde in Anlehnung an die Merkmalsklassifikation von Hoffmann (1986 S. 71 f.) vorgenommen (vgl. Kapitel 4):

- *Sensorische Kriterien* repräsentieren *anschauliche Merkmale*, die man sehen, riechen, fühlen etc. kann; z.B. Blattfarbe, Blattform, Blütengröße, Früchte, Geruch, Stacheligkeit, Blühhäufigkeit.

- *Ereignisbezogene Kriterien* „Pfleger einer Pflanze“ repräsentieren Kriterien, die sich auf die Pflege der Pflanzen beziehen: z.B. Robustheit, Wasser-, Wärme-, Lichtbedarf, Wuchshöhe.
- *Ereignisbezogene Kriterien* „Auswahl einer Pflanze“ repräsentieren Kaufkriterien, wie Besonderheit (Standard-Exotik), aber auch floristische Verwendung;
- *Hierarchisch-kategoriale Kriterien* beziehen sich auf Ober-, Unter- bzw. Nebenbegriffe, wie die Pflanzenfamilie, Herkunft.

Tabelle 7.3-4 zeigt die Nennungen der Stichproben für die jeweilige Merkmalskategorie (absolute und prozentuale Nennungen).

Tabelle 7.3-4. Absolute & prozentuale Nennungen für die jeweilige Merkmalskategorie

	KundInnen				Fachkräfte	
	ohne Kauforientierung		mit Kauforientierung		Gärtner/ VerkäuferInnen	
	absolut	%	absolut	%	absolut	%
o Sensorische Kriterien	105	46,8%	36	36,4%	24	28,6%
o Pflege-Ereigniskriterien	62	27,5%	35	35,3%	41	48,8%
o Auswahl- Ereigniskriterien	24	10,6%	20	20,2%	11	13,1%
o Kategoriale Kriterien	34	15,1%	8	8,1%	8	9,5%
insgesamt	225	100%	99	100%	84	100%

Bei der Berechnung des Chi²-Tests zeigen sich signifikante Unterschiede auf dem 5% Signifikanz-Niveau zwischen den KundInnen-Stichproben mit und ohne Zielorientierung. $\chi^2_{\text{berechnet}} = 10,98$ $\chi^2_{\text{krit (Df=3)}} = 7,81$. Zwischen Fachkräften und den KundInnen zeigt sich sogar ein signifikanter Unterschied auf dem 0,5% Signifikanz-Niveau $\chi^2_{\text{berechnet}} = 13,13$ $\chi^2_{\text{krit (Df=3)}} = 12,83$ (vgl. Bortz, S. 839). Die Zielorientierung „Bitte stellen Sie sich vor, Sie möchten eine Pflanze kaufen.“ hat somit eine Wirkung darauf, welche Sortierkriterien die KundInnen beim Sortieren nennen. Die KundInnen, die aufgefordert werden die Pflanzen im Kontext eines Pflanzenkaufs zu sortieren, beschreiben ihre Sortierschritte häufiger mit Pflege- und Auswahl-Kriterien und weniger mit Kriterien bezüglich Aus-

sehen oder Pflanzenfamilie als jene KundInnen, die nur gebeten werden, die Pflanzenab- bildungen zu sortieren. Auch die unterschiedliche Handlungserfahrung hat eine signifi- kante Wirkung auf die Nennung der Sortierkriterien. Fachkräfte benennen ihre Sortier- schritte mehr nach Kriterien bezüglich Pflege und weniger bezüglich Aussehen, Auswahl und kategorialer Familienzugehörigkeit.

Ergebnisse des Sortierverhaltens Die Sortiererergebnisse des Hierarchischen Sortierens wurden mit der NMDS und der HCA ausgewertet. Hierfür wurden in einem *ersten Schritt* mittels des *Scree-tests* die *Stresswerte* für die NMDS-Konfigurationen berechnet. Für alle Versuchspersonengruppen liefert der Scree-test bereits für die 2-dimensionalen NMDS-Konfigurationen Stresswerte, die auf dem 1% Niveau signifikant sind (vgl. Gigerenzer 1981, S. 353): Die Stresswerte der Stichproben sind kleiner als der kritische Stresswert von = 0.306 für 24 Objekte für 2 Dimensionen auf dem 1% Signifikanzni- veau, wie Tabelle 7.3-5 zeigt. Das festgelegte Signifikanzniveau von 5% ist unterschrit- ten und die Gefahr einer irrtümlichen Zurückweisung der Nullhypothese minimiert.

Tabelle 7.3-5. Stresswerte

Stresswerte für die 2-dimensionalen NMDS-Konfigurationen		
Stichprobe 1: KundInnen <i>ohne</i> Zielorientierung	0.073 *	
Stichprobe 2: KundInnen <i>mit</i> Zielorientierung	0.147 *	
Stichprobe 3: Fachkräfte	0.061 *	* signifikant 1% Niveau
*Stresswert ist nach der Tabelle von Gigerenzer (1981, S.353) als zufallskritisch abgesichert, da er für 24 Objekte kleiner als der kritische Stresswert von = 0.306 ist.		

Die Signifikanz der Stresswerte zeigt an, dass die begrifflich-semantiche Ähnlichkeits- beurteilung eine personenübergreifende stabile Struktur ergibt. Die Personen einer Gruppe beurteilen die Zimmerpflanzen nicht nach individuell unterschiedlichen Kriteri- en, sondern es gibt überzufällig konsistente Kriterien für die Beurteilung der Zimmer- pflanzen innerhalb einer Gruppe. Dabei zeigt die Gruppe der Fachkräfte mit einem Stresswert von 0.061 die größte Übereinstimmung in der Ähnlichkeitsbeurteilung.

In einem *zweiten Schritt* wurde die *Rangkorrelation* zwischen den Gruppen berechnet, um zu prüfen, wie stark die Sortierergebnisse der drei Gruppen miteinander korrelieren bzw. sich unterscheiden (Tab. 7.3-6). Hierfür wurden die empirischen Distanzen mit der Rangkorrelation nach Spearman (vgl. Siegel 1976, Bortz 1984) berechnet. Die Rangkorrelation nach Spearman eignet sich als Korrelationsmaß für ordinalskalierte Merkmale.

Tabelle 7.3-6. Rangkorrelationen

Rangkorrelationen nach Spearman	
zwischen den Gruppen	
- KundInnen <u>ohne</u> Zielorientierung (Ko)	
- KundInnen <u>mit</u> Zielorientierung (Km)	$r_{Ko/Km} = 0.88^*$ $r_{\text{Best-Ma\ss}} = 0.78^*$
zwischen den Gruppen	
- KundInnen <u>ohne</u> Zielorientierung (Ko)	
- Fachkräfte (Fach)	$r_{Ko/Fach} = 0.84^*$ $r_{\text{Best-Ma\ss}} = 0.70^*$
zwischen den Gruppen	
- KundInnen aus Grafing <u>ohne</u> Zielorientierung	
- KundInnen aus Waldtrudering <u>ohne</u> Zielorientierung	$r_{Ko/Km} = 0.91^*$ $r_{\text{Best-Ma\ss}} = 0.83^*$
* signifikant auf 1% Niveau	

Die Ergebnisse der Rangkorrelationen zeigen, dass die Gruppen eine sehr hohe Übereinstimmung in der Ähnlichkeitsbeurteilung haben: Die Sortierergebnisse der KundInnen mit und ohne Zielorientierung korrelieren zu 78%, sie unterscheiden sich somit nur zu 22%. Die Sortierergebnisse der KundInnen und der Fachkräfte korrelieren zu 70% überein, sie unterscheiden sich daher zu 30%.

In einem *dritten Schritt* wurde die *Hierarchische Clusteranalyse* (HCA) nach Johnson für alle drei Stichproben berechnet. Tabelle 7.3-7 gibt einen Überblick über die Ergebnisse der Hierarchischen Clusteranalyse. Alle drei Stichproben stimmen darin überein, dass sie strikt zwischen Blüh- und Grünpflanzen trennen. Die *KundInnen mit und ohne Zielorientierung* unterscheiden sich auch innerhalb der Blüh- und Grünpflanzen kaum. Bei den Blühpflanzen differenziert mal die eine, mal die andere Gruppe feiner. Bei den Grünpflanzen-Cluster unterscheiden sich die Gruppen in der Beurteilung der Yucca. Ansonsten stimmen die Cluster weitgehend überein. Auf diesen geringen Unterschied

hat bereits der hohe Rangkorrelationskoeffizient von $r_{\text{Best-Ma\ss}} = .78$ hingedeutet. Die *Fachkräfte* unterscheiden auch strikt zwischen Blüh- und Grünpflanzen. Bei den Blühpflanzen gibt es aber kein Subcluster, das mit denen der KundInnen identisch ist. Bei den Grünpflanzen gibt es mehr Übereinstimmung mit den KundInnen. Auf die Gründe der Ähnlichkeitsbeurteilung wird in der Interpretation eingegangen.

Tabelle 7.3-7. Ergebnisse der Hierarchischen Clusteranalyse für die drei Stichproben

Subcluster der Clusteranalyse für die 3 Stichproben		
1. KundInnen ohne Zielorientierung	2. KundInnen mit Zielorientierung	3. Experten
Blühpflanzen	Blühpflanzen	Blühpflanzen
Usambara, Cyclamen, Azalee	Usambara, Cyclamen, Azalee	Usambara, Cyclamen, Azalee Solanum
Aechmea, Kaktus, Solanum	Aechmea, Kaktus, Solanum	Kamelie, Kaktus, Hibiskus
Kamelie, Bougainvillea, Hibiskus	- Kamelie, Bougainvillea - Hibiskus	- Bougainvillea, Stephanotis - Passionsblume
- Passionsbl., Stephanotis - Orchidee	Passionsblume, Stephanotis, Orchidee	Aechmea, Orchidee
Grünpflanzen	Grünpflanzen	Grünpflanzen
Croton, Dieffenbachia	Croton, Dieffenbachia	- Croton, Dieffenbachia, - Efeu
Monstera, Philodendron, Gummibaum	Monstera, Philodendron, Gummibaum	Monstera, Philodendron, Gummibaum
Efeu, Benjamin	Yucca , Efeu, Benjamin	Yucca
Araucaria, Farn	Araucaria, Farn	Araucaria, Farn, Benjamin
- Phönix, Kokospalme - Yucca	Phönix, Kokospalme	Phönix, Kokospalme

Legende: Die Grauschattierung und der Doppelpfeil \leftrightarrow veranschaulichen die Übereinstimmung der Cluster. Fettmarkierung zeigt Unterschiede. Die Reihenfolge der Cluster hat keine Bedeutung.

In einem *vierten Schritt* wurden die 2-dimensionalen *NMDS-Konfigurationen* für die Stichproben berechnet. Die NMDS-Lösung der Fachkräfte stellte eine „degenerierte Lösung“ dar (vgl. Gigerenzer 1981, S. 346): Die Pflanzen wurden in wenige kompakte Cluster organisiert, weil die Anzahl gleich großer Distanzen relativ zur Gesamtzahl der Distanzen zu groß ist. Deshalb wird der von Gigerenzer empfohlene Weg beschritten und zwei getrennte NMDS für das Grobcluster Blüh- und Grünpflanzen berechnet. Das Gleiche wird bei den KundInnen gemacht, um sie mit den Fachkräften besser vergleichen zu können. Aufgrund der hohen Übereinstimmung von KundInnen *mit* und *ohne* Zielorientierung wurden die KundInnen *mit* Zielorientierung nicht weiter verfolgt. Es wurden nur die Ergebnisse der KundInnen *ohne* Zielorientierung getrennt berechnet.

Die Abbildungen 7.3-3 und 7.3-4 zeigen die NMDS-Lösungen mit den integrierten HCA Ergebnissen von Fachkräften und KundInnen für die *Blühpflanzen*. (Die Punkte stellen die Pflanzen dar. Sie sind beschriftet. Die Lage der Punkte ist Ergebnis der NMDS-Berechnung. Die Kreise bzw. Ovale sind die integrierten Ergebnisse der HCA.)

Fachkräfte Blühpflanzen NMDS & HCA

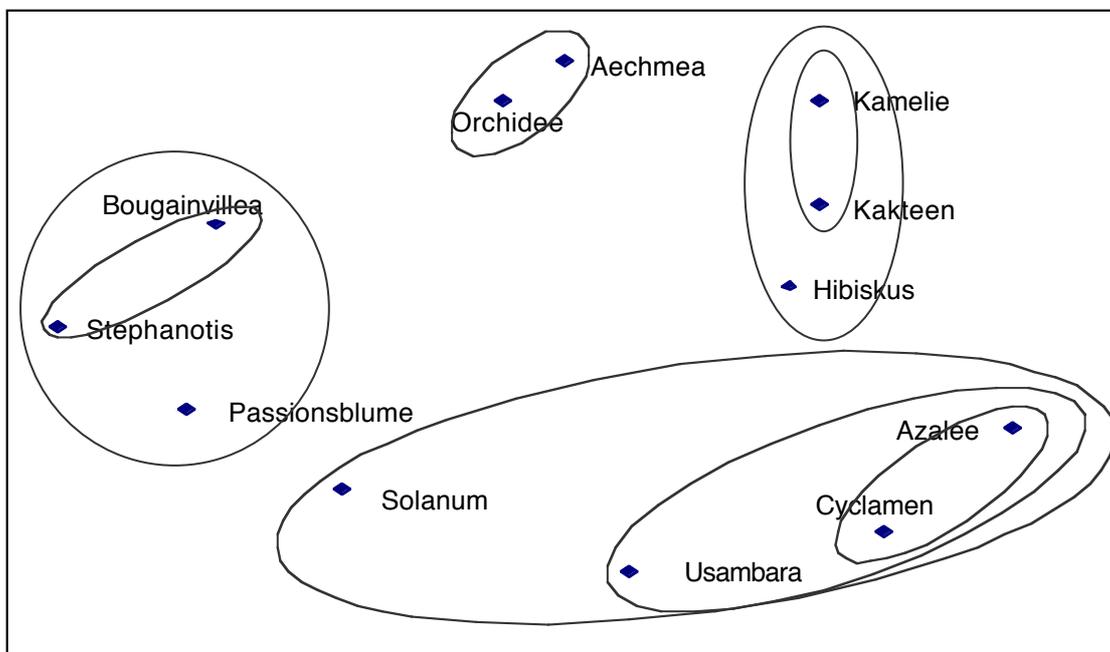


Abbildung 7.3-3. NMDS-Konfiguration mit integrierten HCA-Ergebnissen (Umkreisungen) für die Stichprobe 3: **Fachkräfte**, nur **Blühpflanzen**

KundInnen Blühpflanzen NMDS & HCA

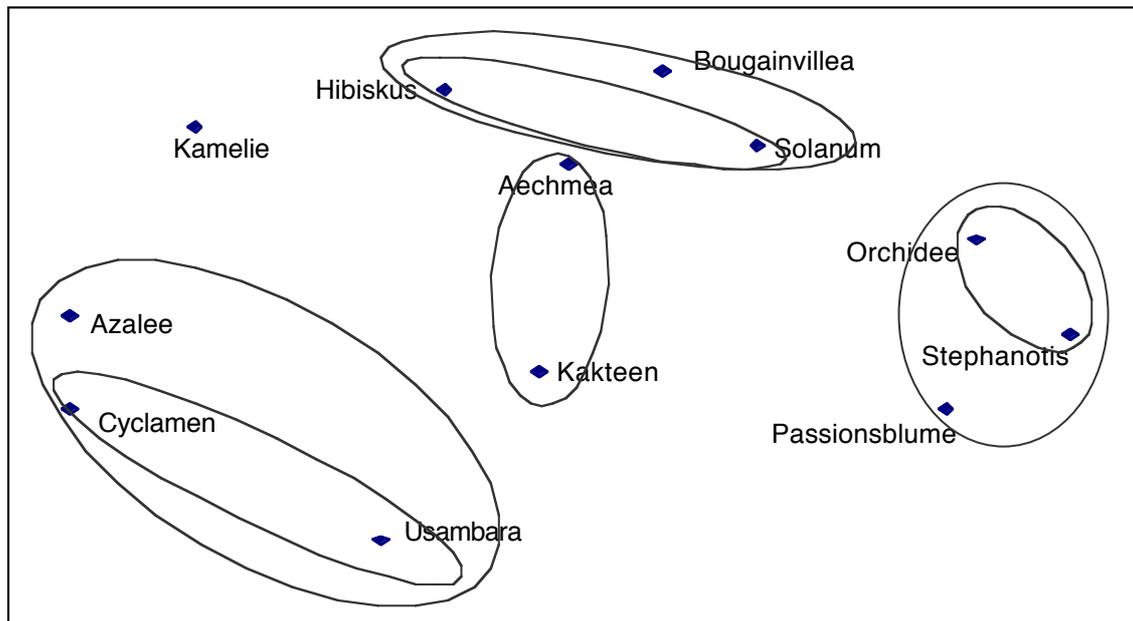


Abbildung 7.3-4. NMDS-Konfiguration mit integrierten HCA-Ergebnissen (Umkreisungen) für die Stichprobe 1: **KundInnen ohne Kauforientierung, nur Blühpflanzen**

Die Abbildungen zeigen anschaulich, dass sich die Fachkräfte und die KundInnen in ihren Sortierergebnissen unterscheiden. Fachkräfte beurteilen z.B. die beiden tropischen Epiphyten (Baumaufsitzer) Aechmea und Orchidee als ähnlich, während die KundInnen die Aechmea zu den Kakteen sortieren. Es gibt aber auch Übereinstimmungen, so z.B. die ähnliche Beurteilung von Azalee, Cyclamen (Alpenveilchen) und Usambara.

Die Abbildungen 7.3-5 und 7.3-6 zeigen die NMDS-Lösungen von Fachkräften und KundInnen für die *Grünpflanzen* mit den integrierten HCA-Ergebnissen. Sie zeigen anschaulich, dass sich die Fachkräfte und die KundInnen in ihrer Gruppenbildung (Kreise) kaum und in der räumlichen Zuordnung nur wenig unterscheiden. Fachkräfte beurteilen z.B. die Gruppe Farn und Araucaria dem Benjamin ähnlich und der Gruppe der Palmen (Kokos und Phoenix). Die KundInnen beurteilen die Gruppe Farn und Araucaria zwischen den Palmen und dem Efeu.

Fachkräfte Grünpflanzen NMDS & HCA

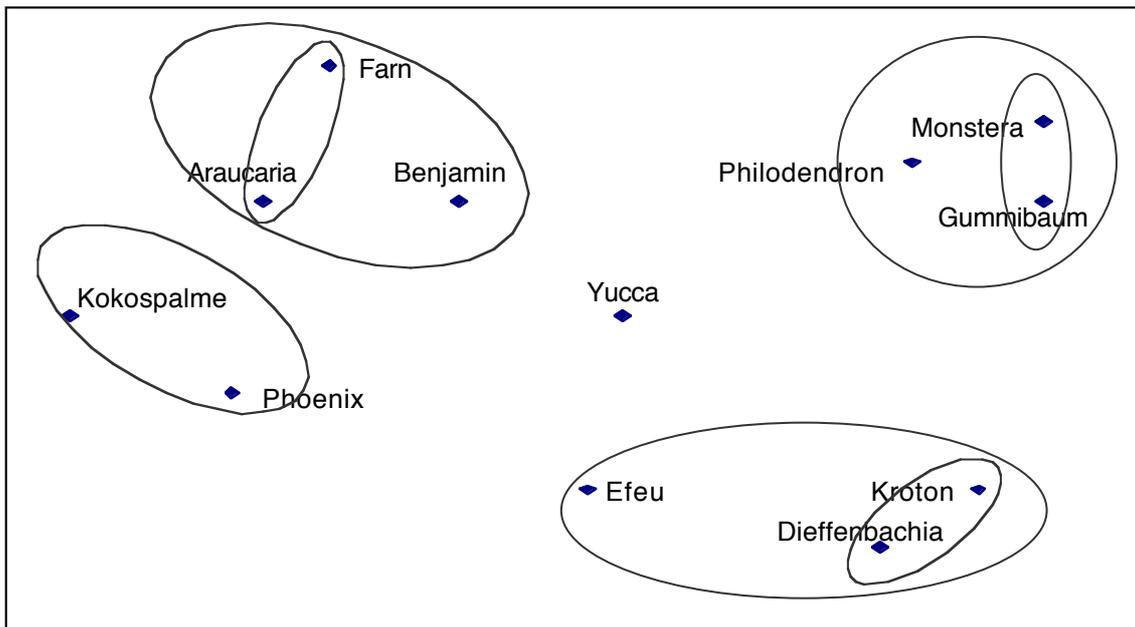


Abbildung 7.3-5. NMDS-Konfiguration mit integrierten HCA-Ergebnissen (Umkreisungen) für die Stichprobe 3: **Fachkräfte**, nur **Grünpflanzen**

KundInnen Grünpflanzen NMDS & HCA

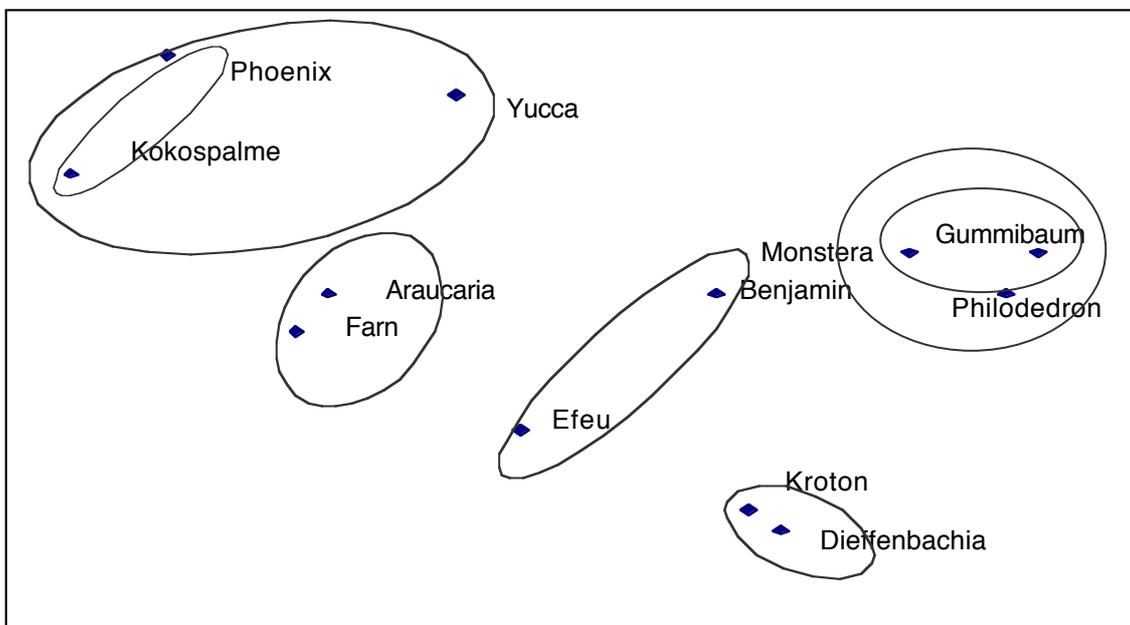


Abbildung 7.3-6. NMDS-Konfiguration mit integrierten HCA-Ergebnissen (Umkreisungen) für die Stichprobe 1: **KundInnen ohne Kauforientierung**, nur **Grünpflanzen**

Lesehilfe: Die unterschiedliche Lage beider Konfigurationen im Koordinatensystem ist ohne Bedeutung. NMDS Konfigurationen sind ohne Koordinatensystem. (vgl. Borg1981)

7.3.4 Nachversuch: Property Fitting Analyse

Methode

Die Property Fitting Analyse wurde angewendet, um Erfahrungen mit einem empirisch gestützten Verfahren zur Interpretation der NMDS- und HCA-Ergebnisse zu sammeln. Je 60 Versuchsperson wurden gebeten die 12 Blühpflanzen oder die 12 Grünpflanzen des Hauptversuchs bezüglich der im Hauptversuch am häufigsten genannten Sortierkriterien in eine Rangreihe zu legen. Tabelle 7.3-8 zeigt die im Hauptversuch häufig genannten Sortierkriterien und beschreibt die Begriffe für das Rating.

Tabelle 7.3-8. Liste der häufigst genannten Sortierkriterien-Nennungen (aus Hauptversuch)

9 häufigst genannte Sortierkriterien für Blühpflanzen	f	$p = \frac{f}{N=56}$	Gegensatzpaare für das Rating
Lichtbedarf	25	0.45	dunkel hell
Robustheit	23	0.41	robust empfindlich
Besonderheit	18	0.32	standard exotisch
Blühhäufigkeit	17	0.30	selten häufig
Wasserbedarf	15	0.27	trocken nass
Blütengröße	12	0.21	klein groß
Blütenanzahl	5	0.09	wenig viel
Stacheligkeit	5	0.09	glatt stachelig
Wärmebedarf	6	0.11	kalt warm

8 häufig genannte Sortierkriterien für Grünpflanzen	f	$p = \frac{f}{N=56}$	Gegensatzpaare für das Rating
Wuchshöhe	32	0.57	klein groß
Lichtbedarf	25	0.44	dunkel hell
Robustheit	23	0.41	robust empfindlich
Wasserbedarf	15	0.26	trocken nass
Herkunft	15	0.26	Süden Norden
Wärmebedarf	6	0.11	kalt warm
Blattgröße	5	0.09	klein groß
Luftfeuchtigkeit	2	0.03	lufttrocken luftfeucht

Legende: f = absolute Häufigkeiten p= relative Häufigkeiten
 N= Versuchspersonenanzahl der 3 Stichproben aus dem Hauptversuch

Untersuchungsmaterial und Instruktion Für den Nachversuch dienten die Pflanzenab- bildungen aus dem Hauptversuch. Die Instruktion lautete: „Bitte nehmen Sie den Stoß Karten. Jede Karte zeigt eine Pflanze. Ich bitte Sie die Pflanzen < z.B. „der Wuchshöhe“ > nach zu ordnen. Fangen Sie bitte mit derjenigen Pflanze an, die Ihrer Meinung nach am < kleinsten > bleibt und legen Sie dann die Pflanzen der < Wuchshöhe > nach aneinan- der, bis zu der Pflanze, die am < größten > wird.“ Die Rangreihe wurde vom Versuchslei- ter notiert. Die Versuchspersonen konnten die Rangreihen ohne Mühe bilden.

Versuchspersonen und Versuchsdurchführung Es wurden insgesamt 153 KundInnen (31 Männer, 122 Frauen) und 7 Fachkräfte (1 Mann, 6 Frauen) untersucht (Tab. 7.3-9) Die Versuchspersonen wurden im Gartencenter Grafing angesprochen.

Tabelle 7.3-9. Versuchspersonen des Property Fitting

Personenmerkmale	Männer	Frauen	Summe
Stichprobe 1: KundInnen	31	122	153
Stichprobe 2: Fachkräfte	1	6	7
insgesamt	32	128	160
Alter			
unter 20	8	14	22
20 – 29	5	26	31
30 – 39	7	42	49
40 – 49	1	19	20
50 – 59	4	11	15
60 – 69	5	13	18
70 und älter	2	3	5

Auswertung Die Auswertung der Ergebnisse zu den 17 Rangreihen (9 für Blüh- und 8 für Grünpflanzen) erfolgte über die Berechnung der mittleren Tendenz; d.h. die arithme- tische Mittelwerte über die Versuchspersonen werden dabei in eine ganzzahlige Rang- ordnung überführt. Mittels des Property Fitting werden diese Ergebnisse in die NMDS- Konfigurationen eingepasst, indem für jede Rangreihe die Regressionsgerade in der NMDS-Konfiguration gesucht wird. Das Property Fitting rotiert die einzelnen Rangrei- hen so lange, bis eine maximale Korrelation mit den Skalenwerten der in der NMDS- Konfiguration abgebildeten Pflanzen erreicht ist. Die Pflanzen in der NMDS- Konfiguration werden auf diese Gerade projiziert und ihre Reihenfolge wird mit der

Rangreihe korreliert. Dieser Korrelationskoeffizient wird durch den Vergleich mit dem kritischen Wert aus der Tabelle von Gaensslen & Schubö (1976, S. 314) verglichen und somit auf seine statistische Signifikanz geprüft. Dimensionen, die eine hohe Korrelation mit der NMDS-Konfiguration aufweisen, haben für das Zustandekommen der Ähnlichkeitsurteile mehr Aussagekraft als diejenigen geringerer Korrelation (Marx 1982, S. 143).

Ergebnisse

Das Property Fitting wurde für KundInnen und Fachkräfte jeweils für Blüh- und Grünpflanzen berechnet. Die kritische, multiple Korrelation zwischen der Beurteilung der Pflanzen (Rangreihen des Property Fitting) und den Skalenwerten der Pflanzen in der NMDS-Konfiguration liegt bei $r_{krit} = .506$ auf dem 5% Signifikanz-Niveau. Tabelle 7.3-10 zeigt die Korrelationswerte für die Blüh- und Grünpflanzen.

Tabelle 7.3-10. Korrelationswerte für die Dimensionen des Property Fitting (geordnet nach der Höhe der Korrelationen)

9 Dimensionen für Blühpflanzen		8 Dimensionen für Grünpflanzen	
KundInnen	Fachkräfte	KundInnen	Fachkräfte
Besonderheit .73*	Blühhäufigkeit .79*	Wuchshöhe .86*	Wasserbedarf .81*
Wasserbedarf .65*	Besonderheit .75*	Wasserbedarf .85*	Blattgröße .70*
Wärmebedarf .64*	Wärmebedarf .72*	Herkunft .68*	Luftfeuchte .66*
Blühhäufigkeit .61*	Blütengröße .71*	Blattgröße .65*	Robustheit .53*
Lichtbedarf .58*	Wasserbedarf .59*	Luftfeuchte .63*	Wuchshöhe .49
Blütengröße .41	Blütenanzahl .57*	Lichtbedarf .61*	Wärmebedarf .31
Stacheligkeit .35	Lichtbedarf .39	Robustheit .27	Lichtbedarf .28
Robustheit .32	Stacheligkeit .36	Wärmebedarf .25	Herkunft .07
Blütenanzahl .31	Robustheit .25		

* sign. Korrelationen $r_{krit} = .506$ für 5% Signifikanz-Niveau (Gaensslen & Schubö 1976)

Abbildungen 7.3-7 und 7.3-8 zeigen die NMDS-Konfigurationen mit den integrierten Property Fitting Dimensionen für die Stichproben *KundInnen* (5 Dimensionen) und

Fachkräfte (6 Dimensionen) für die Blühpflanzen. Es wurden nur die Dimensionen mit hohen Korrelationen eingetragen.

Für die KundInnen lässt sich feststellen: die Ordnungsdimension Besonderheit (.73) mit den Polen „standard und exotisch“ lässt sich bei den Blühpflanzen am Besten einpassen. Die Dimension Besonderheit korreliert hoch mit den Dimensionen Blühhäufigkeit ($r = 0.96$), Wasserbedarf ($r = 0.94$) und Wärmebedarf ($r = 0.94$), sodass diese wenig zur Differenzierung des Begriffsfeldes beitragen. Die Dimension Lichtbedarf (.58) mit den Polen hell und dunkel lässt sich am schlechtesten einpassen. Aus botanischen Gesichtspunkten stimmen jedoch die Dimensionen Wasser-, Wärme und Lichtbedarf nicht! Einige Pflanzen haben andere Pflegeanforderungen als hier von den KundInnen beurteilt! Die Cyclame und die rankende Stephanotis brauchen viel Licht, die Aechmea aus den halbschattigen oberen Etagen des Regenwaldes braucht tatsächlich weniger Licht als die Kakteen. Die KundInnen beurteilten die Pflanzen aber anders!

Für die Fachkräfte zeigt das Property Fitting: bei den Blühpflanzen lassen sich Blühhäufigkeit (.79), Besonderheit (.75), Wärmebedarf (.72) und Blütengröße (.71) gut einpassen. Betrachtet man diese Ergebnisse auf dem Hintergrund der botanischen Eigenheiten der Pflanzen, dann stimmen die Dimensionen Besonderheit, Wärmebedarf und Blütengröße noch ungefähr. Die anderen Dimensionen (Wasserbedarf, Blütenanzahl) stimmen aus botanischer Sicht nicht! Obwohl dies die Ergebnisse der Fachkräfte sind, die ja eigentlich die Pflanzen aus botanischer Sicht richtig beurteilen müssten! Das Usambaraveilchen z.B. braucht wenig Wasser, liegt aber bei dem Pol „viel Wasser“. Die Kakteen brauchen am wenigsten Wasser, liegen nicht am Pol „trocken“. In der Interpretation werden diese Widersprüche erörtert.

Die Ergebnisse für die Grünpflanzen sind genauso problematisch! Sie stimmen aus botanischen Gesichtspunkten auch nicht. Auf die Darstellung der Ergebnisse für die Grünpflanzen wird daher ganz verzichtet.

KundInnen Blühpflanzen NMDS & HCA

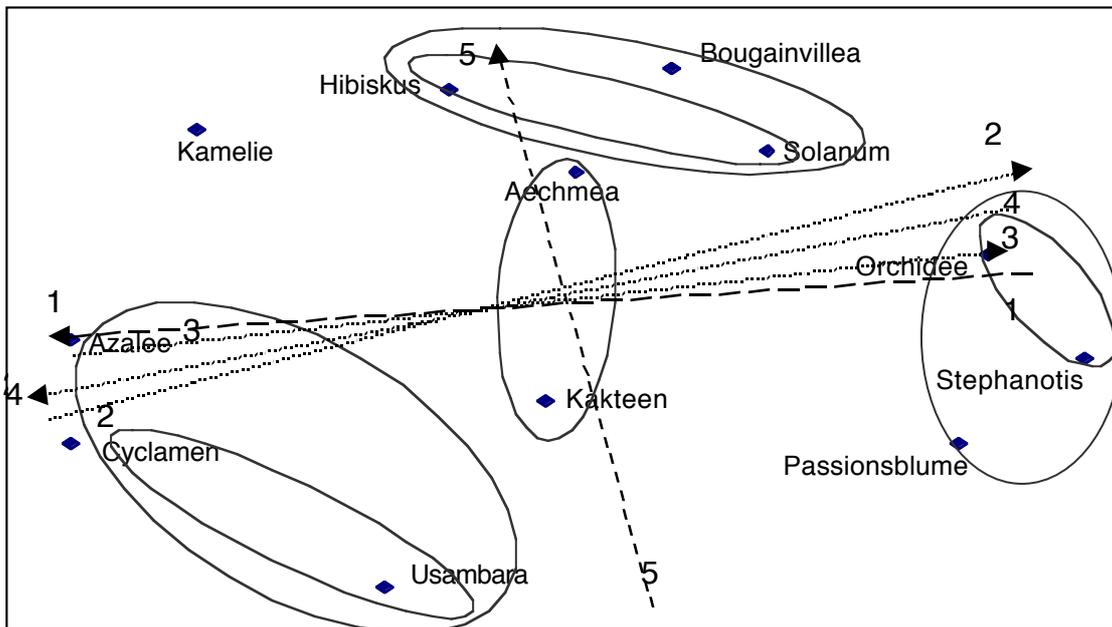


Abbildung 7.3-7. NMDS-Konfiguration mit integrierten HCA (Umkreisungen) und Property Fitting (Pfeile) für die Stichprobe **KundInnen, Blühpflanzen**; Dimensionen: 1 = Blühhäufigkeit, 2 = Besonderheit, 3 = Wärmebedarf, 4 = Wasserbedarf, 5 = Lichtbedarf; Pfeilende zeigt jeweils die größere Ausprägung

Fachkräfte Blühpflanzen NMDS & HCA

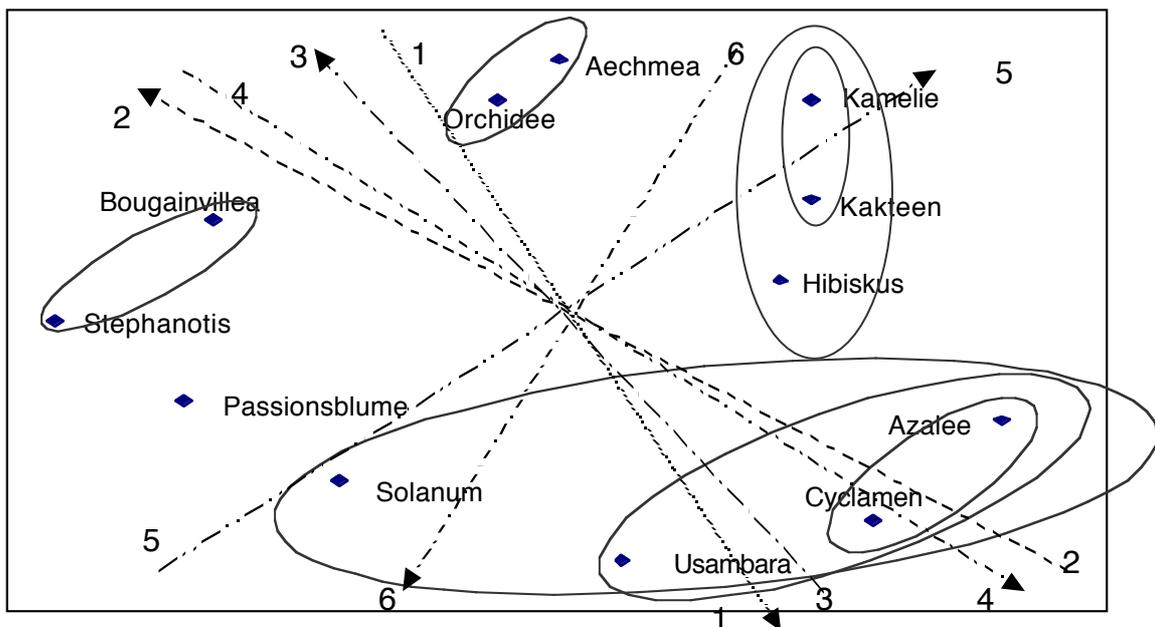


Abbildung 7.3-8. NMDS-Konfiguration mit integrierten HCA (Umkreisungen) und Property Fitting (Pfeile) für die Stichprobe **Fachkräfte, Blühpflanzen**; Dimensionen: 1= Blühhäufigkeit, 2= Besonderheit, 3= Wärmebedarf, 4=Wasserbedarf, 5= Blütengröße, 6= Blütenanzahl; Pfeilende zeigt jeweils die größere Ausprägung

Interpretation der Property Fitting-Ergebnisse

Die Ergebnisse des Property Fitting tragen nicht dazu bei, die Sortierergebnisse tiefergehend zu verstehen und gar zu erklären. Zum einen korrelieren die Dimensionen untereinander hoch, sodass ihr Erklärungswert sinkt. Zum anderen lassen sich die Dimensionen rechnerisch zwar gut einpassen, aber sie sind aus botanischer Sicht falsch. Der geringe Beitrag der Property Fitting Analyse zum tieferen Verständnis der Ähnlichkeitsbeurteilung lässt sich durch Folgendes erklären: Das Property Fitting arbeitet mit den verbalen Sortierkriterien-Nennungen aus dem Hauptversuch. Es zeigte sich aber bereits, dass die verbalen Nennungen nicht mit dem tatsächlichen Sortierverhalten übereinstimmen. Es wird vermutet, dass die Versuchspersonen beim des Sortierens oberflächliche Kriterien und nicht die wesentlichen Kriterien nennen: u. U. sind sie auch nicht in der Lage, ihre Kriterien zu verbalisieren. Aus diesem Grund wird Abstand von den Ergebnissen des Property Fitting genommen und es erfolgt eine zweite Nachuntersuchung.

7.3.5 Befragung und Literaturrecherche für die Interpretation

Um die Sortierergebnisse zu verstehen, wurde für jede Pflanze in einer Pflanzen-Enzyklopädie Pflanzenfamilie, Herkunft, Wasser- Licht- und Wärmebedarf ermittelt (Brikell 1999). Auch wurde einem Gärtner die NMDS-Konfigurationen und integrierten HCA-Cluster vorgelegt. Er wurde im offenen Interview zu jeder Zimmerpflanze nach den selben Kategorien, nach der floristischen Verwendung, nach der Warenpräsentation und nach dem Preis gefragt, sowie nach möglichen Gründen für die Ergebnisse der Ähnlichkeitsbeurteilung. Die Ergebnisse werden in der Interpretation dargestellt.

7.3.6 Gesamtschau und Interpretation der Ergebnisse

Interpretation in Hinblick auf die Hypothese: Personenübergreifende Ähnlichkeitsbeurteilung Die Ergebnisse weisen darauf hin, dass es ausgeprägte, personenübergreifend ähnliche Begriffsstrukturen für Zimmerpflanzen gibt: Die Ergebnisse des Scree-test – die

Stresswerte zeigen für die 2-dimensionalen NMDS-Konfigurationen signifikante Werte auf dem 1% Niveau für alle drei Stichproben. Die signifikanten Stresswerte bestätigen die Annahme, dass die Personen jeder Stichprobe die Zimmerpflanzen weniger nach individuellen Kriterien, sondern vielmehr nach übereinstimmenden Kriterien beurteilen. In den Ergebnissen der NMDS und HCA zeigen sich eindeutige Konfigurationen bzw. Cluster. Die Hypothese 1 wird durch die Ergebnisse bestätigt. Dies Ergebnis entspricht den Theorieannahmen, obwohl zu Beginn der Untersuchung starke Zweifel auftauchten, ob die KundInnen zur Objektklasse Zimmerpflanzen gemeinsames begriffliches Wissen haben. Dass die Fachkräfte übereinstimmen, wurde erwartet. Denn es wurde davon ausgegangen, dass Zimmerpflanzen Gegenstand ihrer täglichen beruflichen Arbeit sind und dass sie über Zimmerpflanzen in der Berufsschule gelernt haben. Dass deren Ähnlichkeitsurteile allerdings aus botanischer Sicht falsch sind, wurde nicht erwartet, doch dazu später mehr bei der Hypothese „Bisherige Handlungserfahrung“.

Interpretation in Hinblick auf die Hypothese: „Aktuelle Zielorientierung“ Die KundInnen mit der Instruktion „Bitte stellen Sie sich vor, Sie möchten eine Pflanze kaufen!“ sortieren die Zimmerpflanzen sehr ähnlich wie die KundInnen, die diese Instruktion nicht erhielten. Die unterschiedliche Instruktion zeigt somit auf der Verhaltensebene des Sortierens kaum Effekte. Dies zeigt die hohe Rangkorrelation zwischen den beiden Stichproben von $r_{\text{Best-Ma\ss}} = 0.78$ sowie die deskriptive Analyse der HCA-Cluster. Beide Gruppen unterscheiden sich nur in der Zuordnung der Yucca. Die KundInnen ohne Zielorientierung ordnen sie den Palmen (Phoenix und Kokospalme) zu, obwohl die Yucca gar keine Palme ist und andere Ansprüche hat, aber auch hoch wächst. KundInnen mit der Kauf-Zielorientierung ordnen sie dem Efeu und dem Benjamin zu. Dass die beiden Stichproben sich nicht unterscheiden, könnte daran liegen, dass die Instruktion zu schwach ist.

Die Instruktion zeigt aber auf der verbalen Ebene einen signifikanten Effekt. Dies ist sehr erstaunlich, denn die KundInnen nannten ihre Sortierkriterien während des Sortierens, um ihr Sortierverhalten zu beschreiben! Und das Sortierverhalten der beiden Gruppen ist nahezu identisch! Doch die KundInnen ohne Zielorientierung nennen signifikant mehr sensorische Kriterien (Blattfarbe, Größe etc.) und kategoriale Kriterien als Grund

ihres Sortierverhaltens sowie weniger Pflege- und Auswahlkriterien. Ganz anders sind die Ergebnisse der KundInnen mit Kauforientierung: sie nennen weniger sensorische und weniger kategoriale Kriterien, aber mehr Pflege und Auswahlkriterien. Die Instruktion scheint demnach einen nachweisbaren Einfluss auf das Verbalverhalten der Versuchspersonen zu haben, aber einen ausgesprochen geringen Einfluss auf das Sortierverhalten.

Wodurch lassen sich die Unterschiede bei den verbalen Nennungen und im Sortierverhalten erklären? Es könnte sein, dass die Instruktion zu schwach ist. Der Effekt im Verbalverhalten könnte auch auf die soziale Erwünschtheit zurückzuführen sein: Die Versuchspersonen nennen Kriterien, von denen sie glauben, dass der Versuchsleiter sie hören will. Es könnte dadurch bedingt sein, dass die verbalen Reaktionen reflektiert werden können, nicht aber die semantischen Ähnlichkeitsurteile im Sortierverhalten, weil sie tieferliegenden, vorbewussten Mechanismen folgen. Die Übereinstimmung im Sortierverhalten kann dann als Ausdruck dessen interpretiert werden, dass die Begriffsstrukturen stabil und unabhängig von der *aktuellen* Zielorientierung sind. Sie bestätigen die Theorieannahmen von Klix und Hoffmann, die davon ausgehen, dass ereignisbezogene Begriffsstrukturen quasi-stabil repräsentiert sind. *Aufgrund des geringen Effektes der Instruktion wird die Hypothese „Die Ähnlichkeitsbeurteilung von Zimmerpflanzen sind unabhängig von der aktuellen Zielorientierung.“ nicht bestätigt, aber auch nicht verworfen.*

Es wird jedoch eine *andere Schlussfolgerung* aus den Ergebnissen gezogen: Es scheint, dass *unterschiedliche Instruktionen keinen bedeutenden Effekt auf das Sortierverhalten haben*, sofern sich die Versuchspersonen in dem für die zu sortierende Objektmenge typischen Umfeld befinden.

Interpretation in Hinblick auf die beiden Hypothesen „Bisherige Handlungserfahrung“ und „Handlungsbezogene Kriterien“ Die Begriffsstrukturen von KundInnen und Fachkräften weisen trotz großer Ähnlichkeit bedeutsame Unterschiede auf, die auf die unterschiedlichen Handlungsziele und -erfahrungen der beiden Gruppen zurückgeführt werden können. Da sich die Unterschiede über die zugrunde liegenden Klassifizierungskrite-

rien erklären lassen, werden im Folgenden die Ergebnisse zu beiden Hypothesen zusammenfassend dargestellt und interpretiert:

Die Ergebnisse der Fachkräfte weisen für alle Objekte eine – nach Gigerenzer (1981) – degenerative Lösung auf. Ihre Sortierergebnisse zeigen im Vergleich zu den KundInnen in der NMDS-Konfiguration und in der HCA-Analyse kompaktere Gruppierungen bzw. Cluster und haben den geringsten Stresswert für alle 24 Zimmerpflanzen (Stresswert von 0.061* für die 2-dimensionale Lösung im Vergleich zu einem signifikanten Stresswert von 0.073* der KundInnen). Dies ist nach Gigerenzer ein typisches Ergebnis für eine Gruppe mit homogenen Vorerfahrungen, wie für eine Expertengruppe (Gigerenzer 1981, S. 346). Die „Fachkräfte“ nennen überdies bei den Sortierkriterien-Nennungen häufiger Verhaltensmerkmale und weniger sensorisch-anschauliche und kategoriale Kriterien als die KundInnen. Auch dies ist nach Gigerenzer charakteristisch für eine Gruppe mit homogenen Vorerfahrungen. Diese Ergebnisse entsprechen den Erwartungen. Die „Fachkräfte“, die sich intensiver aufgrund der fachlich beruflichen Aufgaben mit Zimmerpflanzen beschäftigen, haben eine ausgeprägtere Begriffsstruktur.

Die Begriffsstrukturen der KundInnen und Fachkräfte stimmen zu ca. 70% überein ($r_{\text{Best-Ma\ss}} = 0.70$). Dies zeigt sich inhaltlich darin, dass beide Stichproben die 24 Zimmerpflanzen auf der ersten Sortierebene strikt nach Grün- und Blühpflanzen trennen. Wichtigstes übereinstimmendes Klassifizierungskriterium ist somit die Blühfähigkeit. Die Übereinstimmung ist aufgrund der unterschiedlichen Verwendung von Grün- und Blühpflanzen nicht weiter verwunderlich. Die Begriffsstrukturen weisen aber einen Unterschied von 30% auf, der sich deutlich in den Konfigurationen der NMDS und in den Clusterstrukturen der Hierarchischen Clusteranalyse zeigt. Bei Blühpflanzen haben KundInnen und Fachkräfte kein identisches Subcluster in der HCA (vgl. Tabelle 7.3-7 Subcluster der Clusteranalyse) und auch bei den NMDS-Konfigurationen zeigen sich andere Ähnlichkeitsrelationen. Bei den Grünpflanzen gibt es größere Übereinstimmungen in den Clustern und Konfigurationen.

Folgende inhaltliche Analyse zeigt, dass dieser Unterschied auf Faktoren zurückzuführen ist, der in den unterschiedlichen, bisherigen Handlungserfahrungen der beiden

Nutzergruppen zu finden ist. Basis der inhaltlichen Analyse ist die Befragung eines Gärtners sowie die Auswertung einer Pflanzen-Enzyklopädie (Brickell 1999).

Innerhalb der *Blühpflanzen* beurteilen die KundInnen z.B. die Orchidee, die Stephanotis und die Passionsblume als ähnlich. Die Orchidee ist eine Aufsitzpflanze, die im feucht-tropischen Regenwald in den halbschattigen, oberen Etagen der Bäume wächst; die Stephanotis ist eine Rankpflanze aus Madagaskar, die Halbschatten braucht und die Passionsblume ist eine Rankpflanze aus dem tropischen Südamerika, die volle Sonne benötigt. Die Ähnlichkeitsbeurteilung ist aus botanischer *und* pflegerischer Sicht gänzlich falsch. Sie lässt sich aber über die Exotik (Fragilität und Schönheit der Blüten) erklären sowie den ähnlich hohen Preis (15.- bis 35.- Euro). Sie eignen sich alle drei als „größeres“ Geschenk, was darauf schließen lässt, dass innerhalb der Blühpflanzen eine Gruppe unterschieden wird, die sich als „größeres Geschenk“ charakterisiert.

Die KundInnen bilden des Weiteren eine Gruppe aus Kaktus und Aechmea. Auch dies ist botanisch-kategorial *und* pflegerisch gänzlich falsch. Die Kakteen brauchen wenig Wasser, volle Sonne und niedrige Luftfeuchte. Die Aechmea ist ein Ananasgewächs aus Südamerika und gehört wie die Orchidee zu den Epiphyten (Halbschatten, hohe Luftfeuchte). Aechmea und Kaktus sehen aber beide stachelig aus, haben eine ähnliche, gehobene Preisklasse und eignen sich als besonderes Geschenk für Liebhaber dieser Pflanzen. Es sind beides keine Pflanzen, die man eben mal verschenkt, weil sie zu speziell in der Pflege, aber auch im Aussehen sind.

Auch den anderen Ähnlichkeitsbeurteilungen der KundInnen liegen keine botanischen, d.h. kategorialen Kriterien, aber auch keine pflegerischen Kriterien zugrunde. Die KundInnen gruppieren z.B. Azalee, Cyclamen (Alpenveilchen) und Usambara-Veilchen. Die Azalee ist ein Rhododendrongewächs, braucht viel Wasser und viel Licht und kühle Temperaturen. Das zarte Usambara-Veilchen, ein Geneseriengewächs, eine ehemals seltene Pflanze aus dem Usambara-Gebirge Ostafrikas, die normalerweise im tropischen Halblight am Fuße hoher Bäume wächst, braucht mäßig Wasser, wenig Licht und hohe Lufttemperatur. Die Cyclame (Alpenveilchen), ein Primelgewächs kommt aus den kühlen Bergen Persiens, verträgt keine Sommerhitze, braucht Halbschatten und viel Wasser.

Doch es gibt eine wichtige Gemeinsamkeit im Verwendungszweck: alle drei Pflanzen sind „kleine Geschenk- und Dekorationspflanzen“! Sie sind ausgesprochen billig (3.- bis 6.- Euro), die KundIn kauft sie für sich oder als kleines Mitbringsel z.B. zum Kaffeekränzchen, sie werden auf den Tisch oder das Fensterbrett gestellt und nach dem Verblühen weggeworfen. Die KundInnen sind nur mit Fachkenntnissen in der Lage, diese Pflanzen selber wieder zum Blühen zu bekommen. Das Kriterium der billigen Schenk-Deko- und Wegwerfpflanze wurde von den KundInnen während des Sortierens nicht genannt – vielleicht aus Gründen der sozialen Erwünschtheit.

Die Blühpflanzen wurden von den KundInnen durchweg nicht nach botanischen Fachkriterien beurteilt. Dies wurde erwartet und entspricht der Hypothese. Die Ergebnisse zeigen, dass der Begriffsstruktur handlungsbezogene Kriterien zugrunde liegen, auch wenn es nicht ganz die erwarteten Handlungskriterien sind: Es wurde erwartet, dass die KundInnen die Pflanzen nach pflegerischen Kriterien beurteilen und dies trifft nicht zu. Die KundInnen beurteilen die Blühpflanzen vielmehr nach kauf-, schenk- und dekorations-orientierten Handlungskriterien. Dies sind zwar auch handlungsbezogene Kriterien, aber keine pflegerischen, sondern eben verwendungszweck- bzw. finalitätsbezogene. Dies stimmt auch mit den Ergebnissen des Vorversuchs überein: die befragten KundInnen geben an, Zimmerpflanzen für sich selbst zu Hause zu kaufen oder um sie zu verschenken. Ihre Beurteilungskriterien bzgl. Zimmerpflanzen bezogen sich auf Kriterien zum „Schenken“ und „Daheim aufstellen & genießen“ sowie auf die Kriterien Kaufpreis, persönliches Gefallen, Aussehen und weniger auf das Pflegen.

Auch die *Fachkräfte* bilden bei den *Blühpflanzen* die Gruppe der „Wegwerf-Billig-Pflanzen“. Sie nennen dieses Kriterium übrigens auch nicht! Sie gruppieren das Usambara, die Azalee und das Alpenveilchen auch als ähnlich und darüber hinaus auch das Solanum, ein südamerikanisches Nachtschattengewächs aus der gleichen Familie wie Kartoffel, Tomate und Tabak, eine anspruchslose, etwas stinkende Pflanze mit lustigen, orangefarbenen Früchten. Auch das Solanum ist billig und ein „Wegwerfprodukt“! Im Verkaufsraum werden alle diese Produkte in der Regel gemeinsam auf so genannten „Sonderangebots- oder Aktionstischen“ präsentiert. Der Ähnlichkeitsbeurteilung der Fachkräfte liegen hier auch keine kategorial oder pflegerischen Kriterien zugrunde, sondern waren-

präsentations- und verkaufsbezogene Handlungskriterien. Dass die KundInnen das Solanum nicht zu Usambara-, Alpenveilchen und Azalee ordnen ist hinsichtlich des Handlungsziels klar: eine etwas stinkende Pflanze verschenkt man nicht und stellt sie nicht auf den Tisch – auch wenn sie hübsch aussieht und billig ist.

Die Fachkräfte gruppieren Bougainvillea, Stephanotis und Passionsblume. Auch dieser Gruppe liegen nicht botanisch-kategoriale oder pflegerische Kriterien zugrunde als vielmehr warenpräsentations- und verkaufsbezogene Handlungskriterien: alle drei Pflanzen ranken, können in Kübeln verkauft werden und haben verkaufsattraktive Blüten, auch wenn die Pflegeansprüche der sonnenhungrigen, anspruchslosen Bougainvilleae einem Wunderblumengewächs aus Südamerika sich von denen der Stephanotis (Halbschatten) und der Passionsblume (volle Sonne) absolut unterscheiden. Überdies ist die Stephanotis eine reine Zimmerpflanze, während die Bougainvillea auch in den Garten gepflanzt werden kann. Auch hier weisen Ähnlichkeitsbeurteilungen auf verwendungszweckbezogene bzw. finalitätsbezogene Kriterien hin.

Die Fachkräfte beurteilen zumindest die stachelige Aechmea und die zarte Orchidee als ähnlich, schließlich sind beide Epiphyten. Beide benötigen auch die gleiche Pflege und die gleichen verkaufsbezogenen Handlungen: nämlich eine intensivere Beratung.

Auch die Begriffsstrukturen *der Fachkräfte und der KundInnen* für *Grünpflanzen* folgen nicht kategorialen, sondern handlungsbezogenen Kriterien: und auch hier sind es nicht pflegerische, sondern kauf- bzw. verkaufs- und warenpräsentationsbezogene und damit verwendungs- bzw. finalitätsbezogene Kriterien. Kroton und Dieffenbachia sind buntblättrige „Arrangement-Pflanzen“, Gummibaum, Philodendron und Monstera sind grüne, großblättrige „Arrangement-Pflanzen“: sie werden gerne gemeinsam in Pflanzkübel gepflanzt für den Wohninnenraum oder das Büro. Auch das Efeu in der Nähe zur Yucca erklärt sich bei den Fachkräften daher, dass beide gerne von den Fachkräften zusammen in Pflanzkübeln arrangiert werden: die Yucca wächst nach oben, der Efeu gibt unten Fülle. Es sind Pflanzen, die KundInnen und Fachkräfte in einem Kontext für das gleiche Handlungsziel verwenden bzw. erleben: für einen Grünpflanzenpflanzkübel. Es sind aber auch Pflanzen, die austauschbar sind hinsichtlich des Handlungsziels: anstatt Kroton wird bzw. ist eine Dieffenbachia mit Efeu und/oder einer Yucca gepflanzt.

Eine mögliche kategoriale, aber auch handlungsbezogene Beurteilung könnte der Palmengruppe zugrunde liegen: Die Fachkräfte beurteilen Phönix und Kokospalme als ähnlich. Beides sind Palmengewächse, beide benötigen volle Sonne und wollen am liebsten 20m in die Höhe und mit den Wurzeln 40m in den Grund wachsen. Doch überdies sind beide ähnlich teuer und werden ähnlich selten abverkauft. Die KundInnen beurteilen die Yucca, eine asiatische, anspruchsloses halbschattenliebendes Agavengewächs –als kein Palmengewächs – ähnlich der Phönix und Kokospalme. Diese Pflanzen eignen sich auch hinsichtlich des Handlungsziels „Pflanzkübel für die Wohnungsdeko“.

Zusammenfassend lässt sich formulieren: Die Begriffsstruktur der KundInnen *und* der Fachkräfte bezüglich Zimmerpflanzen folgt nicht kategorialen Kriterien (Familienähnlichkeit) der Pflanzensystematik, sondern handlungsbezogenen Kriterien und hier insbesondere verwendungszweck bzw. finalitätsbezogenen Kriterien. Da KundInnen und Fachkräfte sich in ihren Handlungszielen im Umgang mit Zimmerpflanzen unterscheiden, unterscheiden sie sich somit auch in den Ähnlichkeitskriterien! Die Begriffsstruktur der KundInnen folgt Handlungskriterien, die sich darauf beziehen, ob die Pflanze sich als großes Geschenk oder kleines Mitbringsel eignet, wie die Pflanze als Wohnungsdekoration zu verwenden ist etc. Dies stimmt auch mit den Befragungsergebnissen der KundInnen des Vorversuchs überein. (Aber nicht mit den Kriterien, die die KundInnen während des Sortierens nannten!) Die Begriffsstruktur der Fachkräfte folgt Handlungskriterien, die sich auf den Abverkauf und die Warenpräsentation (schnell drehender Aktionsartikel, langsam drehende und beratungsintensive Standard- oder Exotikartikel, Artikel geeignet für Pflanzarrangement etc.) beziehen. Die Begriffsstruktur wird demnach geleitet von handlungsbezogenen Kriterien, die sich auf den Zweck bzw. die Finalität beziehen im Sinne von „Was kann man mit den Pflanzen tun? Wozu eignen sie sich?“ und weniger vom prädikats- oder aktionsorientierten Kriterium: Wie muss man die Pflanze pflegen und behandeln?

Diese Ergebnisse stimmen mit der Hypothese überein, die aus den Theorieannahmen von Klix. et al. abgeleitet wurde. Sie lassen aber noch eine spezifischere Annahme formulieren: den Ähnlichkeitsbeurteilungen, also der Begriffsstruktur, liegen handlungszielbezogene Kriterien zugrunde. Es werden solche Pflanzen ähnlich beurteilt, die in einem

gemeinsamen Handlungskontext hinsichtlich eines Handlungsziels verwendet wurden bzw. die hinsichtlich des gleichen Handlungsziels austauschbar sind, d.h. funktional äquivalent sind. Dies legt die Vermutung nahe, dass Begriffsstrukturen nicht nur handlungsbezogen im Sinne von prädikats- bzw. tätigkeitsorientiert gebildet werden, sondern auch hinsichtlich Handlungszielen. Handlungsziele scheinen eine Ordnung stiftende Funktion für die Begriffsstruktur zu haben. Dies erscheint nachvollziehbar vor der Annahme, Begriffe dienen der Verhaltenssteuerung, würde aber für die Begriffstheorie von Klix et al. bedeuten, dass man sie umformulieren müsste. Denn dort dominiert ein prädikats- und aktionsorientiertes Kriterium für ereignisbezogene Begriffsstrukturen: Ordnung stiftende Funktion hat dort der Tätigkeitsbegriff und nicht das Handlungsziel.

7.3.7 Beurteilung der verwendeten Methoden

Die Methoden, die im Vor- und im Nachversuch eingesetzt wurden, sollen nun beurteilt werden, wie weit sie sich für ein generelles Untersuchungsdesign eignen, um die Begriffsstruktur für Objektklassen zu untersuchen.

Beurteilung des Vorversuchs zur Auswahl der zu sortierenden Elemente Im Vorversuch wurden KundInnen und Fachkräfte befragt, um 24 bekannte und für das Sortiment repräsentative Zimmerpflanzen auszuwählen. Das inhaltliche Ziel wurde erreicht. Dennoch waren die offenen Fragen unbefriedigend in der Auswertung, weil die qualitative Auswertung viel Unsicherheit mit sich brachte. Eine klare Trennung in ein nondirektives Interview zur Ermittlung der Urteilsdimensionen und anschließende geschlossene Fragen (bzw. Ranking) für eine quantitative Ermittlung erscheinen sauberer und zuverlässiger.

Beurteilung der Property Fitting Analyse zur Objektivierung der Interpretation der NMDS und HCA Ergebnisse Die Property Fitting Analyse zeigt sich problematisch. Sie liefert Erklärungsdimensionen, die auf den häufigsten Sortierkriterien-Nennungen basieren. Die Untersuchung zeigte jedoch, dass sich die verbalen Nennungen der Sortierkriterien vom faktischen Sortierverhalten unterscheiden. Es wurden wesentliche Sortierkriterien nicht genannt. Möglicherweise liegt das daran, dass das benötigte Wort nicht zur

Verfügung steht, dass die Personen ungeübt sind in der Verbalisierung ihrer Urteile oder aber, dass Sortierkriterien aus Gründen der sozialen Erwünschtheit nicht genannt werden. Es taucht aber auch der Verdacht eines induktionistischen Zirkelschlusses auf: durch den Versuch Verhaltensäußerung mittels anderer Verhaltensäußerungen zu erklären, lassen sich keine Erklärungs- und Prognoseargumente finden. Deshalb scheint es ratsam auf eine Property Fitting Analyse zu verzichten und die Ergebnisse hypothesenorientiert zu interpretieren.

7.3.8 Zusammenfassung

Ausgangspunkt war die theoriegeleitete Hypothese, dass sich personenübergreifend stabile Ähnlichkeitsbeurteilungen für Zimmerpflanzen finden lassen, dass es aber Unterschiede bei KundInnen und Fachkräften gibt. 56 Personen, davon 45 KundInnen und 11 Fachkräfte, sortierten 24 Zimmerpflanzen mittels der Methode des Hierarchischen Sortierens. Die Ergebnisse wurden mit der nonmetrischen Multidimensionalen Skalierung und der Hierarchischen Clusteranalyse ausgewertet. Für die Interpretation wurde ein Nachversuch mit einer Property Fitting Analyse an 160 Personen durchgeführt.

Für die KundInnen und Fachkräfte lassen sich ausgeprägte Ähnlichkeitsbeurteilungskriterien für Zimmerpflanzen finden. Die Ähnlichkeitsbeurteilungen beider Gruppen stimmen in der Basis überein: Beide Gruppen trennen scharf nach Blüh- und Grünpflanzen. Innerhalb der Blüh- und Grünpflanzen gibt es allerdings bedeutsame Unterschiede bei den KundInnen und Fachkräften, die sich auf die unterschiedlichen Handlungsziele und -erfahrungen zurückführen lassen. Der Ähnlichkeitsbeurteilung der KundInnen liegen handlungsbezogene Kriterien zugrunde, die sich auf den Verwendungszweck der Pflanze als Geschenk für sich oder andere und die Zimmerdekoration beziehen: So gibt es bei den Blühpflanzen z.B. die Gruppe der preiswerten Geschenk-Deko-Wegwerfpflanzen für sich oder andere, die Gruppe der hochpreisigen, exotischen Geschenkpflanzen und die Gruppe der Liebhaber-Sammelpflanzen. Der Ähnlichkeitsbeurteilung der Fachkräfte liegen dagegen Kriterien zugrunde, die sich auf den Abverkauf und die Warenpräsentation beziehen. So gibt es bei ihnen z.B. die Gruppe der hochpreisigen Rank-

und Kübelpflanzen, der preiswerten, schnell drehenden Sonderangebotspflanzen oder der tropischen Spezialitäten.

Der Ähnlichkeitsbeurteilung liegen keine kategorialen (familienähnlichen) Kriterien zugrunde. Aus botanischer als auch aus pflegerischer Sicht liegen sowohl bei KundInnen als auch bei den Fachkräften durchwegs falsche Klassifizierungen vor. Die Ergebnisse zeigen, dass die Ähnlichkeitsbeurteilung beider Stichproben handlungsbezogenen Kriterien folgt, wenngleich sich die Handlungskriterien nicht auf die Pflege (Licht-, Wasserbedarf etc.), sondern vielmehr auf das Handlungsziel und die Verwendung beziehen.

Durch eine Veränderung der Instruktion bei den KundInnen konnte kein bedeutsamer Unterschied im Sortierverhalten erreicht werden. Ob dies Ausdruck einer stabilen Begriffsstruktur ist, konnte nicht mit Sicherheit entschieden werden, da die Instruktion eventuell nicht ausreicht, um eine Zielorientierung zu verändern. Allerdings zeigten sich bei den unterschiedlichen Instruktionen bedeutsame Unterschiede in den Sortierkriterien-Nennungen, die aber nicht mit dem Sortierverhalten übereinstimmen. Die Property Fitting Analyse brachte keine wirklichen Erklärungen zur Interpretation der Ähnlichkeitsurteile und erscheint im Nachhinein überflüssig.

7.4 Ermittlung einer käuferorientierten Anordnung von Süßgebäck und Überprüfung der Ergebnisse in einem Feldexperiment

7.4.1 Einleitung

Problemstellung Die folgende Untersuchung beschäftigt sich mit einem Anordnungsproblem des Einzelhandels. Wie müssen Artikel im Regal angeordnet sein, damit die KundInnen sich schnell und einfach zurechtfinden? Oder ist es egal, wo welche Artikel liegen, Hauptsache die Anordnung ist nur immer gleich, damit die KundInnen lernen können, wo sie was finden? Die vorliegende Studie untersucht diese Fragestellung stellvertretend an der Warengruppe Süßgebäck. Konkret stellen sich hier die Fragen: Gibt es eine „kunden-orientierte“ Anordnung, z.B. für Butterkekse, Schokowaffeln und Löffelbiscuit? Welche Rolle spielt eine Anordnung nach Marken? Welche Auswirkungen hat eine „kundInnen-orientierte“ Anordnung auf das Verhalten der KundInnen? Unterstützt sie wirklich das Such- und Findeverhalten der KundInnen?

Anlass der Studie Hintergrund dieser Untersuchung war ein Auftrag der mittelständischen Lebensmittel-Einzelhandelskette Feneberg GmbH im Allgäu. Der Auftrag lautete, ein neues, „kundInnen-orientiertes“ Supermarkt-Konzept zu entwickeln. Die Geschäftsführer und Inhaber hatten die Vision, dass ein Supermarkt, der die Kundinnen – die meisten Kunden eines Supermarktes sind Frauen – unterstützt und ihren Bedürfnissen entgegenkommt, dem Unternehmen einen Marktvorteil bieten würde: in einer intensiveren Kundenbindung, in einer höheren Kundenzahl und vielleicht sogar in höheren Umsätzen. Bei der Führungsmannschaft: dem Vertriebsleiter, dem Einkaufsleiter und dem Verkaufsflächenplaner konnten sie sich mit ihrer Vision nicht ohne weiteres durchsetzen: alte Gewohnheiten, Bedenken vor Umsatz- und Ertragsverlusten sowie Mangel an Know-how hielten sie davon ab, tradierte Einzelhandelsregeln zu verlassen und Neues auszuprobieren.

Bei der Regalbelegung hatte man schon seit Jahrzehnten die „traditionellen“ Regeln der Hersteller-orientierten Regalbelegung befolgt: Alle Produkte eines Herstellers bzw. einer Marke werden hier im Regal zusammen aufgestellt. Hintergrund dieser Regel war

die Überzeugung, der Kunde sei markentreu und orientiere sich zunächst am Hersteller bzw. der Marke und dann erst am Inhalt des Produktes. Teure Produkte müssen in die Sicht- und in die Greifzone des Regals, preiswerte Produkte in die Bück- oder Reckzone (unter 0,8 m und über 1,60m) (vgl. Vance Packard 1957), damit der Kunde eher teure als preiswerte Produkte sieht und kauft. Die geschäftsführenden Inhaber hofften auf neue, psychologische Erkenntnisse. Sie erteilten den Auftrag, "KundInnen-freundliche" Supermarkt-Konzepte zu entwickeln mit der Vorgabe, die Auswirkungen sowohl auf betriebswirtschaftliche Kennziffern (Absatz, Umsatz, Ertrag) als auch auf das Kundenverhalten empirisch zu testen.

Die vorliegende Studie ist eine von vielen zur Entwicklung einer KundInnen-orientierten *Artikelanordnung* im Supermarkt und zeigt stellvertretend am Beispiel der Warengruppe "Süßgebäck", wie die Anordnung der Artikel im Regal für die KundInnen optimiert werden kann.¹ Für den interessierten Leser sei erwähnt: die Artikelanordnung ist nur ein "Ansatzpunkt", um eine Verkaufsfläche "KundInnen-orientiert" zu gestalten. Andere Ansatzpunkte sind zum Beispiel die Lauf- und Wegeführung, die Sortimentsplatzierung, die Sortimentsgestaltung, die Preisgestaltung, die Beleuchtung, die Farbgestaltung und natürlich nicht zu vergessen das Verhalten der Verkaufsmitarbeiter.

Ziele der Studie Diese Studie verfolgt zwei Ziele:

- 1) In einem ersten Experiment soll die Begriffsstruktur der Supermarkt-KundInnen für Süßgebäck untersucht werden. Es sollte dabei erstens geprüft werden, ob die KundInnen darin übereinstimmen, welche Süßgebäck-Artikel sie als begrifflich-semantisch ähnlich bzw. unähnlich beurteilen. Und es sollte zweitens überprüft werden, ob die KundInnen in ihrer Ähnlichkeitsbeurteilung der traditionellen, Hersteller-orientierten Regalbelegung folgen und alle Artikel eines Herstellers (z.B. Bahlsen, Filia, Leibnitz) als ähnlich beurteilen oder ob sie die Artikel nach eigenen, KundInnen-spezifischen Kriterien beurteilen.
- 2) In einer Evaluations-Studie soll überprüft werden, ob eine „Begriffsstruktur-orientierte“ Anordnung der Süßgebäck-Artikel im Regal das Such- und Findeverhalten der KundInnen unterstützt. Die Ergebnisse sollten als Vorlage für eine „Begriffsstruktur-

¹ Die Untersuchung wurde von Petra Kinateder im Rahmen ihrer Diplomarbeit durchgeführt.

orientierte“ Regalanordnung dienen: semantisch ähnliche Süßgebäck-Artikel sollten entsprechend räumlich nahe, semantisch unähnliche sollten entsprechend entfernt angeordnet werden. Es sollte dann überprüft werden, welche Effekte eine „Begriffsstruktur-orientierte“ Süßgebäck-Artikelanordnung auf das Kundenverhalten und den Abverkauf hat. Um zu prüfen, ob die „Begriffsstruktur-orientierte“ Süßgebäck-Anordnung die KundInnen besser bei ihrem Finde-, Auswahl- und Kaufverhalten unterstützt als die „Hersteller-orientierte“, musste überdies geklärt werden, welches Verhalten die KundInnen als „besser“ bewerten.

Untersuchungsdesign Die Studie besteht aus zwei empirischen Untersuchungen: In einem *Sortierexperiment* mit 51 Personen wurde die Begriffsstruktur für Süßgebäck ermittelt. In einem *Feldexperiment* wurde dann an 1047 Personen untersucht, welche Auswirkungen eine Regalbelegung, die der Begriffsstruktur der KundInnen entspricht, auf das Verhalten der KundInnen und auf den Abverkauf von Süßgebäck hat. Abbildung 7.4-1 gibt einen Überblick über das Untersuchungsdesign.

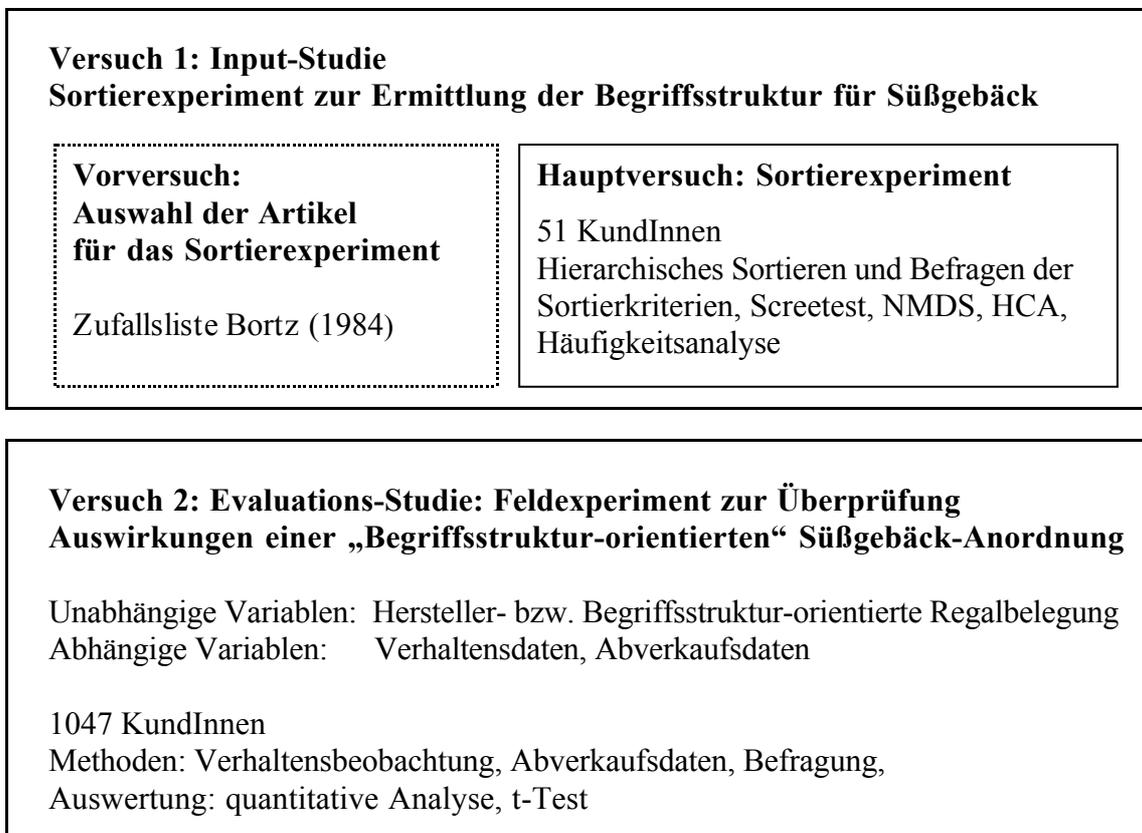


Abbildung 7.4-1. Untersuchungsdesign der Studie Süßgebäck

7.4.2 Input-Studie: Erhebung der Begriffsstruktur von Süßgebäck

Einleitung

Das Sortierexperiment sollte klären, ob sich für die KundInnen eine Übereinstimmung in den Begriffsstrukturen finden lässt und wenn ja, wie diese aussieht, ob sie ein Abbild der gelernten Regalanordnung im Feneberg-Supermarkt ist oder eigenen, kundenspezifischen Klassifizierungskriterien folgt. Die Begriffsstruktur wurde als semantische Ähnlichkeit operationalisiert. Entsprechend der Theorie von Klix (1971), Hoffmann (1983, 1986) et al. wurde angenommen, dass die Begriffe und Begriffsstrukturen abhängig vom Handeln sind und daher Handlungserfahrungen repräsentieren. Personen mit ähnlichen Handlungserfahrungen bezüglich Süßgebäck müssten darin übereinstimmen, welche Süßgebäckartikel semantisch ähnlich bzw. unähnlich sind. Basierend auf der Theorie von Klix (1971), Hoffmann (1983,1986) et al. wurde des Weiteren davon ausgegangen, dass die Begriffsstruktur der KundInnen handlungsbezogenen Kriterien folgen müsste. Handlungsbezogene Kriterien werden definiert als Kriterien, die mit dem Geschmack und mit der Verwendung von Süßgebäck zu tun haben.

Es wurde daher angenommen, dass die KundInnen z.B. Gebäckmischungen als semantisch ähnlich beurteilen und deutlich von z.B. Butterkekse unterscheiden. Gebäckmischungen haben einen spezifischen Handlungskontext (Kaffeetrinken), der sich vom Handlungskontext für Butterkekse unterscheidet. Butterkekse sind Kekse für zwischendurch, für Kinder oder auf Ausflügen.

Würden die KundInnen allerdings das Süßgebäck entsprechend der Marke bzw. des Herstellers als ähnlich beurteilen und z.B. alle Süßgebäck-Artikel von Filia, Bahlsen etc. zusammenfassen – unabhängig davon, ob es sich um einen Butterkeks, Nussblätter, Waffeln oder eine Gebäckmischung handelt – dann würde dies der Annahme widersprechen. In diesem Falle würde die Begriffsstruktur der KundInnen vielmehr die Regalanordnung von Feneberg wiedergeben: in den Feneberg-Supermärkten wurden die Süßgebäckartikel traditionell nach Marken bzw. Herstellern (z.B. Leibnitz, Bahlsen, Brandt, Filia etc.) angeordnet. Die Begriffsstruktur würde somit ein Abbild der wahrgenommenen Realität im Supermarkt darstellen und nicht ein Abbild der Handlungsrealität.

In dieser Studie wurden die folgenden Hypothesen untersucht:

Hypothese 1: Personenübergreifende, stabile Begriffsstruktur Die KundInnen haben eine personenübergreifende, d.h. überindividuelle Begriffsstruktur für Süßgebäck, d.h. die semantische Ähnlichkeitsbeurteilung von Süßgebäck folgt gemeinsamen Kriterien.

Hypothese 2: Handlungsbezogene Kriterien Der Ähnlichkeitsbeurteilung von Süßgebäck-Artikeln liegen handlungsbezogene Kriterien der KundInnen und nicht kategoriale Kriterien zugrunde. Die semantische Ähnlichkeitsbeurteilung erfolgt nach den Kriterien: „Geschmack“ und „Verwendung“. Die Ähnlichkeitsbeurteilung erfolgt nicht nach dem Kriterium „Herstellerfirma“ bzw. „Marke“.

Methode

Erhebungsmethode Die semantisch-begriffliche Ähnlichkeit wurde mit der Methode des Hierarchischen Sortierens untersucht (Marx & Hejj 1989; vgl. Kapitel 6). Die Versuchsperson wird im Sinne eines forced choice aufgefordert, nach einem ihr passend erscheinenden Kriterium Süßgebäck-Artikel zunächst in zwei Gruppen zu sortieren, dann diese zwei Gruppen wieder in zwei Gruppen usw., bis jeder Süßgebäckartikel für sich allein steht (Abb. 7.4-2). Die Gruppen müssen dabei nicht gleich groß sein.

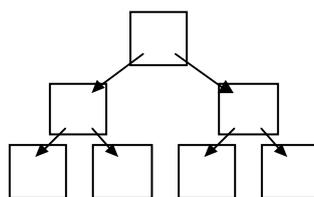


Abbildung 7.4-2. Veranschaulichung der Sortierschritte beim Hierarchischen Sortieren. Die Objekte werden so lange in Gruppen geteilt, bis sie allein stehen.

Nach jedem Sortierschritt wurde die Versuchsperson nach ihrem Sortierkriterium gefragt. Diese Nennungen sollten bei der Auswertung der Sortiererergebnisse herangezogen werden, um die Ergebnisse zu interpretieren und die Interpretation nicht nur den Annahmen des Untersuchers zu überlassen.

Untersuchungsmaterial Die Methode des Hierarchischen Sortierens erlaubt nur eine begrenzte Anzahl von zu beurteilenden Objekten (ca. 25), um die Versuchsperson und die Auswertung nicht zu überlasten. Das Süßgebäck-Sortiment des Feneberg Supermarktes in Kempten umfasste jedoch insgesamt 91 Artikel. Die 25 Artikel für das Experiment wurden nach dem Zufallsprinzip ausgewählt. Jedem Artikel wurde eine Nummer gegeben. Die Auswahl erfolgte dann nach der in Bortz (1984 S. 543) veröffentlichten Liste von Zufallszahlen. Dieses Verfahren schließt einen tendenziellen Einfluss durch den Versuchsleiter aus, ist schnell und damit ökonomisch. Es hat aber den Nachteil, dass nicht kontrolliert werden kann, ob die Auswahl repräsentativ für das Gesamtsortiment ist. Tabelle 7.4-1 zeigt die für das Sortierexperiment ausgewählten 25 Süßgebäck-Artikel. Da die Süßgebäck-Artikel handlich sind, wurden den Versuchspersonen die originalen Artikel auf einem Tisch für das Hierarchische Sortieren vorgelegt.

Tabelle 7.4-1. Ausgewähltes Süßgebäck für das Sortierexperiment n= 25 Artikel

Nr.	Artikel	Gewicht	Preis
1.	Filia Gebäckmischung	400 g	DM 2.45
2.	Brandt Zum Tee Gebäck	200 g	DM 2.49
3.	Filia Schoko-Keks zartbitter	125 g	DM 1.99
4.	Filia Nussblätter	125 g	DM 1.99
5.	Brandt Vollkorn Müsli Keks	200 g	DM 1.99
6.	Brandt Wiener Schnitten	200 g	DM 2.49
7.	Teekrun Haselnussgebäck	100 g	DM 2.98
8.	Italia Meringen	250 g	DM 2.45
9.	Italia Löffelbisquit	200 g	DM 2.49
10.	De Beukelaer Butterkekse	250 g	DM 2.29
11.	Filia Spritzringe	400 g	DM 1.99
12.	Filia Kaffeekränze	250 g	DM 1.25
13.	Leibnitz Butterkekse	200 g	DM 1.99
14.	Leibnitz Butterkekse	150 g	DM 1.69
15.	Bahlsen Grandioso	125 g	DM 2.99
16.	Bahlsen Romeo	150 g	DM 2.99
17.	Bahlsen Bobbes Apfel	125 g	DM 1.69
18.	Bahlsen Bobbes Erdbeere	125 g	DM 1.69
19.	Leibnitz Schoko-Keks Nuss	125 g	DM 2.49
20.	Azora Orangen-Kekse	125 g	DM 1.99
21.	Bahlsen Ouverture	250 g	DM 4.99
22.	Bahlsen Picadilly	150 g	DM 3.50
23.	Leibnitz Schoko-Keks Vollmilch	125 g	DM 2.49
24.	Leibnitz Schoko-Keks zartbitter	125 g	DM 2.24
25.	Bahlsen Messino Schoko/Orange	150 g	DM 2.69

Auswahl der Versuchspersonen und Durchführung Die KundInnen-Stichprobe wurde in Kempten in einem Supermarkt der Firma Feneberg untersucht. Auf dem Tisch in der Nähe des Süßgebäck-Regals lagen die 25 ausgewählten Süßgebäck-Artikel. Es wurden KundInnen angesprochen, die in der Nähe des Regals vorbeikamen. Die StudentInnen-Stichprobe wurde in München in der Ludwig-Maximilian-Universität in der Zeitungslesehalle angesprochen und in einen nahe gelegenen Raum geführt, wo sie einen Tisch mit den Süßgebäckartikeln vorfanden. Insgesamt wurden 51 Personen untersucht, davon 34 Kunden und Kundinnen und 17 Studenten und Studentinnen. Die Versuchspersonen wurden zufällig ausgewählt. Der hohe Frauenanteil in der Stichprobe ergab sich zufällig und ist gerechtfertigt, da die KundInnen eines Supermarktes in der Mehrzahl Frauen sind. Tabelle 7.4-2 beschreibt die Stichprobe.

Tabelle 7.4-2. Versuchspersonen für das Sortierexperiment mit Süßgebäck

Vpn-Merkmale	Frauen	Männer	Gesamt
KundInnen	26	8	34
StudentInnen	14	3	17
Insgesamt	40	11	51
Alter			
< 20	1	1	2
20-29	15	5	20
30-39	9	3	12
40-49	6	0	6
50-59	6	0	6
60-69	3	2	5
Berufe			
Hausfrau	15	0	15
StudentInnen aus 6 Fakultäten	14	3	17
Diverse Berufe	10	4	14
Rentner	1	2	3
Arbeitslos	0	1	1
Keine Angaben	0	1	1

Instruktion und Dokumentation Nachdem eine Versuchsperson an den Tisch herantreten war und Zeit hatte, sich zu orientieren, wurde sie mit folgenden Worten instruiert:

„Herzlichen Dank, dass Sie am Experiment teilnehmen. Sie sehen hier 25 Süßgebäcke. Ich möchte Sie bitten, diese in zwei Gruppen zu unterteilen. Die Gruppen müssen nicht gleich groß sein und es gibt keine richtigen oder falschen Lösungen. Mich interessiert nur, welche Produkte Sie zu einer Gruppe zusammenfassen.“ Nach jedem Sortierschritt wurde die Versuchsperson gefragt, nach welchem Kriterium sie ihre Wahl getroffen habe. Dann wurde sie gebeten: „Bitte trennen Sie jede dieser zwei Gruppen wieder in zwei Gruppen, so dass Sie vier Gruppen haben.“ Die Instruktion zum weiteren Trennen der Gruppen und anschließendem Nennen des Sortierkriteriums erfolgte so lange, bis alle Produkte einzeln auf dem Tisch lagen. Auf Zwischenfragen, wonach sie die Artikel trennen solle, antwortete die Versuchsleiterin: „Gehen Sie nach einem Ihnen sinnvoll erscheinenden Kriterium vor.“ oder „Das bleibt Ihnen überlassen.“ Die Sortierexperimente wurden immer von der gleichen Versuchsleiterin durchgeführt. Sie notierte die Ergebnisse des Hierarchischen Sortierens und der Kriteriumsennung für jede Versuchsperson in einem eigenen Protokollbogen.

Ergebnisse des Sortierexperimentes

Auswertungsmethoden Die *Verbaldaten*, d.h. die Nennungen der Sortierkriterien, wurden nach Häufigkeiten ausgewertet. Für die *Sortiererergebnisse* des Hierarchischen Sortierens wurde eine Distanzmatrix erstellt. Die Originaldaten wurden mit der Nonmetrischen Multidimensionalen Skalierung (NMDS) ausgewertet, um die Ähnlichkeitsurteile in einer räumlichen Konfiguration abzubilden (vgl. Gigerenzer 1981, Borg 1981). Zur Prüfung der Anpassungsgüte der räumlichen Darstellung an die Originaldaten wurde mittels des Scree-tests der Stresswert berechnet. Ein signifikanter Stresswert besagt, dass die räumliche Konfiguration die Ähnlichkeitsurteile der Versuchspersonen gut wiedergibt. Zur Berechnung der NMDS und des Scree-tests wurde das Programm MINISSA von Lingoes & Roskam (1975) verwendet. Die Originaldaten der Distanzmatrix wurden darüber hinaus mit der Hierarchischen Clusteranalyse nach Johnson (1967) mit dem Programm Clustan der SPSS-X Version ausgewertet. Es wurde das Minimum-Modell verwendet. Des Weiteren wurden anhand der Originaldaten die Rangkorrelationskoeffizienten von Spearman für die KundInnen- und StudentInnen-Stichprobe berechnet.

Entscheidungsregel zum Testen der Hypothesen Als Signifikanzniveau wurde bei allen Auswertungsverfahren das 5% Niveau festgelegt. Dieses Signifikanzniveau ist für die Stichprobe (51 Personen) als genügend hoch zu bewerten um sicherzugehen, dass die Nullhypothese irrtümlich abgelehnt wird. Ein irrtümliches Beibehalten der Nullhypothese (β -Fehler) würde dem Unternehmer mehr Nachteile (bzgl. Kundenzufriedenheit, Orientierung, eventuell Abverkauf,) bringen als eine irrtümliches Ablehnen.

Ergebnisse der verbalen Nennungen der Sortierkriterien Jede Nennung wurde pro Versuchsperson nur einmal gewertet, auch wenn die Person das Kriterium öfter nannte. Die Nennungen wurden vier Kategorien – Geschmacksbezogene Kriterien, Marken- bzw. Herstellerbezogene Kriterien, Verwendungsbezogene Kriterien, Diverse Kriterien – zugeordnet und ausgewertet (siehe Tab. 7.4-3).

Tabelle 7.4-3. Genannte Sortierkriterien sortiert nach Klassen

Genannte Sortierkriterien	Anzahl	Prozent		
Kekse mit Schokolade	30	17,4 %		
Einfache Kekse	21	12,3 %		
Gefüllte Kekse	12	6,9 %	⇒ Geschmacksbezogene Kriterien Summe: 48,6 %	
Kekse mit Frucht	11	6,0 %		
Kekse mit Nussgeschmack	4	2,3 %		
mit Zucker bestreut	3	2,0 %		
lockere/feste Kekse	2	1,7 %		
Hersteller	11	6,0 %		⇒ Herstellerbezogene Kriterien Summe: 13,8 %
Verpackung	13	7,8 %		
Persönliche Vorliebe	18	10,5 %	⇒ Verwendungsbezogene Kriterien Summe: 16,8 %	
Gesunde Kekse	4	2,3 %		
Verwendungszweck	4	2,3 %		
Kekse für Kinder	2	1,7 %		
Gebäckmischungen	11	6,0 %	⇒ Diverse Kriterien Summe: 20,8 %	
Hochwertiges Gebäck	10	5,3 %		
Gebäckform	9	5,1 %		
Bekanntheit	5	2,7 %		
Waffeln	2	1,7 %		
insgesamt	172	100 %		

Lesebeispiel: 30 von 51 Versuchspersonen nannten mindestens einmal, dass sie nach dem Schokoladenanteil trennen.

Das häufigst genannte Sortierkriterium heißt Schokoladenanteil mit 17,4%, gefolgt von dem Kriterium Einfachheit mit 12,3% und der persönlichen Vorliebe 10,5%. Ordnet man die Nennungen den Kategorien Geschmack, Verwendung, Hersteller und Diverse zu, dann beziehen sich 48,6% der Nennungen auf Geschmackskriterien (z.B. Schoko, Frucht, Nuss, Zucker etc.). Nur 16,8% der Nennungen beziehen sich auf Verwendungskriterien (z.B. für Kinder, Verwendungszweck, Persönliche Vorliebe). 13,8% der Nennungen beziehen sich auf Hersteller bzw. Verpackung und 20,8% der Kriterien sind Diverse, die sich nicht eindeutig den anderen Kategorien zuordnen lassen, sondern das Gebäck nach anderen Kriterien beschreiben (z.B. Gebäckform, Gebäckmischung, Waffeln, etc.). Die Ergebnisse zeigen, dass Hersteller-bezogene Kriterien am seltensten genannt werden. Mehr als 60% der Nennungen beziehen sich auf Geschmacks- bzw. Verwendungskriterien.

Ergebnisse der Hierarchischen Clusteranalyse Abbildung 7.4-3 zeigt die Ergebnisse der Hierarchischen Clusteranalyse. Die Clusteranalyse zeigt deutlich zwei Großcluster: das obere Großcluster mit *allen Süßgebäckartikeln ohne Schokolade*, das untere mit *allen Süßgebäckartikeln mit Schokolade*, unabhängig davon, von welchem Hersteller sie sind. Der Schokoladenanteil ist somit auch beim Sortierverhalten das wichtigste Sortierkriterium: sogar das Mürbteiggebäck Bahlsen Romeo – mit anteilig wenigen kleinen Schokoladenwürfeln – ist dem Süßgebäck mit Schokolade zugeordnet. Innerhalb des „schokoladenfreien“ Großclusters gibt es Subcluster, die unterschiedliche Kekstypen zusammenfassen, z.B. die Subcluster „Butterkekse“, oder „Kekse mit Fruchtgeschmack“. Die Nussblätter sind noch den Schoko-Keks zu zugeordnet, stellen aber den Übergang zu dem schokoladenhaltigen Süßgebäck dar. Innerhalb des „Schoko“ Großclusters gibt es Subcluster nach unterschiedlichem Schokoladenanteil bzw. Qualitätsklassen, wie die Subcluster mit Schoko-Butterkeksen, Gebäckmischungen oder hochwertigem Kaffeegebäck. Auf den ersten zwei Klassifizierungsebenen wird nicht nach Hersteller klassifiziert. Erst innerhalb der Sub- bzw. Subsubcluster ist das Süßgebäck nach Herstellern sortiert.

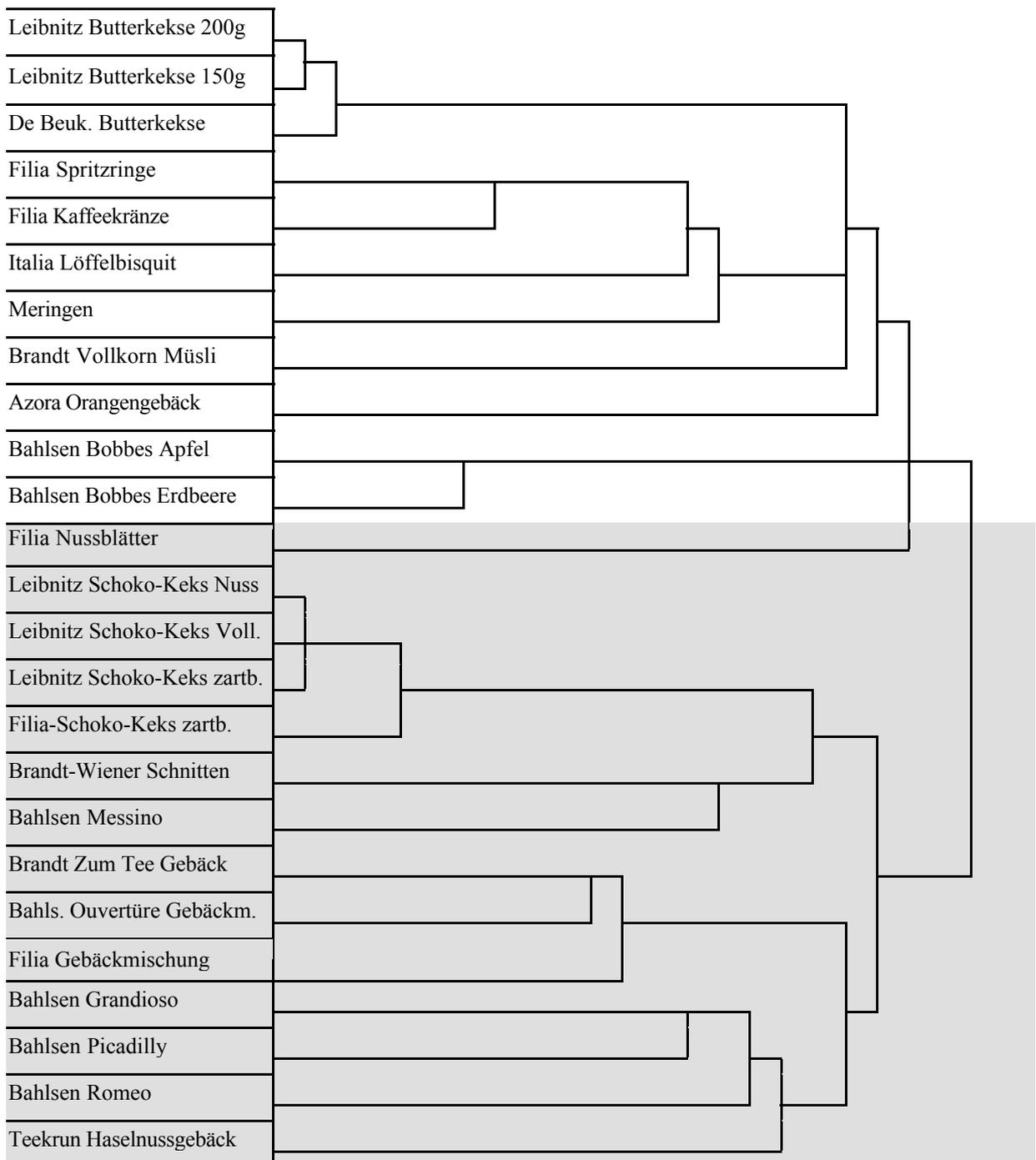


Abbildung 7.4-3. Ergebnisse der Hierarchischen Clusteranalyse von Süßgebäck (Minimum-Modell). Das Großcluster mit schokoladehaltigem Gebäck ist grau hinterlegt.

Ergebnisse des Scree-tests Die NMDS berechnete für Versuchspersonendaten 1- bis 5-dimensionale euklidische Lösungen. Mittels des Scree-tests wurden die Stresswerte für die einzelnen Lösungen berechnet. Die Stresswerte sind aus Abbildung 7.4-4 ersichtlich. Die 2-dimensionale NMDS-Konfiguration liegt im Knick der Kurve und ist somit nach

dem Ökonomieprinzip diejenige Lösung, die die Ähnlichkeitsurteile mit dem geringsten Stress eingepasst hat. Sie hat einen Stresswert von 0.14665, der nach der Tabelle von Gigerenzer (1981, S. 353) mit einer 1%igen Irrtumswahrscheinlichkeit als signifikant gilt. Dieser Stresswert lässt die Schlussfolgerung zu, dass die Versuchspersonen nach ähnlichen Kriterien die Süßgebäck-Artikel beurteilt haben.

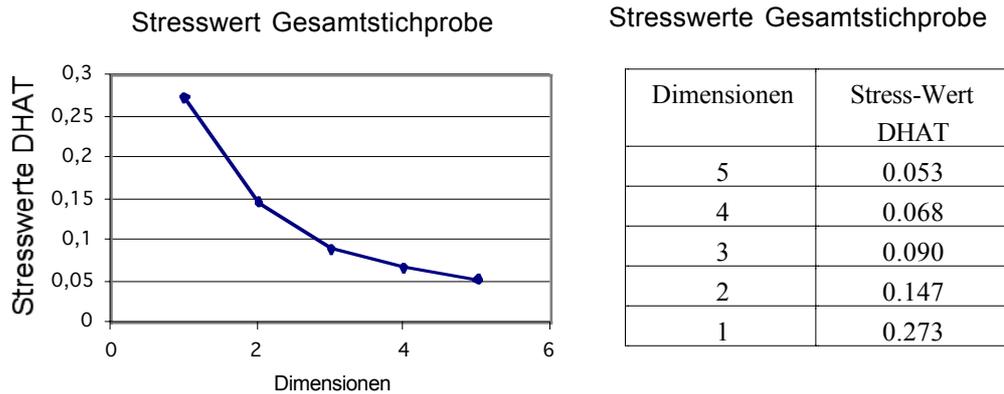


Abbildung 7.4-4. Stresswerte für die Gesamtstichprobe in Diagramm und Tabelle

Ergebnisse der NMDS-Konfiguration Abbildung 7.4-5 zeigt die Ergebnisse der 2-dimensionalen NMDS-Konfiguration für die Gesamtstichprobe.

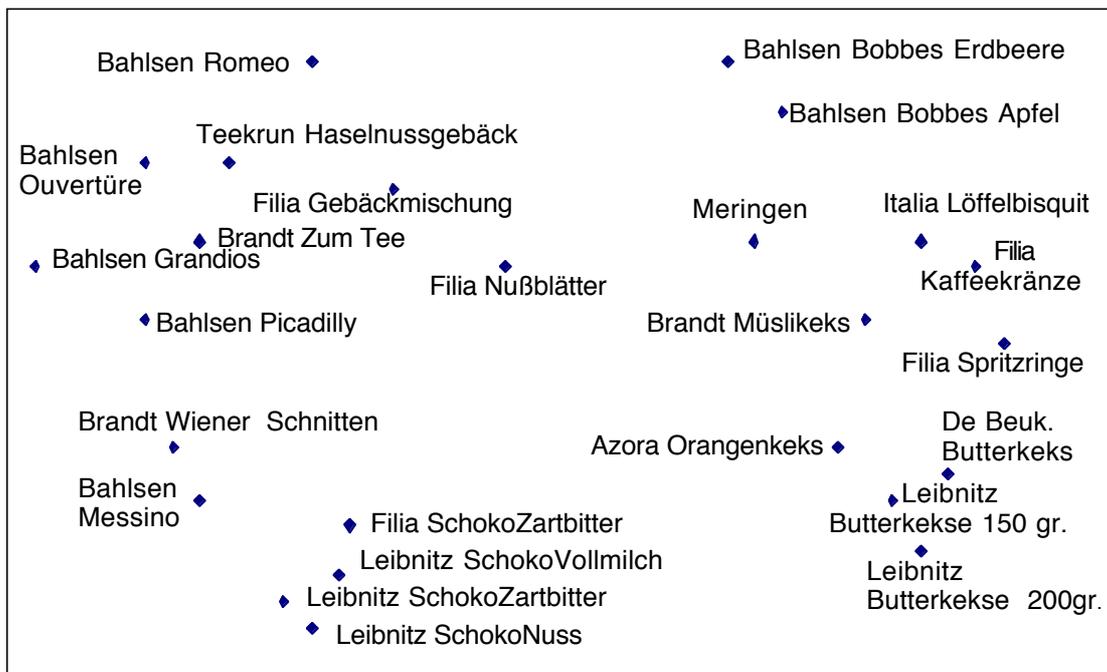


Abbildung 7.4-5. Ergebnisse der 2-dimensionalen NMDS-Konfiguration für Süßgebäck

Die räumlichen Distanzen entsprechen den Ähnlichkeitsbeurteilungen. Die NMDS-Darstellung zeigt die gleichen zwei Gruppen wie die HCA-Ergebnisse: links das schokoladenhaltige Süßgebäck, rechts das schokoladenfreie Süßgebäck. Auf der Verhaltensebene ist der Schokoladenanteil auch das Hauptunterscheidungskriterium. Die Nussblätter sind bei dieser Auswertung – anders als bei der HCA-Auswertung – näher bei dem schokoladenhaltigem Süßgebäck, was ein Indiz für die Unsicherheit der Versuchsperson ist, wie dieses Gebäck zu beurteilen ist und dass es wohl einen Übergang zwischen beiden Gruppen darstellt.

Innerhalb der beiden großen Gruppen zeigen sich Untergruppen: links oben die Gebäckmischungen und das hochwertige Teegebäck, links unten die Schoko-Butterkekse, dazwischen als Übergang die Waffeln und Messino, ein orangengefülltes Schokogebäck.

Rechts oben die hochwertigen schokofreien Kekse mit Fruchtgeschmack. Dann die hochwertigen schokofreien Kekse, wie Kaffeekränze und Spritzringe, Löffelbiscuit, Meringen und ein Müsli-Vollkorn-Keks. Rechts unten die reinen Butterkekse und als Übergang Azora, ein trockener Keks mit leichtem Orangengeschmack. Er ist in der HCA-Auswertung den anderen Fruchtgeschmack-Keksen zugeordnet.

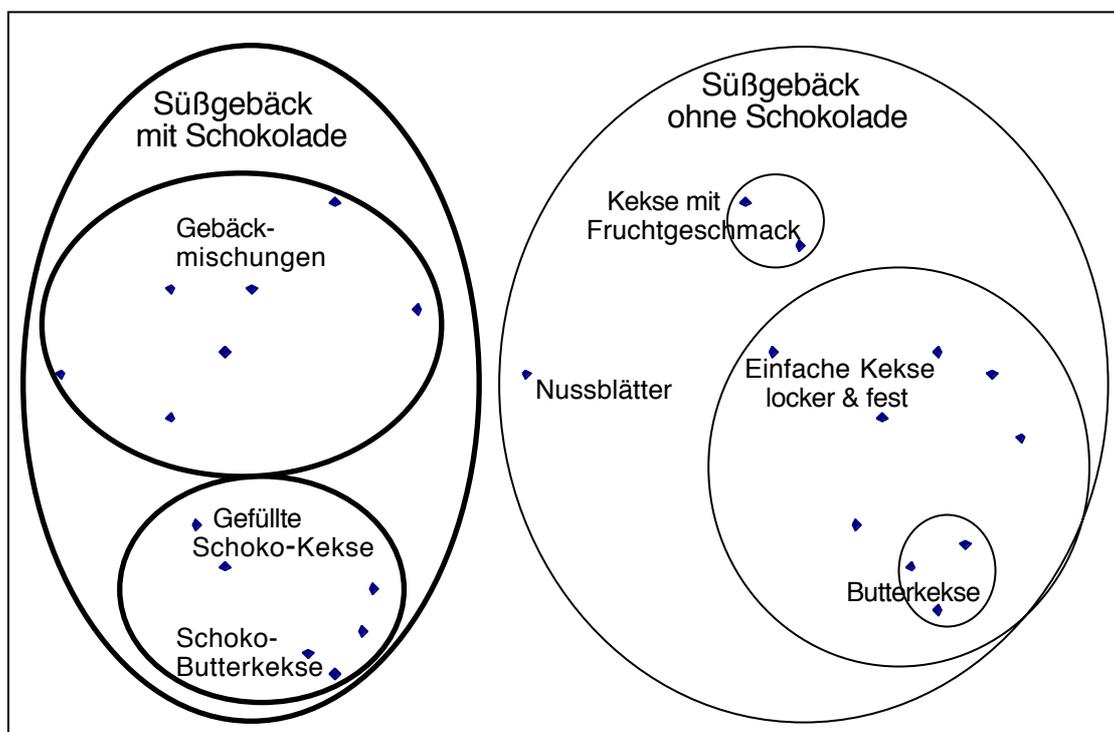


Abbildung 7.4-6. 2-dimensionale NMDS-Konfiguration mit integrierten HCA-Ergebnissen (Kreise)

Integrierte Ergebnisse von HCA und NMDS Abbildung 7.4-6 zeigt die Ergebnisse der Clusteranalyse, die in die NMDS-Konfiguration eingetragen sind. Das, was in der Clusteranalyse in einer Gruppe ermittelt wurde, wird in Kreisen zusammengefasst.

Die integrierte Darstellung zeigt, dass die Ergebnisse von HCA und NMDS bezüglich der semantischen Ähnlichkeit der meisten Süßgebäckartikel übereinstimmen. Unterschiede zeigen beide Auswertungen bei der Beurteilung von Nussgebäck und Azora, den trockenen Keksen mit Orangengeschmack. Unterschiede zeigen sich auch in der Deutlichkeit, in der die Gebäckmischungen von hochwertigem Teegebäck unterschieden werden.

Das Cluster "Süßgebäck mit Schokolade" unterteilt sich zunächst wieder in zwei Subcluster, in Süßgebäck mit viel Schokolade (unten) und Süßgebäck mit weniger Schokolade. Tabelle 7.4-4 listet noch einmal auf, welcher Süßgebäck-Artikel welchen Gruppen zugeordnet ist.

Tabelle 7.4-4. Darstellung der begrifflichen Clusterstruktur für das Süßgebäck

<i>Süßgebäck mit Schokolade</i>	<i>Süßgebäck ohne Schokolade</i>
Süßgebäck mit viel Schokolade	Süßgebäck mit Fruchtgeschmack
Schoko-Butterkekse Schoko Leibniz Vollmilch Schoko Leibniz zartbitter Schoko Leibniz Nuss Filia Schoko+Keks Zartbitter	Kekse mit Frucht ohne Schoko Bahlsen Bobbes Apfel Bahlsen Bobbes Erdbeere
Gefülltes Schokogebäck Brandt Wiener Schnitten (Waffeln) Bahlsen Messino	Brandt Vollkorn Müsli Keks
Hochwertiges Süßgebäck zum Tee	„Einfache“ Kekse
Gebäckmischungen Filia Gebäck Mischung Brandt Zum Tee Gebäckmischung Ouvertüre Teekrun Haselkracker	„Einfache“ Kekse locker&fest Meringen Eierlöffelbiscuit Filia Spritzringe Filia Kaffeekränze
Hochwertiges Gebäck Bahlsen Grandioso Bahlsen Romeo Bahlsen Picadilly	Butterkekse De Beukelaer Butterkekse Leibnitz Butterkekse 250 gr. Leibnitz Butterkekse 150 gr.

Interpretation der Ergebnisse des Sortierexperimentes

Im Folgenden werden die Ergebnisse im Hinblick auf die Hypothesen interpretiert und diskutiert.

Interpretation im Hinblick auf die Hypothese: Personenübergreifende Ähnlichkeitsbeurteilung Die Ergebnisse bestätigen die Hypothese, dass Süßgebäck nach ähnlichen, personenübergreifenden Kriterien beurteilt wird. Der Stresswert für die 2-dimensionale NMDS-Konfiguration (0.147) ist auf dem 1% Niveau signifikant, was darauf schließen lässt, dass die KundInnen in hoch ähnlich geurteilt haben. In der NMDS-Konfiguration und in der HCA zeigen sich eindeutige, stabile Gruppierungen bzw. Cluster. Die Ergebnisse werden dahingehend interpretiert, dass die Personen eine ähnliche Begriffsstruktur bezüglich Süßgebäck haben.

Interpretation im Hinblick auf die Hypothese: Handlungsbezogene Kriterien Die Ergebnisse der HCA und NMDS sowie die Nennungen der Sortierkriterien zeigen eindeutig, dass Geschmackskriterien im Vordergrund stehen. Wichtigstes Kriterium ist, ob ein Süßgebäck einen Schokoladenanteil hat, dann folgen Kriterien, welchen Geschmack es hat und wofür man es verwendet, und erst dann – quasi als letztes Kriterium, wird nach Hersteller unterschieden. Das zentrale Kriterium, "Süßgebäck mit Schokolade" versus "ohne Schokolade" lässt sich als Geschmackskriterium interpretieren, aber auch als Verwendungskriterium: ein schokoladenhaltiges Süßgebäck erfordert – aufgrund von möglichen Schokoladenflecken – anderes Essverhalten und andere Verzehrsituationen. Ein Keks ohne Schokolade ist einfacher in der Handhabung – vor allem für Kinder – und praktischer für Anlässe, in denen Schokoladenflecken vermieden werden sollen. Andere Untergruppen sind eindeutig von Handlungsaspekten bestimmt: Das hochwertige Gebäck und die Gebäckmischungen sind typisches Gebäck für den Kaffee bzw. Tee.

Ordnungskriterien – wie sie die KundInnen im Süßgebäckregal vorfinden (z.B. Hersteller, Verpackung und Preis) – spielen – wenn überhaupt – nur eine untergeordnete Rolle bei der semantischen Beurteilung von Süßgebäck. Erst innerhalb der Sub- bzw. Subsubcluster, wie im Subcluster Kaffeegebäck, werden alle Bahlsen-Kekse zusammengefasst. Oder innerhalb des Subsubclusters Butterkekse sind die Leibnitz-Butterkekse

ähnlicher als das Filia Butterkeks. Hersteller bzw. Marke bestimmen die Begriffsstruktur nicht "primär".

Die Ergebnisse des Sortierexperimentes lassen den Schluss zu, dass die Ähnlichkeitsbeurteilungen und somit die Begriffsstrukturen der KundInnen nicht ein Ergebnis der Realitätswahrnehmung und der Gewöhnung von Supermarkterfahrungen sind. Die Ergebnisse weisen eindeutig darauf hin, dass die Begriffsstrukturen Geschmacks- und Handlungskriterien folgen. Beides sind Kriterien, die auf die Konsequenzen, die Finalität der Handlung, verweisen, im Sinne von: Welche Konsequenzen, welches Ziel will ich mit dem Süßgebäck erreichen? Welches geschmackliche Erlebnis will ich bewirken?

Die Ergebnisse zeigen auch, dass die Handlung „Wo finde ich welchen Artikel im Regal“ die Begriffsstrukturen der KundInnen für Süßgebäckartikel nicht bestimmt. Das Handlungsziel „Kaufen“ hat keine strukturbestimmende Wirkung, sondern nur die Handlungsziele „Geschmack“ oder „Verwendung“ des Süßgebäcks. Die Ergebnisse legen den Schluss nahe, dass die Begriffsstrukturen ein Abbild der Handlungsstruktur unter der ordnenden Maßgabe des Handlungsziels bzw. der Handlungszwecks sind. Und das Ordnung stiftende Handlungsziel scheint nicht das Finden des Gegenstandes im Regal zu sein, sondern die Verwendung!

7.4.3 Evaluations-Studie: Überprüfung der begriffsstruktur-orientierten Anordnung von Süßgebäck auf das Kundenverhalten

Einleitung

Die folgende Untersuchung hatte das Ziel zu überprüfen, ob eine Anordnung, die der Begriffsstruktur der KundInnen entspricht, tatsächlich Vorteile für die KundInnen hat. Es wird angenommen, dass die KundInnen ihre Erwartungen, wo sich ein bestimmter Süßgebäckartikel befindet, auf der Basis ihrer Begriffsstrukturen bilden. Es wird ferner angenommen, dass eine Anordnung der Süßgebäckartikel, die der „Kunden-spezifischen“ Begriffsstruktur entspricht, die KundInnen in ihren Such- und Findeprozessen unterstützt. Es wird dementsprechend folgende Arbeitshypothese aufgestellt: Süßgebäckartikel, die im Regal entsprechend der Begriffsstruktur der KundInnen für Süßgebäck ange-

ordnet sind, erleichtern den KundInnen die Such- und Findeprozesse. Um diese Hypothese zu überprüfen, wurde dieses Feldexperiment durchgeführt.

Operationalisierung der unabhängigen Variablen Die unabhängige Variable wird durch zwei Artikelanordnungen im Regal definiert: die herkömmliche „Hersteller-orientierte“ Artikelanordnung und eine neue „Begriffsstruktur-orientierte“ Artikelanordnung.

„Hersteller-orientierte“ Artikelanordnung Zunächst wurde die übliche Artikelanordnung untersucht, die die KundInnen schon seit Jahren gewohnt waren. Hier ist das wichtigste Sortierkriterium der Hersteller oder die Marke (z.B. Bahlsen, Filia, Brandt etc.), d.h. alle Süßgebäck-Artikel eines Herstellers oder einer Marke stehen nebeneinander im Regal. Diese Regalbelegung wird „Hersteller-orientierte“ Artikelanordnung genannt. Abbildung zeigt 7.4-7 die „Hersteller-orientierte“ Artikelanordnung, die im Feldexperiment untersucht wurde:

1. Regaleinheit		2. Regaleinheit		3. Regaleinheit		4. Regaleinheit	
Filia		Brandt				Bahlsen	
		Div.	Diverse	Diverse			
		Div.	Diverse	Div.			
		Filia		De Beukelaer			

Abbildung 7.4-7. „Hersteller-orientierte“ Regalanordnung, die die KundInnen seit Jahren gewohnt waren (Div. = Süßgebäck von diversen Herstellern)

Begriffsstruktur-orientierte Artikelanordnung Für die „Begriffsstruktur-orientierte“ Artikelanordnung wurden die Ergebnisse des Sortierexperimentes auf die insgesamt 91 Süßgebäckartikel übertragen. Die Regalanordnung berücksichtigt, dass die KundInnen von links an das Regal herantreten, denn dies wurde in einem Prätest im Untersuchungsmarkt (signifikant) ermittelt. Die neue Regalbelegung sollte vier Regaleinheiten (je 1,20 m) umfassen, da das Süßgebäckregal im Untersuchungsmarkt vier Regaleinheiten groß ist. Bei der Realisierung mussten faktische Bedingungen des Regals

(Regalbodenbreite und –höhe) und des Abverkaufs (abverkaufsgerechter Platzbedarf pro Artikel) berücksichtigt werden, sodass aus Platzgründen die "Einfache Kekse locker & fest" (z.B. Müsli Keks von Brandt (Nr.5), Filia Spritzringe (Nr. 11) und Filia Kaffeekränze nicht im Regal 2 als Untergruppe der "Einfachen Kekse" neben der Artikelgruppe "Kinderkekse" liegen können, sondern in das Regal 1 zu "Kekschen mit Frucht ohne Schoko" geordnet werden mussten. Abbildung 7.4-8 zeigt die neue, "Begriffsstrukturorientierte" Regalanordnung:

<i>Süßgebäck ohne Schokolade</i>		<i>Schokoladenhaltiges Süßgebäck</i>	
1. Regaleinheit	2. Regaleinheit	3. Regaleinheit	4. Regaleinheit
Waffeln	Butterkekse Einfache, trockene Kekse Kinderkekse	Gebäck- mischungen	Schokowaffeln
Kekse mit Frucht Kekse locker & fest		Süßgebäck mit unterschiedlichem Schokoladenanteil	Gefülltes Schoko- gebäck Schoko- Butterkekse

Abbildung 7.4-8. "Geschmacksorientierte" Regalanordnung für die 2. bis 7. Untersuchung.

Operationalisierung der abhängigen Variablen Als abhängige Variablen wurden beobachtbare Verhaltensvariablen definiert. In einem Prätest mit 205 Personen in einem Feineberg-Supermarkt in Memmingen wurde eine Verhaltensbeobachtung durchgeführt, um diejenigen Beobachtungskriterien zu definieren, die Such- und Findeverhalten indizieren. Außerdem wurde das Beobachtungsprotokoll entwickelt. Das Verhalten sollte ohne Videobeobachtung erfassbar sein, da Videobeobachtungen Technik- und vor allem sehr auswertungsintensiv sind.

Als Beobachtungsprotokoll wurde eine maßstabsgerechte Abbildung des Regals verwendet. Vor dem realen Regal wurden am Boden für die KundInnen unauffällige Markierungen angebracht, die auch auf dem Protokollbogen eingezeichnet waren, um das Verhalten exakt protokollieren zu können. Der Prätest ergab die Beobachtungskriterien (Tab.7.4-5): Aufenthaltsdauer, Anzahl der Stopps, Wegstrecke, Anzahl der berührten Produkte, Absatz, Umsatz.

Tabelle 7.5-5. Abhängige Variablen für das Feldexperiment Süßgebäck

Abhängige Variablen	Operationalisierung
Aufenthaltsdauer in Sekunden	Gemessen wird die Zeit in Sekunden, die die KundIn vor dem Regal im markierten Bereich verweilt. Die Zeitmessung stoppt, wenn sie den Bereich verlässt. Kommt sie zurück, läuft die Zeit weiter.
Anzahl der Stopps ja/nein	Gemessen wird, wenn eine KundIn stehen bleibt. Die Anzahl der Stopps der KundIn wurde in den maßstabgerechten Protokollbogen eingezeichnet.
Wegstrecke in Meter	Der Weg der KundIn wurde in dem maßstabgerechten Protokollbogen eingetragen und abgemessen. Die Bodenmarkierungen dienten als Orientierungspunkte.
Anzahl der berührten Produkte ja/nein	Ein Produkt gilt als berührt, wenn die KundIn es mit der Hand berührt, unabhängig davon, ob sie es dann anhebt, kauft oder zurückstellt.
Absatz in Stückzahl	Gemessen wird die Anzahl der gekauften Produkte. Ein Produkt gilt als gekauft, wenn die KundIn es in ihren Wagen legt und dort belässt.
Umsatz in DM	Für die Süßgebäck-Produkte, die in den Wagen gelegt wurden, wurde der Preis ermittelt.

Aufgrund von Plausibilitätsüberlegungen und Gesprächen mit einigen KundInnen wurde davon ausgegangen, dass sich ein besseres Such- und Findeverhalten wie folgt darstellt: die Aufenthaltsdauer ist kürzer, die Anzahl der Stopps ist geringer, die Wegstrecke ist kürzer, die Anzahl der berührten Produkte ist geringer, die Anzahl der gekauften Produkte und der Umsatz bleiben gleich. Es muss in einer Kundenbefragung abgeklärt werden, wie die KundInnen diese Verhaltenstendenzen selber beurteilen und ob sie es tatsächlich als Ausdruck eines besseren Such- und Findeverhaltens beurteilen.

Hypothese Im Feldexperiment wurde somit folgende Hypothese getestet: Das Süßgebäck-Regal mit der „Begriffsstruktur-orientierten“ Artikelanordnung hat eine signifikante Auswirkung auf das Kundenverhalten: die Aufenthaltsdauer ist kürzer, die Anzahl der Stopps ist geringer, die Wegstrecke ist kürzer, die Anzahl der berührten Produkte ist geringer, die Anzahl der gekauften Produkte und der Umsatz bleiben im Vergleich zu einer Hersteller-orientierten Artikelanordnung gleich.

Methode

Erhebungsmethode und Dokumentation Als Erhebungsmethode diente eine *systematische Verhaltensbeobachtung*. Beobachtungskriterien und Beobachtungsprotokoll wurden wie beschrieben in einem Prätest mit 205 Personen entwickelt. Zusätzlich zu den abhängigen Variablen wurden pro Person das Geschlecht und das geschätzte Alter erfasst. Für jede der beiden Artikelordnungen wurde ein Beobachtungsbogen erstellt. Auf dem Beobachtungsbogen war das Regal mit jedem Artikel abgebildet. Zusätzlich war der Boden vor dem Regal mit für die Versuchspersonen kaum sichtbaren Rasterpunkten markiert, damit der von der Person begangene Weg genau protokolliert werden konnte. Das Kundenverhalten wurde pro Kunde in einem Beobachtungsbogen festgehalten. Zusätzlich wurden das Geschlecht und das geschätzte Alter erfasst. Um das Kundenverhalten eindeutig interpretieren zu können, wurde eine *standardisierte Befragung* durchgeführt, um zu ermitteln, was Aufenthaltsdauer, Stopps, Wegstrecke, Anfassen der Produkte für die KundInnen bedeutet. Für die Befragung wurde ein geschlossener Fragebogen entwickelt.

Untersuchungsort und Durchführung Die Untersuchungen wurden im Feneberg Markt in Memmingen durchgeführt. Der Feneberg-Markt in Kempten, in dem das Sortierexperiment durchgeführt wurde, wurde nicht gewählt, um die "Begriffsstruktur-orientierte" Artikelanordnung an einer anderen Stichprobe zu testen. Die Untersuchung fand an sieben Messzeitpunkten statt. Am 1. Messzeitpunkt wurde das Kundenverhalten bei der herkömmlichen, "Hersteller-orientierten" Artikelanordnung als Vergleichswert erhoben. Danach wurde das Regal entsprechend der "Begriffsstruktur-orientierten" Artikelanordnung umgeräumt und blieb dann für den gesamten weiteren Untersuchungszeitraum von acht Wochen unverändert. Der 2. Messzeitpunkt fand 14 Tage nach der Umräumung statt, damit die KundInnen Zeit hatten, sich an die neue Regalordnung zu gewöhnen. Die weiteren Messzeitpunkte fanden in einem Abstand von jeweils 14 Tagen statt, jeweils an einem Freitag. Dies ist der Wochentag mit der höchsten Kundenfrequenz im Supermarkt (ca. 2500 Kundinnen pro Tag). Die Beobachtungen fanden vormittags und nachmittags statt. Am letzten Messzeitpunkt wurde die standardisierte Befragung durchgeführt. Tabelle 7.4-6 gibt einen Überblick über die Durchführung.

Tabelle 7.4-6. Versuchsplan für das Feldexperiment

Messzeitpunkte	Untersuchte Regalordnung	KundInnen
1. Kundenbeobachtung Baseline-Erhebung Freitag 11. Juni vormittags	herkömmliche Ordnung "Hersteller-orientiert"	239
2. Kundenbeobachtung Freitag 18. Sept. vormittags	"Begriffsstruktur-orientiert"	137
3. Kundenbeobachtung Freitag 18. Sept. nachmittags	"Begriffsstruktur-orientiert"	129
4. Kundenbeobachtung Freitag 02. Okt. vormittags	"Begriffsstruktur-orientiert"	140
5. Kundenbeobachtung Freitag 02. Okt. nachmittags	"Begriffsstruktur-orientiert"	190
6. Kundenbeobachtung Freitag 16. Okt. vormittags	"Begriffsstruktur-orientiert"	104
7. Kundenbeobachtung Freitag 16. Okt. nachmittags	"Begriffsstruktur-orientiert"	108
Standardisierte Befragung Freitag 16. Okt. nachmittags	Fragebogen	24

Die Beobachtungen wurden immer von der gleichen Versuchsleiterin durchgeführt. Die Versuchsleiterin trug einen weißen Feneberg-Verkaufskittel und hielt ein Clipboard mit den Beobachtungsbögen in der Hand. Sie sah somit wie eine Verkäuferin aus, die Ware disponiert. Die Befragung fand mündlich statt und die Versuchsleiterin notierte die Antworten für jede Person in einem Fragebogen.

Versuchspersonen Insgesamt wurden 1047 Personen beobachtet (846 Frauen und 201 Männer). Von diesen 1047 Personen kauften 300 KundInnen (= 28,6 %) Süßgebäck. Das Süßgebäckregal war so im Markt gelegen, dass die KundInnen auf dem Weg zur Kasse fast unweigerlich daran vorbeikamen. Die Personen wurden aufs Geratewohl ausgewählt. Tabelle 7.4-7 enthält die Daten für die 7 Messzeitpunkte:

Um zu prüfen, wie weit die Stichproben vergleichbar sind, wurden Konfidenzintervalle basierend auf der Stichprobe 1 berechnet. Das 95% Konfidenzintervall wird für die Merkmale Männer/Frauen sowie Käufer/Nichtkäufer in keiner der Stichproben überschritten. Stichproben 4 und 5 weichen hinsichtlich der Altersverteilung von den übrigen Stichproben ab. Stichprobe 4 weist mehr 50-59 jährige, Stichprobe 5 mehr 40-49 jährige Personen auf. Dies kann durch Schätzfehler der Versuchsleiterin beeinflusst sein.

Tabelle 7.4-7. Versuchspersonendaten für das Feldexperiment
"Begriffsstruktur-orientierte" Regalordnung

Untersuchungszeitpunkte	1. Base-line	2.	3.	4.	5.	6.	7.	Summe
Anzahl der beobachteten Personen								
Personen	239	137	129	140	190	104	108	1047
Verteilung Frauen und Männer								
Frauen	189	118	93	105	154	92	94	845
Männer	50	19	36	35	36	12	14	202
Süßgebäck-Käufer und Nichtkäufer								
Süßgebäck-Käufer	70	35	37	47	47	33	31	300
Nichtkäufer	169	102	92	93	143	71	77	747
Prozentualer Anteil Käufern und Nichtkäufer								
Süßgebäck-Käufer	29%	26%	29%	34%	25%	32%	29%	
Nicht-Käufer	71%	74%	71%	66%	75%	68%	71%	
Anteil an Käufern nach Geschlechtern								
Frauen	79%	86%	73%	75%	81%	88%	88%	
Männer	21%	14%	27%	25%	19%	12%	13%	
Anteil an Nicht-Käufern nach Geschlechtern								
Frauen	76%	74%	65%	68%	68%	73%	64%	
Männer	24%	26%	35%	32%	32%	27%	36%	

Ergebnisse des Feldexperimentes

Die beobachteten Verhaltensdaten (Zeit, Weg, Anzahl der Stopps, Berührungen, Absatz und Umsatz) wurden nach Häufigkeiten ausgewertet und es wurden die Mittelwerte berechnet. KundInnen, die kein Süßgebäck kauften (Nicht-KäuferInnen), wurden nur im Hinblick auf die Anzahl der berührten Produkte ausgewertet. Tabelle 7.4-8 zeigt pro Untersuchungszeitpunkt die Anzahl der Versuchspersonen sowie für jede abhängige Variable die summierten Werte und die errechneten Mittelwerte für jede Stichprobe pro Messzeitpunkt. Mittels des t-Tests wurden die Ergebnisse der "Hersteller-orientierten" Regalanordnung (1. Untersuchung) und der "Begriffsstruktur-orientierten" Regalanordnung (2.-7. Untersuchung) auf Unterschiede geprüft.

Tabelle 7.4-8. Ergebnisse der Untersuchungen für Süßgebäck-KäuferInnen
Summenwerte und Mittelwerte

Treatment 1: "Hersteller-orientierte" Regalordnung		Treatment 2: "Begriffsstruktur-orientierte" Regalordnung					
Untersuchungs- zeitpunkte	1. Base- line	2.	3.	4.	5.	6.	7.

Aufenthaltsdauer der Süßgebäck-KäuferInnen vor dem Regal

Summe in Sekunden	3153	1542	1444	2041	2046	880	948	*sign. 5%
Mittelwert in sek.	45,0	42,8	40,1	43,4	42,8	26,7	30,6	

Wegstrecke der Süßgebäck-KäuferInnen vor dem Regal

Summe in Meter	759	334	382	512	481	264	276	*sign. 5%
Mittelwert in m	8,3	7,1	7,9	8,2	7,5	6,0	6,8	

Anzahl an Stopps der Süßgebäck-KäuferInnen

Summe der Stopps	168	74	86	88	78	46	39	*sign. 5%
Mittelwert	2,4	2,1	2,3	1,9	1,6	1,4	1,3	

Anzahl der berührten Produkte der Süßgebäck-KäuferInnen

Summe der berührten Produkte	29	10	2	19	15	9	10	*sign. 5%
Mittelwert in Stück	0,41	0,28	0,06	0,40	0,30	0,27	0,32	

Anzahl der berührten Produkte der Nicht-KäuferInnen

Summe der berührte Produkte	2	1	1	3	9	4	6	*sign. 5%
Mittelwert in Stück	0,01	0,01	0,06	0,01	0,03	0,06	0,05	

Absatz der Süßgebäck-KäuferInnen

Summe der gekaufte Produkte	109	52	56	87	83	49	40
Mittelwert in Stück	1,56	1,4	1,56	1,85	1,77	1,49	1,29

Umsatz der Süßgebäck-KäuferInnen

Umsatzssumme in DM	247,26	113,49	127,39	201,84	195,70	121,63	90,11
Mittelwert in DM	3,53	3,24	3,44	4,29	4,16	3,69	2,91

Annäherung ans Regal

...von rechts	32	15	17	16	25	17	7
...von links	38	20	20	31	22	16	24

Der t-Test zeigt signifikante Unterschiede auf dem 5% Niveau für die Verhaltensvariablen Zeit, Weg, Anzahl der Stopps, Anzahl der berührten Artikel. Die Süßgebäck-KäuferInnen halten sich signifikant kürzer vor dem Regal auf, sie legen signifikant weniger Wegstrecke vor dem Regal zurück, sie bleiben weniger oft stehen und sie fassen weniger Artikel an. Der Absatz und der Umsatz verändern sich nicht signifikant. Die Nicht-KäuferInnen berühren signifikant mehr Produkte. Um zu prüfen, ob es einen Zusammenhang zwischen den Variablen gibt, wurden Rangkorrelationen berechnet. Tabelle 7.4-9 zeigt die Rangkorrelationskoeffizienten.

Tabelle 7.4-9. Rangkorrelationen zwischen den abhängigen Variablen

	Zeit	Weg	Stopps	Berühr.	Absatz
Wegstrecke	.88*				
Zeitdauer	.57	.69			
Berührungen	.70	.54	-.06		
Absatz	.71	.76	.43	.29	
Umsatz	.37	.36	-.01	.21	.78

* signifikant 5 % Niveau

Zeit und Weg korrelieren signifikant auf dem 5% Niveau: Je mehr Wegstrecke eine KundIn geht, desto länger hält sie sich vor dem Süßgebäck-Regal auf. Andere Variablen wie z.B. die Aufenthaltszeit und der Absatz (Anzahl der gekauften Produkt) oder die Aufenthaltszeit und die Anzahl der berührten Produkte korrelieren hoch (0.70 bzw. 0.71), aber nicht signifikant.

Kundenbefragung Um die Interpretation der oben dargestellten Ergebnisse empirisch abzusichern, wurde eine Nachuntersuchung durchgeführt. Ziel war es, abzusichern, was eine kurze Verweildauer, eine kurze Wegstrecke, wenige Stopps und wenige Berührungen für die KundInnen bedeuten. Ob dies z.B. als Interessenlosigkeit gegenüber dem Süßgebäck-Regal zu deuten ist, oder ob dies Ausdruck eines orientierungsfreundlichen, übersichtlichen Regals ist, in dem sich die KundInnen gut zurechtfinden.

Es wurden 24 KundInnen (22 Frauen, 2 Männer) mit einem Fragebogen befragt, die Antworten waren geschlossen und die Befragten konnten sich zwischen 2 Antwortalternativen entscheiden. Tabelle 7.4-10 zeigt die Ergebnisse der Befragung. Die KundInnen

hatten es sehr eilig und hatten wenig Bereitschaft, den Fragebogen auszufüllen. Da nicht alle Personen jede Frage beantworteten, ergeben die Summen nicht immer 24.

Tabelle 7.4-10. Ergebnisse der standardisierten Befragung zur Interpretation der Verhaltensvariablen

Ergebnisse der Standardisierten Befragung von 24 KundInnen (22 Frauen, 2 Männer) Nicht alle Personen beantworteten jede Frage		Antwort- häufigkeit
Aufenthalts- dauer	Wenn ich lange vor dem Regal stehe, heißt das für mich: o Ich finde mich schlecht zurecht.	12
	o Das Regal ist interessant, ich will es mir ansehen.	11
Weg- strecke	Wenn ich oft auf und ab gehe, heißt das für mich: o Das Regal interessiert mich, ich will mir alle Produkte ansehen.	06
	o Ich muss erst alles ansehen, damit ich weiß, wo ich das, was ich suche, auch finde.	17
Anzahl der Stoppes	Wenn ich oft stehen bleibe, heißt das für mich: o Ich finde mich schlecht zurecht.	6
	o Ich will mir in Ruhe alles ansehen.	16
Anzahl der berührten Produkte	Wenn ich viele Artikel berühre, heißt das für mich: o Ich vergleiche die Produkte bevor ich sie kaufe.	14
	o Ich kann das, was ich suche, nicht finden.	4
Beurteilung von Regal- eigenschaf- ten	Welche Regal-Eigenschaften sind Ihnen am wichtigsten?	
	o Es soll interessant sein.	0
	o Es soll übersichtlich sein.	15
	o Artikel gleicher Hersteller sollen zusammenliegen.	02
	o Preisvergleich soll möglich sein.	06
o Was gleich schmeckt, soll zusammenliegen.	0	

Interpretation der Kundenbefragung Langes Verweilen vor dem Regal wird nicht eindeutig beurteilt: Es kann daher sowohl ein Ausdruck des „sich-nicht-Zurechtfindens“, aber auch ein Ausdruck des Interesses sein. Anders ist es mit der Länge der Wegstrecke: Häufiges Hin- und Hergehen wird als Ausdruck des sich „nicht-Zurechtfindens“ gedeutet. Häufiges Stoppen vor dem Regal wird eher als ein Bedürfnis „sich alles anzusehen“ gedeutet und weniger als ein „schlechtes Zurechtfinden“. Häufiges Berühren von Produkten wird eher als ein „Produktvergleich“ interpretiert und weniger als ein „nicht Zurechtfinden“. Die letzte Frage zeigt die Präferenz der meisten KundInnen: Das Regal

soll übersichtlich sein (15). Die Möglichkeit für einen Preisvergleich wird von weitaus weniger KundInnen als wichtig erachtet (6). Nur 2 KundInnen ist es wichtig, dass alle Artikel von einem Hersteller nebeneinander sortiert sind. Keiner KundIn ist es wichtig, dass das Regal interessant ist. Dies weist darauf hin, dass die KundInnen an Übersichtlichkeit interessiert sind. Keiner Kundin ist es wichtig, dass die Artikel entsprechend der Geschmackslogik sortiert sind. Dies weist darauf hin, dass die KundInnen dies nicht mit einer Sortierung entsprechend des Geschmacks gleichsetzen. Das kann mehrere Gründe haben: Entweder, finden die KundInnen eine Geschmacksortierung nicht übersichtlich oder aber sie können nicht konkretisieren, was das Kriterium für eine übersichtliche Sortierung ist bzw. sie würden es nicht so benennen.

Interpretation der Ergebnisse des Feldexperimentes

Die Ergebnisse des Feldexperimentes zeigen, dass die "Begriffsstruktur-orientierte" Regalbelegung das Verhalten der KundInnen im Vergleich zur „Hersteller-orientierten“ Regalbelegung signifikant beeinflusst: Die Süßgebäck-KäuferInnen verbringen weniger Zeit vor dem "Begriffsstruktur-orientierten" Regal. Sie legen weniger Wegstrecke vor dem Regal zurück, sie bleiben weniger oft stehen und fassen weniger Produkte an. Die NichtkäuferInnen berühren signifikant mehr Produkte. Der Absatz und der Umsatz verändern sich nicht signifikant.

Die Ergebnisse können – anhand der Kundenbefragung – unterschiedlich interpretiert werden: einmal, dass die Such- und Findeprozesse schneller ablaufen, da sich die KundInnen besser zurechtfinden und zum anderen, dass das „Begriffsstruktur-orientierte“ Regal aber auch uninteressanter ist und zu weniger Erkundung einlädt. Beide Interpretationsmöglichkeiten widersprechen sich jedoch im Grunde nicht: eine Ordnung, die übersichtlich ist, erweckt weniger Interesse, erfordert aber auch weniger Orientierungsaufwand als eine unübersichtliche Anordnung. Eine übersichtliche Anordnung erlaubt es aber, sich schneller zu orientieren und die Auswahl schneller zu tätigen.

Die kürzere Verweildauer, die kürzere Wegstrecke, die geringere Anzahl der Stopps und der berührten Produkte der Süßgebäck-KäuferInnen bei gleichem Abverkauf und Umsatz bestätigen diesen Schluss. Die KundInnen können also bei der „Begriffsstruk-

tur-orientierten“ Anordnung mit einem geringeren Zeitaufwand und weniger Suchaktivitäten ihren Einkauf tätigen : Sie müssen nicht so viel hin- undher laufen, um sich zu orientieren. Sie müssen nicht oft stehen bleiben, um sich alles in Ruhe anzusehen, und sie müssen weniger oft Produkte berühren, bevor sie sich für eines entscheiden können. Die KundInnen müssen für ihren Einkauf weniger Zeit und Suchaktivitäten aufwenden.

Die Ergebnisse bestätigen die Annahme, dass eine Artikelanordnung im Regal, die der Begriffsstruktur entspricht, die Such- und Findeprozesse der KundInnen unterstützt. Und die Ergebnisse zeigen, dass die KundInnen von einer „Begriffsstruktur-orientierten“ Regalanordnung nicht dahingehend „verführt“ werden, dass sie signifikant mehr kaufen. Die KundInnen profitieren also eindeutig, wenn die Artikel in einem Regal entsprechend ihrer Begriffsstruktur angeordnet sind.

7.4.4 Gesamtschau und Interpretation der Ergebnisse

Die Untersuchung hatte ihren Ausgangspunkt in den beiden theoretisch fundierten Annahmen, dass a) KundInnen ihre Erwartungen darüber, wo sie welches Süßgebäck im Regal finden auf der Basis ihrer Begriffsstrukturen über Süßgebäck bilden und dass b) eine Anordnung von Süßgebäck, die der Begriffsstruktur der KundInnen entspricht, deren Such- und Findeverhalten verbessert.

Basierend auf der Begriffstheorie von Klix (1971, 1984, 1992), Hoffmann (1983, 1986) et al. wurden Vorhersagen abgeleitet, welchen Regeln und Gesetzmäßigkeiten die Begriffsstrukturen (operationalisiert als Ähnlichkeitsbeurteilungen) bezüglich Süßgebäck folgen müssten: nämlich, dass sich für eine KundInnen-Stichgruppe eine personenübergreifende Ähnlichkeitsbeurteilung für Süßgebäck-Artikel finden lässt und dass diese Ähnlichkeitsbeurteilung handlungsbezogenen Klassifizierungskriterien wie Geschmack bzw. Verwendung und nicht marken- bzw. herstellerbezogenen Kriterien folgt.

Die *Ergebnisse des Sortierexperimentes* bestätigten diese theorie-basierten Vorhersagen: es ließ sich eine personenübergreifende Ähnlichkeitsbeurteilung für Süßgebäck bei den KundInnen erheben. Und die Ähnlichkeitsbeurteilungen folgten nicht marken- oder her-

stellerbezogenen Kriterien, sondern eindeutig Geschmacks- und Verwendungskriterien. (Das wichtigste Klassifizierungskriterium ist der Schokoladenanteil. Innerhalb des schokoladenhaltigen und schokoladenfreien Süßgebäcks klassifizieren die KundInnen nach verwendungs- und geschmacksbezogenen Kriterien.)

Dieses Ergebnis entspricht der theorie-basierten Vorhersage und es widerspricht eindeutig dem Klassifizierungssystem, dass die KundInnen seit Jahren im Regal gewöhnt waren. Das Süßgebäck war traditionell nach Marken und Herstellern im Regal angeordnet. Die Ähnlichkeitsbeurteilungen der KundInnen sind nicht ein Abbild der faktischen Süßgebäckanordnung im Regal. Die gewohnte Regalanordnung hat sich nicht als klassifizierungsrelevant für die Begriffsstruktur ausgewirkt. Klassifizierungsrelevant sind vielmehr handlungsbezogene Aspekte, die auf den Geschmack und die Verwendung (Finalität) zielen. Oder um es provokativer zu sagen: die Begriffsstruktur ist nicht ein Abbild der Realität, sondern ein Abbild der Handlungsstruktur unter der ordnenden Maßgabe des Handlungszieles bzw. -zwecks.

Diese Aussage behauptet nicht, dass KundInnen sich nicht daran erinnern können, wo das eine oder andere Süßgebäck in ihrem gewohnten Supermarkt zu finden ist. Die KundInnen wurden nicht gefragt „Können Sie beschreiben, wo in Ihrem gewohnten Supermarkt die Butterkekse im Regal liegen?“ Wahrscheinlich hätte sogar ein Teil der KundInnen diese Frage richtig aus dem Gedächtnis oder konkreter gesagt: aus ihrem episodischen (raum-zeitlich spezifischem) Wissen abgerufen. Nein, die Ergebnisse sagen vielmehr, dass sich diese über lange Zeit vorgefundene und gewohnte Süßgebäckanordnung im Regal nicht in den grundlegenden Klassifizierungsdimensionen der Ähnlichkeitsbeurteilung wiederfindet. Die Ergebnisse bestätigen die Annahme, dass die Ähnlichkeitsbeurteilungen handlungsbezogenen Kriterien, insbesondere Finalitätskriterien folgen.

Die *Ergebnisse des Feldexperiments* bestätigen die Annahme, dass eine Anordnung von Süßgebäck, die der Begriffsstruktur der KundInnen entspricht, deren Such- und Finderverhalten verbessert. Bei einer Begriffsstruktur-orientierten Artikelanordnung im Regal brauchen die KundInnen weniger Zeit, eine kürzere Wegstrecke, weniger Stopps und weniger Produktberührungen, um ihren Einkauf zu tätigen als bei einer marken- bzw.

herstellerorientierten Anordnung. Die Ergebnisse zeigen, wie bedeutsam, effektiv und kundenfreundlich eine Regalanordnung ist, die den Begriffsstrukturen der KundInnen entspricht. Und die Ergebnisse lassen überdies den Schluss zu, dass in der Such- und Orientierungssituation vor dem Regal spontan auf begrifflich-semantisches Wissen zurückgegriffen wird. Es wird vermutet, dass ein Rückgriff auf episodisches Wissen, im Sinne der Frage "Wo sind die Butterkekse in diesem Supermarkt?", einen anderen Verarbeitungsprozess darstellt, der kognitiv aufwändiger ist bzw. sich nicht so spontan vollzieht wie der Rückgriff auf begriffliches Wissen. Oder aber es steht kein bzw. wenig episodisches Wissen für die spezifische Anordnung der Süßgebäck-Artikel zur Verfügung. Sprich: die KundInnen lernen einfach nicht eine gesamte Anordnung von Süßgebäck-Artikeln im Regal. (Was nicht heißt, dass sie sich die Position des einen oder anderen Artikel nicht merken.) Dies würde Gedächtnisexperimenten entsprechen, die zeigen, dass „sinnfreie“ Chunks, (sinnfreie Anordnungen) sehr schwer zu erlernen sind.

Bewertung des Problemlösungsweges Betrachtet man abschließend die Ergebnisse vor dem Hintergrund des ganzen Problemlösungsweges, dann ist das Vorgehen als erfolgreich zu bewerten. Mittels dieses theorie-basierten, empirischen Vorgehens lässt sich tatsächlich eine Anordnung finden, die das Such- und Findeverhalten der KundInnen verbessert.

Bewertung aus der Sicht des Handelsunternehmens Bleibt noch die Frage zu stellen, ob sich der zeitliche und finanzielle Aufwand für ein Handelsunternehmen lohnt. Schließlich bedeutet dies Kosten für die Untersuchungen, um die Begriffsstruktur der KundInnen zu erheben, aber auch Personalkosten für Umräumarbeiten in den Regalen der Lebensmittelmärkte. Was hat also ein Handelsunternehmen von einer "Kundinnen-orientierten" Regalordnung – bei der KundInnen sich schneller und besser zurechtfinden, aber nicht mehr kaufen? Hans Feneberg, der erste Auftraggeber für Studien dieser Art, erklärte mir dies mit dem einfachen Satz: Die "Produkte", die das Feneberg-Handelsunternehmen "anbietet und verkauft" sind die Feneberg-Märkte und nicht einzelne (Süßgebäck-)Artikel. Sein Interesse ist es, dass das Verkaufsflächenkonzept als Ganzes stimmt. Wenn das Konzept stimmt, dann sind die KundInnen zufrieden, dann kommen sie wieder und dann kaufen sie auch wieder. Und über die größere Kundenbindung ist die Existenz des Unternehmens im harten Wettbewerb mit anderen Lebensmittel-Handelsunternehmen gesichert.

Inwiefern seine Rechnung aufgegangen ist, kann hier – aufgrund von Geheimhaltungsvereinbarungen – nicht belegt werden. Zumindest wurde für alle Sortimente und Warengruppen eine ”kundinnenorientierte” Regalbelegung ermittelt und im Markt umgesetzt. Ich erhielt auch von anderen Einzelhandelsketten Aufträge zur Entwicklung ”kundInnenorientierter” Sortiments- und Regalbelegungen. Verkaufsflächenplaner großer deutscher Lebensmittel-Handelsunternehmen „pilgerten“ später ins Allgäu, um zu studieren und zu kopieren, was die „Kemptner“ da gemacht haben. Eine Einzelhandelskette, die sich die Psychologen sparen wollte, machte dabei folgenden Fehler: sie sortierten auch die Schokolade-Tafeln nach der ”Geschmackslogik”: Vollmilch-Schokoladen zu Vollmilch-Schokoladen, Haselnuss-Schokoladen zu Haselnuss-Schokoladen usw. Doch ausgerechnet Schokolade kauft die KundIn zuerst dem nach Hersteller: eine Milka ist eben eine Milka – und keine Ritter Sport!

7.4.5 Beurteilung der verwendeten Methoden

Methoden des Sortierexperimentes Die Auswahl der 25 Süßgebäck-Artikel mittels Zufallsliste ist sehr einfach, Zeit sparend und wenig aufwändig. Glücklicherweise erwies sich die Auswahl für die Umsetzung als nicht problematisch. Die gefundenen Klassifizierungskriterien ließen sich ohne große Probleme auf das Gesamtsortiment mit 91 Artikeln übertragen. Für Untersuchungen an anderen Objektklassen sollte geprüft werden, ob es nicht notwendig ist, die Auswahl dahingehend zu kontrollieren, dass die ausgewählten Artikel alle wichtigen Objektkategorien repräsentieren. Die Methode des Hierarchischen Sortierens ist für die Versuchspersonen nicht schwierig, aber langwierig. Die Methode erwies sich als aufwändig in der Erstellung der Distanzmatrixen.

Methoden des Feldexperimentes Die Beobachtungskategorien und der Beobachtungsbogen erwiesen sich einfach in der Durchführung und in der Auswertung. Probleme brachte der Fragebogen: die KundInnen hatten es eilig und sie zeigten wenig Geduld und Ausdauer. Wenige KundInnen beantworteten jede Frage. Es ist daher abzuwägen, ob eine Fragebogen-Befragung notwendig ist. Wenn ja, sollten wenige nondirektive Interviews stattfinden, um die wesentlichen Fragen herauszuarbeiten. Oder aber die KundIn ist in

einem standardisierten Interview mit vorgefertigten Antworten zu befragen. Die Fragen müssen einfach, kurz und leicht zu beantworten sein. Das Bedürfnis der KundIn nach Kürze und Einfachheit muss berücksichtigt werden.

7.4.6 Zusammenfassung

Für KundInnen von Süßgebäck lässt sich mittels Ähnlichkeitsbeurteilung eine personenübergreifende Begriffsstruktur finden. 51 KundInnen sortierten 25 Süßgebäck-Artikel in einem *Sortierexperiment* mittels der Methode des Hierarchischen Sortierens. Die Daten wurden mit NMDS und HCA ausgewertet. In den Ergebnissen zeigt sich, dass das wichtigste Klassifizierungskriterium für Süßgebäck der Schokoladenanteil, gefolgt vom Verwendungszweck ist. Schokoladenanteil und verwendungsbezogene Kriterien (hochwertiges Teegebäck, Kinderkekse etc.) verweisen darauf, dass die Begriffsstruktur handlungsbezogenen Kriterien folgt, die auf den Handlungszweck bzw. das Ziel verweisen (geschmackliches Erleben, Verwendungszweck) und nicht auf den Kauf.

Diese Klassifizierungskriterien unterscheiden sich von den Ordnungskriterien, die die KundInnen seit Jahren im Verkaufsraum gewöhnt waren. Der Handel sortierte die Süßgebäck-Artikel nach Marken bzw. Herstellern (Bahlsen, Leibnitz, De Beukelaer, Filia etc.). Die Ergebnisse weisen darauf hin, dass die Begriffsstruktur kein Abbild der sensorisch wahrgenommenen Realität ist, sondern ein Abbild der Handlung unter der ordnenden Maßgabe des Handlungszieles bzw. des Handlungszwecks.

In einem *Feldexperiment* mit 1047 KundInnen wurde untersucht, ob eine Regalanordnung, die der Begriffsstruktur der KundInnen folgt, deren Such- und Findeverhalten verbessert. Als Vergleich diente die Regalanordnung, die die KundInnen seit Jahren gewöhnt waren und bei der die Artikel wie oben dargestellt nach Marken bzw. Herstellern zusammengefasst wurden. Passt man die Artikelanordnung im Regal der Begriffsstruktur der KundInnen an, so verbringen die KundInnen signifikant weniger Zeit vor dem Regal, laufen weniger hin und her, bleiben weniger stehen und berühren weniger Produkte. Die neue Regalanordnung zeigte keine signifikanten Veränderungen des Kaufverhaltens. Die Ergebnisse weisen darauf hin, dass die KundInnen durch eine Regal-

anordnung, die ihrer Begriffsstruktur entspricht, bei ihren Such- und Findeprozessen besser unterstützt werden als durch eine Regalanordnung, die sie bereits seit Jahren kennen. Eine kundInnenfreundliche Anordnung, die die Such- und Findeprozesse unterstützt, ist demnach nicht die gewohnte Anordnung, die die KundInnen schon lange kennen, sondern jene, die ihren Begriffsstrukturen entspricht.

8. Interpretation der Ergebnisse der vier Studien im Hinblick auf die Hypothesen

Welche Erkenntnisse lassen sich ziehen, wenn man die Ergebnisse der vier Studien in der Zusammenschau betrachtet? Gibt es gemeinsame, studien-übergreifende Erkenntnisse hinsichtlich nutzerorientierter Anordnungen? Diese Fragen werden im Folgenden untersucht. Dabei werden die Ergebnisse der vier Studien für jede Hypothese gemeinsam betrachtet und ausgewertet. Im *ersten* Abschnitt werden Ergebnisse zu Hypothese 1 gemeinsam interpretiert, im *zweiten* Abschnitt zu Hypothese 2, im *dritten* Abschnitt zu Hypothese 3, im *vierten* Abschnitt zu Hypothese 4 und im *fünften* Abschnitt erfolgt eine Zusammenfassung.

8.1 Gesamtinterpretation der Ergebnisse zu Hypothese 1

Hypothese 1 lautete: *Für Personen einer Nutzergruppe lässt sich eine personenübergreifende Begriffsstruktur finden. Die Begriffsstrukturen unterschiedlicher Nutzergruppen unterscheiden sich.* Ausgangspunkt war die Gesetzmäßigkeit „Begriffe und Begriffsstrukturen sind abhängig vom Handeln und sie repräsentieren früher gemachte Handlungserfahrungen.“ (vgl. Kapitel 5). Basierend auf dieser Gesetzmäßigkeit wurde die Prognose formuliert, dass die Begriffsstruktur abhängig davon ist, welche Vorerfahrungen die Person mit den Objekten gemacht hat. Die Begriffsstruktur wurde als semantische Ähnlichkeit operationalisiert und mittels einer Ähnlichkeitsbeurteilung erhoben. In jeder der vier vorgestellten Studien wurde diese Prognose im Sinne einer Hypothese untersucht, nämlich, dass sich für Personen einer Nutzergruppe gemeinsame begrifflich-semantische Ähnlichkeitsbeurteilungen für die genutzten Objekte finden lassen und dass sich die Ähnlichkeitsbeurteilungen bei unterschiedlichen Nutzergruppen unterscheiden.

Die Ergebnisse aller vier Studien – *der Süßgebäck-, der Zimmerpflanzen-, der Einkaufszentren- und der Flughafen-Studie* – bestätigen diese Hypothesen: Für die KundInnen, NutzerInnen oder Fachkräfte lassen sich jeweils gemeinsame, personenübergreifende Ähnlichkeitsbeurteilungen (Klassifizierungskriterien) für die untersuchten Objekte fin-

den. Wurden zu einer Objektklasse zwei Nutzergruppen untersucht, wie bei der Zimmerpflanzen- und der Flughafenstudie, dann unterschieden sich auch diese Gruppen (KundInnen und Fachkräfte, Viel- und Wenigflieger) in ihren Ähnlichkeitsbeurteilungen.

In allen vier Studien wurden Sortierexperimente zur Ähnlichkeitsbeurteilung der Objekte eingesetzt. Die Ergebnisse wurden mit der Nonmetrischen Multidimensionalen Skalierung und Hierarchischen Clusteranalyse ausgewertet. In allen vier Studien ergaben sich für die untersuchten Nutzergruppen bedeutsame, überzufällige 2-dimensionale NMDS-Konfigurationen, deren Stresswerte (Alienation) auf dem 1% Niveau signifikant sind, was dafür spricht, dass den Konfigurationen stabile Strukturen zugrunde liegen. Darüber zeigten sich für alle Nutzergruppen eindeutige Gruppierungen in der NMDS-Konfiguration und eindeutige Cluster in der HCA. In der Zimmerpflanzen- und Flughafenstudie zeigten sich jeweils beim Vergleich von zwei unterschiedlichen Nutzergruppen (KundInnen und Fachkräfte, bzw. Viel- und der Wenigflieger) wenn auch nur geringfügige, so doch Unterschiede in den Clustern und Konfigurationen.

Aus diesem Ergebnis wird folgende *allgemeine Schlussfolgerung* gezogen: Die Ergebnisse der vier Studien zeigen somit, dass die Personen einer Nutzergruppe die Objekte nicht nach individuellen, einmaligen Kriterien beurteilten. Hätten sie die Objekte nach individuellen Kriterien beurteilt, dann wären die Stresswerte nicht signifikant gewesen und in den NMDS-Konfigurationen und HCA Clustern hätte es keine Gruppierungen gegeben, sondern gleiche Distanzen. Die Ergebnisse weisen also auch darauf hin, dass nicht episodisches (= individuelles) Wissen getestet wurde.

Das Gegenteil – die Aussage begriffliches Wissen sei „universelles“, „allgemein gültiges“ Wissen – trifft aber auch nicht zu! Die Ergebnisse zeigen nämlich, dass die Personen mit ähnlichen Handlungsvorerfahrungen die Objekte ähnlich beurteilten. Würden begriffliche Wissensstrukturen universal im Sinne von allgemein übereinstimmend sein, dann hätte es keine Unterschiede bei den verschiedenen Nutzergruppen geben dürfen. Es zeigten sich aber Unterschiede in den Begriffsstrukturen bei den unterschiedlichen Nutzergruppen. Begriffsstrukturen scheinen somit weder individuell im Sinne von einzigartig zu sein noch universal im Sinne von allgemein gültig. Keines der

beiden Extreme trifft zu. Die Ergebnisse bestätigen vielmehr die Gesetzmäßigkeit, dass die Begriffsstruktur davon abhängt, was eine Person bisher getan hat. Und die Ergebnisse zeigen, dass Personen, die ähnliche Handlungen, Handlungsziele, -abfolgen etc. mit den Objekten gemacht haben, eine Übereinstimmung in der Begriffsstruktur haben und die Objekte nach anderen Kriterien beurteilen als Personen, die andere Handlungen, Handlungsziele und -abfolgen mit den Objekten gemacht haben.

Dies rückt die Bedeutung von Nutzergruppen in den Blickpunkt. Nutzergruppen wurden definiert als Personen mit ähnlichen Handlungszielen, -folgen, -kontexten (vgl. Kapitel 6). Zwischen Individuum und Masse lassen sich somit Nutzergruppen ausmachen. Dies hat eine große Bedeutung mit praktischen und theoretischen Konsequenzen. Für die Praxis es ist z.B. notwendig, die Personen derjenigen Nutzergruppen zu identifizieren und zu untersuchen, für die man die nutzerorientierte Lösung entwickeln will.

8.2 Gesamtinterpretation der Ergebnisse zu Hypothese 2

Hypothese 2 lautete: *Der begrifflichen Ähnlichkeit liegen handlungsbezogene Kriterien zugrunde.* Die zweite Hypothese bezog sich auf die Kriterien, nach denen die Objekte klassifiziert werden, also auf die inhaltlichen Aspekte der semantischen Nähe der Begriffe. Ausgehend von der Theorie von Klix et al. wurde die Prognose gestellt, dass die Ähnlichkeitsbeurteilungen handlungsbezogenen Kriterien folgen und nicht kategorialen. Die NutzerInnen müssten also alle Objekte, die in einem ähnlichen Handlungs- und Verwendungskontext stehen, als ähnlich beurteilen. Die Ähnlichkeit dürfte also nicht dadurch bestimmt sein, dass die Objekte der gleichen kategorialen Objektklasse (z.B. Pflanzenfamilie, Herstellerfirma, Fluggesellschaft) angehören.

Hintergrund dieser Prognose sind empirische Befunde, die darauf hinweisen, dass handlungsbezogene Begriffsbeziehungen (quasi)stabil abgespeichert werden, während kategoriale Begriffsbeziehungen erst bei Bedarf hergestellt werden (vgl. Klix 1984, 1992). Handlungsbezogene Begriffsbeziehungen sind demnach diejenigen, die direkt und unmittelbar zur Verfügung stehen und die für spontane Erwartungen im Sinne von

spontanen Inferenzen ohne kontrollierte, kognitive Informationsverarbeitungsprozesse herangezogen werden. Die Ergebnisse aller vier Studien bestätigen diese Hypothese:

Einkaufszentrum-Studie. Der Ähnlichkeitsbeurteilung von Einzelhandelsgeschäften liegen eindeutig handlungsbezogene Kriterien und somit thematische Handlungsfelder zugrunde. Die KundInnen beurteilen so z.B. alle Einzelhandelsgeschäfte, die mit „Essen und Kochen“ zu tun haben (z.B. Bäcker, Metzger, Feinkost, Supermarkt etc.) als semantisch ähnlich. Beispiele für andere handlungs- und ereignisbezogene Ähnlichkeitsbeurteilungen sind die Handlungsfelder: „Den Körper pflegen und gesunden“ (Drogerie, Apotheke, Parfümerie), „Schreiben und Lesen“ (Zeitungen, Schreibwaren, Buchhandlung), „Sehen und hören“ (Schallplatten, Optiker, Foto) „Sich und andere beschenken“ (Blumen, Geschenkartikel) oder „Den Haushalt machen“ (Haushaltswaren, Glas und Porzellan) etc. Auch die Nachbarschaft der Handlungsfelder bestimmt sich durch handlungsbezogene Kriterien: neben dem Handlungsfeld „Essen und Kochen“ liegt z.B. das Handlungsfeld „Den Haushalt machen“. Neben dem Handlungsfeld „Sich bekleiden“ liegen die Handlungsfelder „Sich schmücken“ oder „Den Körper pflegen und gesunden“.

Darüber hinaus zeigen sich Kreuzklassifikationen: Die Handlungsfelder „Sich Bekleiden“ (Damen-, Herren-, Kinder-, Sportbekleidung) und „Spielen und Werken“ (Handarbeiten, Heimwerken, Spielwaren) sind über die „Kindergeschäfte“ Spielwaren und Kinderbekleidung semantisch benachbart. So kann man auch von einem Handlungsfeld „Fürs Kind“ sprechen, das die Handlungsfelder „Bekleiden“ und „Spielen und Werken“ überlagert. Die Beispiele zeigen, dass auch der Begriffsstruktur von KundInnen für Einzelhandelsgeschäfte handlungsbezogene Kriterien zugrunde liegen.

Flughafen-Studie. Auch den Ähnlichkeitsurteilen der Flughafengäste liegen handlungsbezogene Kriterien zugrunde. Sowohl die Vielflieger als auch die Wenigflieger unterscheiden drei Gruppen von Serviceeinrichtungen, die sich als Ankunft, Abflug und Zentralbereich beschreiben lassen. Die Serviceeinrichtungen des Zentralbereichs liegen dabei – abhängig von der Bedeutung fürs Fliegen – eher ankunftsnahe (z.B. Blumen),

abflughnah (z.B. Zeitungen, Reisebedarf) oder wirklich flugfern (z.B. Fundbüro, Polizei, Arzt). Unterschiede zwischen den Gruppen gibt es in der differenzierten Zuordnung.

Viel- und Wenigflieger unterscheiden sich z.B. in der Beurteilung der Fluggesellschaften im Ankunfts- und Abflugbereich: Wenigflieger unterscheiden nur zwischen Linie und Charter. Vielflieger unterscheiden eindeutig zwischen innerdeutscher Fluglinie, internationaler Fluglinie und Charter, was daran liegt, dass Vielflieger aus geschäftlichen Gründen häufig innerdeutsch fliegen. Hier, wie bei der unterschiedlichen Beurteilung diverser Geschäfte und Serviceeinrichtungen, zeigt sich, dass die Vielflieger andere Handlungsziele und Abläufe mit dem Flughafen verbinden als Wenigflieger: Vielflieger fliegen in der Regel zu geschäftlichen Terminen, fliegen häufiger innerhalb von Deutschland als Wenigflieger und stehen vor und nach dem Flug unter einem anderen zeitlichen Termin- und Handlungsdruck. Deshalb benötigen sie nach der Ankunft auch Serviceeinrichtungen wie Gepäckaufbewahrung, Stadt- und Hotelinformation, Transfer, Meeting-Point und Parkleitsystem in einem anderen Maße als Wenigflieger.

Zimmerpflanzen-Studie. Die KundInnen und Fachkräfte von Zimmerpflanzen unterschieden zunächst strikt zwischen Grün- und Blühpflanzen. Dies könnte auf den ersten Blick als kategoriales Kriterium interpretiert werden. Doch Grün- und Blühpflanzen werden auch in anderen Handlungskontexten verwendet. Innerhalb der Grün- und Blühpflanzen beurteilen sowohl die Fachkräfte als auch die KundInnen die Pflanzen eindeutig nicht nach kategorialen Kriterien (z.B. Pflanzenfamilie).

Im Gegenteil: aus botanischer Sicht beurteilen beide Gruppen Pflanzen hinsichtlich ihrer Ähnlichkeit als falsch. Für die KundInnen lassen sich bei den Blühpflanzen vier Gruppen finden, die sich beschreiben lassen als „Preiswerte Blühpflanze als Geschenk für sich oder andere“ (Azalee, Cyclame, Usambara), „Hochpreisige, wertvolle Geschenk-Blühpflanzen“ (Orchidee, Stephanotis, Passionsblume), „Stachelige Liebhaber-Blühpflanzen“ (Kaktus, Aechmea), „Südliche Fenster- oder Terrassenpflanzen“ (Bougainvillea, Hibiskus, Kamelie, Solanum).

Den semantisch-begrifflichen Ähnlichkeitsbeurteilungen liegen somit Handlungsziele zugrunde, die sich darauf beziehen, was die KundIn mit dem Kauf der Blühpflanze er-

reichen will: ein Mitbringsel, ein Hochwertiges Geschenk, ein Liebhabergeschenk oder eine Blühpflanze für das sonnige Fenster oder die Terrasse. Dabei kann die Blühpflanze „Geschenk“ für sich oder aber andere sein. Dabei können diese begrifflich ähnlichen Pflanzen in einem gemeinsamen Kontext erfahren worden sein oder aber sie waren austauschbar, d.h. funktional äquivalent hinsichtlich des Handlungsziels

Für die Fachkräfte lassen sich bei den Blühpflanzen auch vier Gruppen finden, die sich jedoch zum Teil oder gänzlich anders definieren: „Preiswerte Angebots-Blühpflanzen“ (Azalee, Cyclame, Usambara, Solanum), „Blühpflanzen auch für Kübelarrangements geeignet“ (Bougainvillea, Stephanotis, Passionsblume), „Hochpreisige Exoten“ (Orchidee, Aechmea), „Mediterrane Pflanzen“ (Hibiskus, Kamelie, Kakteen). Den semantisch-begrifflichen Ähnlichkeitsbeurteilungen der Fachkräfte liegen Handlungsziele zugrunde, die sich darauf beziehen, in welchem gemeinsamen Kontext die Pflanzen erfahren wurden oder was aus floristischer und verkäuferischer Sicht mit den Pflanzen erreicht werden sollte: Angebotsware, Arrangementpflanzen etc.

Süßgebäck-Studie. Die Kriterien, die der Ähnlichkeitsbeurteilung der KundInnen von Süßgebäck zugrunde liegen, lassen sich als Geschmacks- und Verwendungskriterien beschreiben. Die KundInnen unterscheiden zunächst zwischen Süßgebäck mit Schokolade und ohne Schokolade. Innerhalb des schokoladenhaltigen Gebäcks unterscheiden sie zwischen „Hochwertigem Teegebäck /Gebäckmischungen“ und „Einfachen Schokokeksen“ (z.B. Schoko-Butterkekse). Innerhalb des schokoladenfreien Gebäcks unterscheiden sie zwischen „Hochwertigem Gebäck“ und „Einfachen Keksen“ (z.B. Butterkekse). Die Klassifizierungen verweisen auf das Handlungsziel (z.B. Gebäck für nachmittägliches Kaffeetrinken, Pausengebäck etc.) bzw. auf die Konsequenzen des Handlungsziels (z.B. geschmackliches Erlebnis, Konsequenzen des Essens).

Die begrifflich-semantische Klassifizierung von Süßgebäck unterscheidet sich damit eindeutig von der Süßgebäck-Klassifizierung im Regal, die die KundInnen seit Jahren gewohnt waren. Das Süßgebäck wurde traditionell nach Hersteller bzw. Marke (z.B. Bahlsen, Filia, De Beukelaer etc.) im Regal zusammengestellt. Trotz dieser seit Jahren gewohnten Hersteller-orientierten Klassifizierung im Regal beurteilten die KundInnen das Süßgebäck im Sortierexperiment nach Geschmacks- und Verwendungskriterien.

Der Ähnlichkeitsbeurteilung liegen damit nicht Kriterien zugrunde, die die KundInnen aus dem Supermarkt kannten, sondern Kriterien, die mit dem Handeln, genauer mit dem Handlungsziel und den Handlungskonsequenzen zu tun haben. Oder anders formuliert: die semantische Ähnlichkeitsbeurteilung ist kein Abbild der wahrgenommenen Realität, sondern repräsentiert die Handlungsziele der Personen.

Zusammenfassung. Die Ergebnisse aller vier Studien zeigen, dass die begrifflich-semantische Ähnlichkeitsbeurteilung Kriterien folgt, die mit den Handlungen der Nutzer zu tun haben. Die begrifflich-semantische Ähnlichkeitsbeurteilung folgt nicht hierarchisch-kategorialen Klassifizierungen. Differenziert man Handlungen nach Handlungszielen, Handlungsfolgen und Handlungskontext, dann lässt sich sogar über alle vier Studien hinweg feststellen, dass die Kriterien Handlungszielen bzw. den Finalitäten der Handlung folgten.

Betrachtet man Studien, in denen zwei Nutzergruppen zum gleichen Objektbereich untersucht wurden, wie die KundInnen und Fachkräfte in der Zimmerpflanzen-Studie und die Viel- und Wenigflieger in der Flughafenstudie, zeigt sich Folgendes: die Nutzergruppen unterscheiden sich nicht vollständig in der begrifflich-semantischen Beurteilung der Objekte. Sie haben gemeinsame und unterschiedliche Klassifizierungskriterien. In beiden Untersuchungen sind die ersten und damit grundlegenden gemeinsamen Klassifizierungskriterien: KundInnen und Fachkräfte unterscheiden beide nach Grün- und Blühpflanzen; Viel- und Wenigflieger unterscheiden nach Ankunft- Abflug- und Zentralbereich. Unterschiede zeigen sich dann in den feineren Klassifizierungen.

Dies zeigt an, dass die beiden Nutzergruppen sich in grundlegenden Handlungszielen mit dem Objektbereich nicht unterscheiden: sowohl Viel- als auch Wenigflieger kommen am Flughafen an bzw. fliegen dort ab. Sowohl KundInnen als auch Fachkräfte verwenden Grünpflanzen in anderen Kontexten und für andere Handlungsziele als Blühpflanzen. Diese triviale Tatsache spiegelt sich den Klassifizierungskriterien der Ähnlichkeitsbeurteilung wieder. Unterschiede zeigen sich in den feineren Handlungszielen: Vielflieger brauchen bei der Ankunft und beim Abflug andere Serviceeinrichtungen als Wenigflieger, um ihre beruflichen Reiseziele zu realisieren. KundInnen von Zimmerpflanzen beurteilen Blühpflanzen mehr nach ihrer Geschenk- und Dekorations-

qualität, Fachkräfte eher danach, ob es ein Angebotsartikel ist, den man besonders präsentiert und ob man damit Pflanzenarrangement machen kann etc.

Aus diesem Ergebnis lässt sich folgende *zweite allgemeine Schlussfolgerung* ziehen: Die Begriffsstruktur ist kein Abbild der wahrgenommenen Realität. Die Begriffsstruktur scheint sich vielmehr spezifisch durch das Handlungsziel und den Handlungszweck, den die Person beim Handeln mit den Objekten verfolgt, zu bestimmen: Objekte, die im gleichen Handlungskontext bzw. für das gleiche bzw. ähnliche Ziel/Zweck verwendet werden, haben eine hohe semantische Ähnlichkeit.

Die Ergebnisse der Studien zeigen, dass Handlungsziel bzw. Handlungszweck eine ordnende Wirkung für die Begriffsstruktur zu haben scheinen. Somit erfährt die Aussage, die Begriffsstruktur ist abhängig von der Handlungserfahrungen, eine Konkretisierung! Objekte bzw. Funktionen, die im gleichen Kontext stehen bzw. das gleiche Ziel bzw. den gleichen Zweck für den Nutzer erfüllen, haben für ihn auch eine hohe semantische Ähnlichkeit. Die Unterschiede in den Handlungszielen können gering sein, was zu geringfügigen Unterschieden in den Klassifizierungskriterien führt. Die Unterschiede können groß sein, was zu gänzlich anderen Klassifizierungskriterien führen kann. Die Schlussfolgerung kann helfen, die Ergebnisse begrifflicher Ähnlichkeitsuntersuchungen schneller und treffender zu interpretieren. Für die Interpretation von semantisch ähnlichen Objekten ist der gemeinsame Handlungskontext und/oder das gemeinsame (oder ähnliche) Ziel bzw. Zweck zu suchen. So könnte z.B. in einer Nachuntersuchung den Nutzern die NMDS-Konfiguration vorgelegt werden mit der Frage „Wofür verwenden Sie die Objekte, die nahe beieinander sind?“

Die Schlussfolgerung regt aber auch zu folgender Überlegung an: um semantisch begriffliche Ähnlichkeiten vorherzusagen, könnte es ausreichen, die Handlungskontexte und Handlungsziele der Nutzergruppe bezüglich der Objekte zu untersuchen. Die unterschiedlichen Handlungsziele müssten dann in Klassifizierungskriterien übersetzt werden und schließlich in eine Anordnung der Objekte. Ob eine Untersuchung der Handlungsziele der Nutzer ein zuverlässigeres und auch kostengünstigeres Verfahren darstellt als das Abtesten der Begriffsstruktur ist allerdings zweifelhaft. Die Interpretation der Er-

gebnisse regt aber auch zu einer Weiterentwicklung der Begriffstheorie von Klix et al. an. Dies erscheint lohnenswert und wird in Kapitel 9 versucht.

8.3 Gesamtinterpretation der Ergebnisse zu Hypothese 3

Hypothese 3 lautete: *Die begriffliche Ähnlichkeit ist unabhängig vom aktuellen Handlungsziel.* Die dritte Hypothese bezog sich auf die Frage, wie stabil Begriffsstrukturen sind, ob Begriffsstrukturen in Abhängigkeit von einem aktuellen Handlungsziel neu umstrukturiert werden oder ob sie stabil und damit unabhängig vom aktuellen Handlungsziel sind. Konkret bedeutet das, dass unabhängig davon, unter welchem Handlungsziel Personen z.B. Zimmerpflanzen beurteilen etc., die begrifflich semantische Ähnlichkeit für die Pflanzen immer die gleiche sein müsste. Basis für diese Annahme sind Aussagen der Begriffstheorie von Klix (1992), der explizit Begriffsstrukturen als quasi-stabil betrachtet. Er nimmt an, dass sie, wenn sie erst einmal gebildet sind, eine Stabilität besitzen, die erst wieder durch erneute Lern- und Begriffsbildungsprozesse verändert werden kann.

Ausgehend von dieser Annahme wurde die Hypothese formuliert, dass KundInnen trotz unterschiedlicher Instruktionen sich nicht in der Ähnlichkeitsbeurteilung der Objekte unterscheiden. Diese Hypothese wurde nur in einer Studie – der Zimmerpflanzen-Studie – geprüft. Die Ergebnisse zeigen, dass die beiden unterschiedlich instruierten KundInnen-Stichproben zwar unterschiedliche Kriterien während des Sortierens nennen, dass sich aber das Sortierverhalten, sprich die Ähnlichkeitsbeurteilung, nicht bedeutsam unterscheidet. Zweifelhaft blieb, ob die Instruktionen ausreichten, um bei den Versuchspersonen unterschiedliche Handlungsziele entstehen zu lassen. Beide Stichproben wurden im Verkaufsraum – also im Kontext des Einkaufens von Zimmerpflanzen – untersucht.

Deshalb wurde aus den Ergebnissen dieser Studie nur die *eingeschränkte Schlussfolgerung* gezogen: Die Instruktion scheint keinen Einfluss auf die Ähnlichkeitsbeurteilung zu haben, sofern die Personen in der für die untersuchte Objektmenge typischen

Umgebung untersucht werden. Eine weiterreichende Schlussfolgerung würde weitere Untersuchungen erfordern.

8.4 Gesamtinterpretation der Ergebnisse zu Hypothese 4

Hypothese 4 lautete: *Eine Objktanordnung wird dann Such- und Findeleistungen der Nutzer verbessern, wenn die Objekte entsprechend der begrifflich-semanticen Ähnlichkeitsbeurteilung der Nutzer angeordnet sind.* Ausgangspunkt dieser Hypothese war die Annahme, dass NutzerInnen die Erwartung, wo sich ein Objekt befindet, auf der Basis ihrer Begriffsstrukturen bilden. Und es wurde weiter angenommen, dass die Nutzer die gesuchten Objekte schneller und einfacher finden, wenn die relationale Anordnung der Objekte im Raum der Begriffsstruktur der NutzerInnen entspricht.

Diese Hypothese wurde in der Süßgebäck-Studie in einem Feldexperiment getestet. Es wurde zunächst die begrifflich-semantiche Ähnlichkeit für Süßgebäck-Artikel an KundInnen erhoben. Die NMDS eignet sich dabei als Auswertungsverfahren sehr gut, weil sie semantiche Ähnlichkeitsbeurteilungen der Nutzer bereits in einer räumlichen Konfiguration abbildet. Dann wurden die Süßgebäck-Artikel im Regal entsprechend der begrifflich-semanticen Ähnlichkeitsbeurteilung der KundInnen angeordnet. Semantiche ähnliche Süßgebäck-Artikel wurden räumlich entsprechend nahe platziert. In einem Feldexperiment wurde dann das Verhalten der KundInnen an zwei Regalanordnungen beobachtet und analysiert: an der bisherigen Hersteller-orientierten Regalanordnung, die die KundInnen schon seit Jahren gewöhnt waren. Hier waren die Artikel nach Marken und Hersteller angeordnet und an der neuen, Begriffsstruktur-orientierten Regalanordnung. Gemessen wurden Verhaltens- und Abverkaufsdaten.

Die Ergebnisse dieser Studie bestätigen die Hypothese: Die KundInnen verbrachten weniger Zeit vor einem „Begriffsstruktur-orientierten“ Regal. Sie liefen weniger oft auf und ab und legten damit weniger Weg zurück. Sie blieben seltener stehen und fassten weniger oft Produkte an. Trotzdem blieben Umsatz und Absatz gleich. Die KundInnen fanden somit in kürzerer Zeit und mit weniger Aufwand ihre gewünschten Produkte im Vergleich zu der seit Jahren gewöhnten Hersteller-orientierten Anordnung.

Die Ergebnisse zeigen, dass eine Anordnung, die den Begriffsstrukturen der NutzerInnen entspricht, Such- und Findeprozesse besser unterstützt als eine altbekannte, gewohnte Anordnung. Die Ergebnisse weisen darauf hin, dass Gewöhnung, Erfahrung, Lernen keine Erfolgsbedingungen für schnelles Finden sind. Auch wenn NutzerInnen oder KundInnen häufig behaupten, sie müssten eine Anordnung nur lange genug kennen, um sich gut auszukennen, gibt es dennoch eine Anordnung, die per se besser ist für ein einfaches und schnelles Suchen und Finden: Eine Anordnung, die der Begriffsstruktur der Nutzer entspricht, ist eine solche Anordnung, die Such- und Findeprozesse erleichtert.

Aus diesem Ergebnis wird die *vierte allgemeine Schlussfolgerung* gezogen: Eine Anordnung von Objekten im Raum, die der Anordnung der Objekte in der Begriffsstruktur der Nutzer entspricht, unterstützt die Such- und Findeprozesse der NutzerInnen.

8.5 Zusammenfassung

Die Ergebnisse aller vier Studien bestätigen die Hypothesen bzw. die Prognosen, die von den Gesetzmäßigkeiten der Begriffstheorie von Klix, Hoffmann et al. abgeleitet wurden. Die Theorie und die Gesetzmäßigkeiten konnten auf das Problem angewendet werden. Es konnte dadurch ein Untersuchungsdesign und Hypothesen entwickelt werden konnten und diese erwiesen sich als geeignet für die Problemlösung. Die Ergebnisse der Untersuchungen lieferten darüber hinaus auch neue Anregungen für eine Theorieentwicklung. Diese wird im nächsten Kapitel vorgestellt.

9. Auswertung der Ergebnisse in Hinblick auf die verwendete Theorie

Im folgenden Kapitel werden die Ergebnisse der Studien für eine Theorieweiterentwicklung ausgewertet. Im *ersten* Abschnitt wird aufgezeigt, dass es wichtig wäre in der Forschung die Handlungsabhängigkeit des begrifflichen Wissens zu thematisieren. Im *zweiten* Abschnitt wird – basierend auf Erkenntnissen der Studien, dass Handlungsziele eine Ordnung stiftende Funktion für Begriffsstrukturen haben – der Versuch unternommen, neue theoretische Annahmen über Begriffe und Begriffsstrukturen zu formulieren. Diese neuen Annahmen ermöglichen im *dritten* Abschnitt eine differenziertere Prognose darüber, wie NutzerInnen die semantische Ähnlichkeit von Objekten beurteilen werden. Der *vierte* Abschnitt gibt eine Zusammenfassung.

9.1 Ein Untersuchungsthema für die Forschung

Obwohl Klix (1971, 1992) und Hoffmann (in seinen frühen Arbeiten 1983, 1986, wie in seinem Modell der antizipativen Verhaltenssteuerung 1993, 2000) betonen, dass Begriffe und Begriffsstrukturen untrennbar mit dem Verhalten verbunden sind, wird diese Erkenntnis m.E. weder konsequent in den theoretischen Aussagen noch im Untersuchungsdesign umgesetzt. Eine Umsetzung in den Hypothesen und im Untersuchungsdesign würde z.B. erfordern, dass die Handlungsvorerfahrungen der Versuchspersonen hypostasiert, erhoben, variiert bzw. kontrolliert, Versuchsgruppen entsprechend parallelisiert werden müssten. Entsprechendes wurde (in den mir bekannten Experimenten) nicht unternommen. Um diese Behauptung zu belegen, werden z.B. die Untersuchungen zu Primärbegriffen von Rosch (1977) und Hoffmann (1988) angeführt.

Primärbegriffe sind innerhalb von Begriffshierarchien als diejenigen definiert, zu denen die Versuchspersonen die meisten anschaulichen Merkmale assoziieren können. Außerdem sind sie die effektivsten Begriffe in der begrifflichen Identifikation, weil sie z.B. die kürzesten Erkennungszeiten haben. Bereits in Kapitel 4 wurde dargestellt, dass es in jeder Begriffsklasse einen Begriff gibt, der typischer ist als andere: die Gitarre zeigt sich als hochtypisches Musikinstrument, der Apfel als hochtypisches Obst (Hoff-

mann & Zießler 1982). Abbildung 9.1 wird noch einmal dargestellt (vgl. Kapitel 4.2. Begriffstheorien), um weitere Beispiele für Primärbegriffe zu zeigen. Dabei erwähnen die Autoren nicht, für welche Gruppe von Personen diese Begriffe Primärbegriffe sind.

Unterbegriff	Mittelbegriff	Oberbegriff	
Schlaggitarre	Gitarre*	Musikinstrument	
Boskopp	Apfel*	Frucht	
Zimmermannshammer	Hammer*	Werkzeug	
Küchentisch	Tisch*	Möbel	
Sportauto	Auto*	Fahrzeug	
Roteiche	Eiche	Baum*	
Seebarsch	Barsch	Fisch*	
Bergadler	Adler	Vogel*	* Primärbegriff

Abbildung 4.2. Primärbegriffe (Prototypen) auf unterschiedlichen Hierarchieebenen (Rosch 1977, zitiert nach Hoffmann 1983 S. 66)

Es wird behauptet, Hoffmann und Rosch haben hier kultur- und nutzergruppenspezifisch geprägte Primärbegriffe aufgedeckt. Um es deutlich herauszustreichen: es wird nicht das Konzept der Primärbegriffe angezweifelt. Doch es wird behauptet, dass Primärbegriffe abhängig von den Handlungsvorerfahrungen der Nutzergruppe sind. Die hier zitierten Primärbegriffe gelten – wie die meisten Ergebnisse der kognitiven Grundlagenforschung, sofern nicht anders vermerkt, – nur für Personen des europäischen, westlichen Kulturraumes, in dem es große Übereinstimmungen in der Ess- und Musikkultur, der Pflanzen- und Tierwelt gibt. Es wird vermutet, dass der Apfel oder die Gitarre z.B. für Chinesen, Mongolen oder Buschmänner keine Primärbegriffe sein können, weil sie traditionell andere Musikinstrumente und klimatisch bedingt andere Früchte haben. Inhaltliche Unterschiede wird es auch bei Nutzergruppen, die unterschiedliche Handlungs- und Nutzungsziele haben, bezüglich der Primärbegriffe geben. So ist mit großer Wahrscheinlichkeit z.B. der hochtypische Primärbegriff von Handyfunktionen mit großer Wahrscheinlichkeit für Handy-Entwickler ein anderer als für jugendliche Handynutzer, Berufstätige oder ältere Personen.

Es wäre interessant, Experimente zu Begriffen vor dem Hintergrund dieser Annahmen an unterschiedlichen Nutzergruppen durchzuführen. Denn begriffliches Wissen hat

eine große Bedeutung sowohl für bewusste, aber vor allem auch für grundlegende, vorbewusste Prozesse der Wahrnehmung (z.B. begriffliches Identifizieren), der Informationsverarbeitung (z.B. Schlussfolgerungen) und somit auch für das Handeln. In der Süßgebäckstudie zeigte sich, dass begriffliches Wissen selbst bei so einfachen und vorbewussten Erwartungsbildungen und Orientierungsprozessen mitwirkt wie der Suche nach Süßgebäck im Supermarkt-Regal. Es ist anzunehmen, dass sehr viele Verhaltensunterschiede durch unterschiedliches begriffliches Wissen und somit durch unterschiedliche Handlungsvorerfahrungen erklärt werden können.

Die Unterschiede zwischen Wissensinhalten und Gesetzmäßigkeiten der Begriffsbildung müssten eindeutiger thematisiert werden. Erstens, um die Gesetzmäßigkeiten der Begriffs- und Begriffsstrukturbildung besser zu erforschen und zu verdeutlichen. Zweitens, um die begrifflichen Wissensinhalte in Abhängigkeit von der Handlungsvorerfahrung und von unterschiedlichen Nutzergruppen besser kennen zu lernen und zu verstehen. Und drittens um die Konsequenzen der Wissensunterschiede für das spontane, aber auch geplante Verhalten besser zu verstehen. Das würde nicht nur einen Erkenntniszuwachs für die Wissenschaft bedeuten. Es würde auch die Arbeit von Praktikern, wie Nutzerforschern, Technikern, aber auch Entscheidern, wie Führungskräften, inspirieren, unterstützen und effektiver machen und vielleicht für mehr Toleranz, Verständnis und Respekt für das Verhalten anderer Ziel- und Nutzergruppen sorgen, wenn klar wird, dass Verhaltensunterschiede auch spontan und vorbewusst entstehen, nicht, weil die Menschen etwas nicht verstehen, sondern weil sie auch unterschiedliche Vorerfahrungen haben und Erwartungen, Entscheidungen und Verhaltensakte eben auch auf der Basis anderer Informationsdaten entstehen.

9.2 Weiterentwicklung der Begriffstheorie von Klix et al.

Nicht nur Klix und Hoffmann, sondern viele Forscher vor ihnen haben bereits behauptet (z.B. Hull 1920, Gibson 1983, Piaget 1971, Rosch 1975, Aebli 1988), dass Begriffe nicht um ihrer selbst Willen gebildet werden und auch nicht um ein möglichst objektives Abbild der Umwelt zu schaffen, sondern vielmehr im Zusammenhang mit erfolgrei-

chen Verhaltenszielen. Dennoch wurde dies m. E. bisher nicht konsequent in Begriffstheorien umgesetzt. Kennzeichnend dafür ist, dass es bislang nicht gelungen ist eine Begriffstheorie zu konstruieren, die voraussagen kann, welche Merkmale für die Begriffsbildung klassifizierungsrelevant sind bzw. welche Begriffe konkret Begriffsstrukturen bilden.

Klix (1971, 1992) liefert zwar differenzierte Annahmen über begriffsbestimmende Merkmalsklassen und ihre Funktionen, aber er wird dahingehend kritisiert, dass er nicht vorhersagen kann, welche Merkmale klassifizierungsrelevant und damit begriffsbestimmend sind. Hoffmann integriert in seiner Theorie der Antizipativen Verhaltenssteuerung die Begriffsbildung in lern- und handlungstheoretische Bezüge und liefert die Annahme, dass funktionale Äquivalenzklassen, d.h. das Wissen, unter welchen Bedingungen welche Handlungen zu welchen Konsequenzen führen, begriffsbestimmend sind, wobei er dies aber nicht näher definiert (vgl. 1993, S. 151). Klix setzt z.B. voraus, dass Klassifizierungsprozesse an rein sensorischen Merkmalen ansetzen (1971, 1992). Hoffmann kritisiert ihn in seiner Theorie der Antizipativen Verhaltenssteuerung und behauptet, Klassifizierungsprozesse setzen an funktionalen Äquivalenzen unter der Maßgabe des Handlungsziels und der Handlungskonsequenzen (1993) an. Die Annahmen sind zu vage, um differenzierte Vorhersagen über die Inhalte von konkreten Begriffen einer Person zu treffen.

Im Folgenden wird der Versuch unternommen, beide Aussagegebäude und die Erkenntnisse dieser Arbeit zu verbinden, dass nämlich *Handlungsziele eine Ordnung stiftende Funktion für Begriffsstrukturen* haben. Es wird eine neue, erweiterte Begriffstheorie formuliert, die differenzierte Aussagen über Begriffe und Begriffsstrukturen liefert und Vorhersagen erlaubt, welche Merkmale klassifizierungsrelevant, d.h. begriffsbestimmend sind und welche Begriffe Begriffsstrukturen ausbilden.

Erweiterte Basisannahmen Begriffe werden bei Merkmalsrepräsentationstheorien – wie die von Klix et al. – als Ergebnis einer Abstraktion und Generalisation invarianter Merkmale definiert. Für die Überarbeitung der Theorie von Klix et al. werden zusätzliche Basisannahmen in das Aussagegebäude integriert: Begriffe und Begriffsstrukturen entstehen während des Handelns in Hinblick auf ein Handlungsziel. Sie repräsentieren

erfolgreiches Handlungswissen. Es werden somit lern- und handlungstheoretische Annahmen hinzugenommen. Somit gibt es insgesamt folgende drei wichtige Basisannahmen:

- *Begriffsbildung wird als Abstraktion und Generalisation invarianter Merkmale* (vgl. Hull 1920, Klix 1971) *definiert*. Ein Begriff definiert sich durch seine Merkmale und (semantischen) Beziehungen zu anderen Begriffen. Es wird angenommen, dass es unterschiedliche Merkmalsklassen sind, die einen Begriff konstituieren, und dass es unterschiedliche Relationsarten gibt, die Begriffe miteinander verbinden.
- *Begriffsbildung*, d.h. Merkmalsabstraktion und Generalisation *entsteht während des Handelns zur Erreichung eines Handlungszieles* – wobei Sprechen und Lesen auch als Handlung angesehen wird. Wesentlicher Richtungsgeber einer Handlung ist das Handlungsziel, wesentliche Attribute einer Handlung sind der Handlungsgegenstand (Objekt), der Handlungskontext (Personen, andere Gegenstände, Ort, Zeit) und die Handlungsabfolgen (hierarchisch und sequentielle Operationen, Bewegungen, Aktionen), die den Ist-Zustand in einen Soll-Zustand (das Handlungsziel) überführen sollen (vgl. Leontjev 1977, Lantermann 1980, Hacker 1986).
- *Begriffsbildung entsteht durch Lernen*. Wesentliche Attribute des Lernens sind Reiz, Reaktion, wahrgenommene Konsequenzen und eine Reiz/Reaktionsverbindung durch positive, verstärkende (wahrgenommene) Konsequenzen.

Im Folgenden werden hierauf basierend Annahmen zu folgenden Fragen formuliert: Was sind Merkmale? Welche Merkmale sind klassifizierungsrelevant und somit begriffsbestimmend? Welche unterschiedlichen Merkmalsklassen gibt es? Welche unterschiedlichen Relationsklassen und Begriffsstrukturen gibt es?

Neue, erweiterte Annahmen über Merkmale In Anlehnung an Klix (1971, 1992) wird angenommen, dass Abstraktionsprozesse an sensorischen Reizwirkungen (= Merkmalen) ansetzen. Dieser Satz wird erweitert durch die Annahme: Merkmalsabstraktionsprozesse setzen an sensorischen Reizwirkungen (=Merkmalen) an, die die Person mit dem Gegenstand während des Erreichens des Handlungsziels wahrnimmt. Die Reizwirkungen (= Merkmale) können sowohl über umwelt- als auch körperbezogenen Sinnes-

systeme wahrgenommen werden (d.h. über Augen, Ohren, Nase, Sinnesorgane der Haut, des Muskel- und Sehnenapparates, das Gleichgewichtsorgan und andere Sinnesorgane) und sie beziehen sich auf die Farbe, Form, Größe, Schwere, Geruch, Oberfläche, Temperatur, Lage, das Bewegungsgefühl, die Emotion oder andere wahrnehmbare Qualitäten, die der Begriffsgegenstand während der Handlungsfolge in dem spezifischen Handlungskontext auslöst. Es wird darüber hinaus behauptet, dass auch mentale, abstrakte Begriffsgegenstände, (z.B. die Begriffe: Satzgegenstand, Freiheit, Konzept, Märchen, Traum) durch sensorisch basierte Reizwirkungen repräsentiert sind, nämlich durch in der Vorstellung geschaffene „sensorisch-imaginierte“ Merkmale, wie durch eine Visualisierung, ein imaginiertes Körpergefühl, eine imaginierte Emotion oder eine imaginierte Tonfolge.

Neue Annahmen zur Klassifizierungsrelevanz von Merkmalen Es wird angenommen, dass die Person jene Merkmale für die Bildung eines Begriffes abstrahiert, generalisiert und als begriffliches Wissen repräsentiert, die *invariant handlungszielrelevant* sind. Was heißt das? Die Person erfährt während des Handelns – während sie mit dem Objekt versucht, ihr Handlungsziel mittels bestimmter Handlungsfolgen zu erreichen – bestimmte Eigenschaften (= Reizwirkungen) des Objektes in dem Handlungskontext. Sie nimmt das Objekt in bestimmten Bezügen zu anderen Gegenständen, Personen an einem bestimmten Ort zu einer bestimmten Zeit etc. wahr. Und sie nimmt bestimmte Veränderungen des Objektes, des Handlungskontextes etc. wahr.

Wird das Handlungsziel erreicht, wirkt dies verstärkend im Sinne eines Reiz-Reaktions-Lernens: die Person lernt mittels Merkmalsabstraktionsprozessen unter den zielführenden Kontextbedingungen und Handlungsabfolgen die invarianten, handlungsrelevanten Merkmale, die zur Zielerreichung führen. Ergebnis dieses Abstraktions- und Lernprozesses sind die invarianten, handlungszielrelevanten Merkmale über die Objekteigenschaften während der Ausgangssituation (Ist-Zustand), der Handlungsfolgen und der Zielsituation (Soll-Zustand) sowie die Merkmalsbezüge zu invarianten, handlungszielrelevanten Personen und Gegenständen des Handlungskontextes. Es sind also nicht alle Reizeigenschaften begriffsrelevant, die die Person auf dem Weg zur Zielerreichung erfährt, sondern es werden nur jene abstrahiert und generalisiert, die sowohl invariant als auch handlungsrelevant für die erfolgreiche Zielerreichung sind.

Bezüglich der Klassifizierungsrelevanz von Merkmalen wird daher folgende Annahme formuliert: Klassifizierungsrelevant (d.h. begriffsbestimmend) sind jene Merkmale, die die Person bei der Erreichung ihres spezifischen Handlungsziels mit dem Objekt – in einer spezifischen Handlungsabfolge, in einem spezifischen Handlungskontext, bei der Überführung vom Ist- in den Soll-Zustand – erfolgsbestimmend und invariant erlebt (vgl. Abbildung 9.2). Zentrale, Ordnung stiftende Funktion hat damit für die Begriffsbildung (d.h. für den Merkmalsatz von Begriffen) das Handlungsziel.

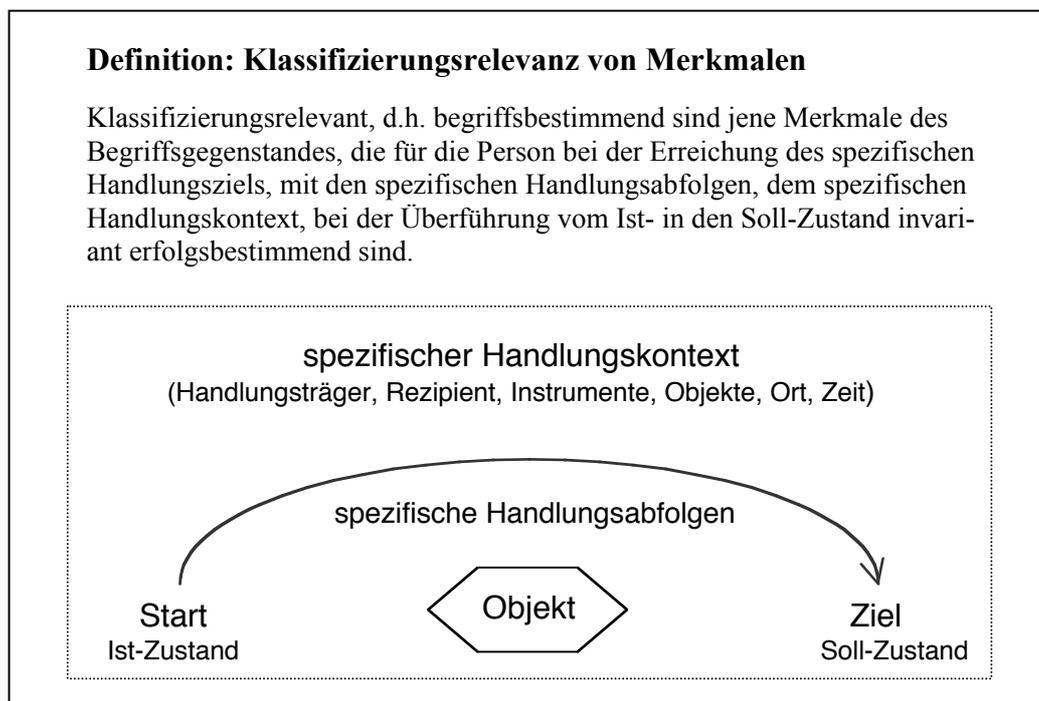


Abbildung 9.2. Visualisierung der Definition: Klassifizierungsrelevanz von Merkmalen für Begriffsbildung

Erweiterte Annahmen zu Merkmals- und Relationsklassen Die Aussage, dass sich ein Begriff durch abstrahierte, invariant handlungszielrelevante Merkmale (Reizwirkungen) konstituiert, ist noch zu allgemein, um Vorhersagen über Begriffe und Begriffsstrukturen zu treffen. Es bedarf differenzierterer Annahmen über Merkmalsklassen, also Arten von Reizwirkungen. In Anlehnung an Klix und die frühen Arbeiten von Hoffmann wird davon ausgegangen, dass der innere Aufbau eines Begriffes bereits seine Beziehungen zu anderen Begriffen und damit die Begriffsstrukturen bestimmt. Es werden sechs Merkmalsklassen unterschieden, die noch eine differenziertere Vorhersage erlauben, welche Merkmale klassifizierungsrelevant für einen Begriff sind. Diese Merkmalsklas-

sen korrespondieren bis auf die erste prinzipiell mit den Merkmalsklassen von Klix et al. Sie unterscheiden sich allerdings inhaltlich und lauten: objektbezogene, partonomische, taxonomische, operative, emotionale und graphisch-phonemische Merkmale. Im Folgenden werden sie erläutert:

- *objektbezogene Merkmale* repräsentieren invariant handlungszielrelevante Reizwirkungen des Begriffsgegenstandes selbst im Handlungsablauf. Diese Merkmalsklasse existiert in der Theorie von Klix et al. nicht. Sie sprechen von einer sensorischen Merkmalsklasse. (In den hier dargestellten Annahmen, wird jedoch davon ausgegangen, dass alle sechs Merkmalsklassen sensorisch definiert sind.) Objektbezogene Merkmale repräsentieren die invarianten, handlungszielrelevanten Reizwirkungen/Eigenschaften des Begriffsgegenstandes selbst, die für die Erreichung des Handlungsziels relevant sind, d.h. während der Begriffsgegenstand durch Handlungsfolgen vom Ist-Zustand in den Ziel-Zustand übergeführt wird.
- *partonomische Merkmale* repräsentieren Beziehungen zu invariant handlungszielrelevanten Gegenständen/Personen, die zur Erreichung des Handlungszieles im gleichen Handlungskontext erfahren werden, wie handlungszielrelevante Personen (z.B. der Handlungsträger, der Handlungsempfänger), handlungszielrelevante Gegenstände (z.B. Werkzeuge, Objekte, Begleitgegenstände), der handlungsrelevante Ort und die Zeit. (Diese Merkmalsklasse wird in Anlehnung an Hoffmann (1993) partonomisch (pars; lat.) genannt, um auszudrücken, dass der Begriffsgegenstand ein Teil einer Gesamtsituation ist.)
- *taxonomische Merkmale* repräsentieren die Beziehungen zu Begriffen, mit denen spezifischere, ähnliche oder allgemeinere Handlungsziele und -konsequenzen erreicht werden, d.h. zu Ober-, Unter-, Neben- und Synonymbegriffen. Solche Ober-, Unter- Neben- und Synonymbegriffe und die dazugehörigen begriffsbestimmenden Merkmale bildet die Person während des Handelns durch eine Differenzierung ihrer Verhaltensziele und den erfahrenden Verhaltenskonsequenzen mit den Gegenständen heraus. (Diese Merkmalsklasse wird taxonomisch (taxon; griech.) genannt, um auszudrücken, dass der Begriffsgegenstand in eine Systematik eingeordnet ist.

Dabei handelt es sich nicht um eine Fachsystematik sondern eine Systematik, die durch unterschiedlich spezifische Handlungsziele bestimmt ist.)

- *operative Merkmale* repräsentieren die Bezüge zu entsprechenden Handlungsfolgen (Verhaltensprogrammen), die den Gegenstand von einem Ist- in einen Soll-Zustand überführen. Solche Verhaltensprogramme können elementare, neuronal organisierte Bausteine für sensomotorische Verhaltensketten (z.B. das Greifen der Finger) sein, oder aber begriffliche Einheiten für komplexere Verhaltensprogramme (z.B. den Keks zum Mund führen oder Krümel von den Kleidern entfernen), die wiederum auf elementare, neuronal organisierte Bausteine für Verhaltensketten verweisen. Operative Merkmale scheinen notwendig, um z.B. die Unmittelbarkeit, mit der viele Handlungsintentionen spontan in tatsächliches Handeln umgesetzt werden, erklären zu können.
- *emotionale Merkmale* repräsentieren Bezüge zu entsprechenden emotionalen Bewertungen, die mit dem Begriffsgegenstand verbunden sind. Emotionale Merkmale scheinen notwendig, um z.B. die Unmittelbarkeit einer gefühlsmäßigen Einstellung bzw. Reaktion zu erklären, die ein imaginiertes Begriff bzw. ein realer Begriffsgegenstand auslösen kann.
- *graphisch-phonemische Merkmale* repräsentieren symbolische Stellvertreter für den Begriffsgegenstand (z.B. Wortmarken, Symbole). Wortmarken und Symbole haben zum einen Stellvertreterfunktion. Zum anderen bieten sie auch eine Möglichkeit, um begriffliches Wissen für die Erreichung sprachlicher und gedanklicher Ziele zu erwerben. Es ist somit ein anderes begriffliches Wissen als das konkrete, durch Handeln erlangte Wissen mit dem Gegenstand, weil es auf anderen Handlungszielen, -abfolgen, -kontexten und somit anderen Handlungserfahrungen beruht. So kann sich eine Person begriffliches Wissen z.B. über Orchideen durch Lesen aneignen, um für eine Prüfung zu lernen, durch die konkrete Pflege dieser Pflanzen und ein Leben mit ihnen zu Hause mit oder aber durch ein Erleben der Pflanzen im tropischen Urwald. In allen Fällen wird sie begriffliches Wissen über Orchideen haben, aber die partonomischen, taxonomischen, emotionalen und operativen Merkmale unterscheiden sich, da das Wissen in einer anderen Handlungssituation mit anderen Handlungs-

zielen, anderen Handlungsfolgen und in anderen Handlungskontexten mit anderen Handlungszielen erworben wurde. Alle drei Wissensinhalte können sich ergänzen und aufeinander aufbauen, sind aber nicht gleichzusetzen.

Begriffe bilden über die Merkmale *Begriffsstrukturen*. Im Folgenden wird nur auf partonomische und taxonomische Begriffsstrukturen eingegangen, die in Anlehnung an Klix und Hoffmann definiert werden. (Auf operative und emotionale Begriffsstrukturen wird hier nicht näher eingegangen. Es bedürfte einer breiteren Auseinandersetzung, die über den Rahmen dieser Arbeit hinausgehen würde.) Wesentlicher Unterschied zu den Annahmen von Klix, Hoffmann et al. bezüglich dieser Begriffsstrukturen ist die ordnungstiftende Funktion der Handlungsziele. Es wird angenommen, dass sowohl partonomische als auch hierarchisch taxonomische Begriffsstrukturen während des Handelns zur Erreichung von Handlungszielen gebildet werden.

- ❖ *Partonomisch, ereignisbezogene Handlungen* repräsentieren die Handlungsbezüge zwischen den Begriffsgegenständen der Handlungssituation, die auf dem Weg zum Handlungsziel verwendet werden. Sie repräsentieren nicht die objektiven, räumlichen (z.B. sichtbaren) Beziehungen von Objekten, Personen, Raum und Zeit, sondern die Bezüge hinsichtlich des Handlungsziels. (Hierin unterscheiden sich diese Annahmen von Schematheorien über Skripts und Frames. s. Kapitel 4). Es wird angenommen, dass der zentrale Kern dieser ereignisbezogenen Begriffsstruktur das Handlungsziel ist und nicht wie z.B. von Klix der Tätigkeitsbegriff (z.B. das Verb).
- ❖ *Taxonomisch, hierarchische Begriffsstrukturen* repräsentieren die Beziehungen von Begriffen mit ähnlichen, spezifischeren oder allgemeineren Handlungszielen und -konsequenzen. Sie verweisen somit auf Ober-, Unter-, Neben-, Synonym-, Komparativbegriffen. Diese Begriffe bilden sich durch einen Wechsel der Handlungsziele (anstatt Alkohol Wein, anstatt Weißwein Rotwein, anstatt Rioja Chianti) mit ähnlichen, spezifischeren oder allgemeineren Konsequenzen. Wenn es beispielsweise nur darum geht, den Hunger zu stillen, ist jede Art von Süßgebäck recht (vgl. Hoffmann 1993 S.149). Geht es aber darum, das passende Süßgebäck zum Kaffeetrinken oder für den Kinderausflug zu wählen, sind genauere Unterscheidungen zu treffen und die Person bildet invariante handlungszielrelevante, taxonomische Merkmale her-

aus, wie Geschmacks-, Konsistenzmerkmale o.Ä., die z.B. Blätterteiggebäck von Butterkeksen unterscheiden lassen. Je differenzierter die Gebäckkennerin die Wahl des Süßgebäcks für bestimmte Anlässe treffen will (für die Tasse Tee mit der Schwiegermutter, für die Konferenzpause, für den Kinderkaffeetisch oder zum Espresso nach einem mehrgängigen Menü), desto differenzierter lernt sie zwischen den Süßgebäck-Sorten anhand von Merkmalen (Geschmack, Bestandteile, Schokoladenanteil, Verarbeitungsqualität, Marken etc.) zu unterscheiden und desto mächtiger ist ihr taxonomisch hierarchisches Begriffsrepertoire; d.h. desto mehr Neben-, Ober-, Unter-, Synonym-, Kontrast- und Komparativ-Begriffe hat sie. Durch differenziertere Verhaltensziele und -konsequenzen entsteht ein differenzierteres Wissen. Ursache des begrifflich taxonomischen Wissens sind somit die Handlungsziele und -konsequenzen, Ergebnis ist der Umfang des begrifflichen Wissens.

Die taxonomischen Begriffsstrukturen unterscheiden sich deutlich in Abhängigkeit vom Handlungsziel und sind damit von der Nutzergruppe abhängig. Klassifizierungssysteme, wie das Linée'sche System des Pflanzenreichs, das hauptsächlich auf der Zahl und Anordnung der Staub- und Fruchtblätter beruht (vgl. Brockhaus 2000), werden von Botanikern gemacht und dient diesen. Systematiken von Wein, die auf Rebsorten, Lagen, Anbaugebieten, Keltereien und Jahrgang basieren, dienen Weinproduzenten und -händlern und ihren Zielen. Systematiken von Lebensmitteln, die z.B. nach Inhaltsstoffen (Mehl-, Zucker-, Fleischprodukte, etc.), Herstellungsart (Frischware, Konserven, Tiefkühl- und Trockenware) oder nach Herstellern unterscheiden, verweisen auf Lebensmittelhersteller und -händler. Eine KundIn, die Süßgebäck kaufen will, kauft in der Regel primär nicht einen Inhaltsstoff, eine Marke oder Verpackungsform, sondern Süßgebäck für ein spezifisches Handlungsziel (Zwischenmahlzeit, für Kinder, für das Kaffeekränzchen etc.). Taxonomische, hierarchische Begriffsstrukturen von Nutzern sind kein Abbild enzyklopädischer oder anderer fachlicher Systematiken – wobei sie diese sehr wohl auch erlernen können. Es wird angenommen, dass taxonomische Begriffsstrukturen eine Repräsentation von Begriffen sind, für unterschiedliche und unterschiedlich spezifische Handlungsziele.

Integration partonomischer und taxonomischer Begriffsstrukturen Die obigen Annahmen definieren die Beziehungen sowohl von partonomischen als auch von taxonomischen Begriffsstrukturen über das Handlungsziel: Partonomische Begriffsstrukturen repräsentieren, in welchen Handlungsbeziehungen die Gegenstände des Handlungskontextes auf dem Weg zur Zielerreichung stehen. Taxonomisch-hierarchische Begriffsstrukturen repräsentieren Bezüge zu Begriffsgegenständen, die gleiche, spezifischere oder allgemeinere Handlungsziele erreichen lassen. Beide Begriffsstrukturen sind durch das Handeln in Bezug auf ein Handlungsziel bestimmt. Beide Begriffsstrukturen sind stark ineinander verwoben.

Folgende Beispiele sollen dies deutlich machen: Einzelhandelsgeschäfte wie das Obst- und Gemüsegeschäft, der Metzger und der Supermarkt stehen sowohl in einer partonomischen als auch in einer taxonomischen Beziehung. Partonomisch, weil die KundIn sie im Handlungskontext „Essen einkaufen“ für das Handlungsziel „Essen & Kochen“ erlebt. Sie geht sowohl zum Bäcker, in die Metzgerei und in den Supermarkt. Taxonomisch deshalb, weil es sich um Nebenbegriffe handelt, die ähnlichen oder gar identischen Handlungszielen dienen und dem Oberziel „Essen & Kochen“ untergeordnet sind. Die KundIn kann in eine Metzgerei oder in einen Supermarkt gehen, um Fleisch zu kaufen. Die Geschäfte sind hinsichtlich des Handlungsziel in etwa substituierbar bzw. funktional äquivalent.

Übereinstimmung mit anderen Annahmen Die hier gemachten Annahmen basieren nicht auf den Axiomen der Booleschen Algebra (trennscharfe Klassen, Zugehörigkeit nach dem Alles-oder-Nichts-Prinzip etc.). Sie ermöglichen daher eine Erklärung und Vorhersage der Phänomene wie Kreuzklassifikation bzw. unscharfe Begriffe. Kreuzklassifikation kann so erklärt werden, dass ein und derselbe Gegenstand für unterschiedliche Handlungsziele unterschiedliche, handlungszielrelevante Merkmale liefert: Ein Baumstamm kann handlungszielrelevante und damit begriffsbestimmende Merkmale eines Schemels aufweisen, wenn man ihn aufrecht stellen kann, er die richtige Höhe (Länge) und den richtigen Durchmesser aufweist und man sich somit darauf setzen kann. Klassifizierungsrelevante Merkmale sind damit Größe und Höhe einer Sitzfläche sowie eine vertikale Stabilität (i.S. von wankt nicht). Der Baumstamm kann – sofern er schnitzrelevante Merkmale wie keine Asteinschlüsse, ebenmäßiger Wuchs,

keine Risse etc. aufweist – die Voraussetzungen für einen Schnitzrohling haben, aus dem ein Künstler eine Holzplastik anfertigen kann. Oder aber er kann Brennholz sein, das noch zerteilt und gesplissen werden muss. In allen drei Fällen sind aufgrund unterschiedlicher Handlungsziele, -folgen, -kontexte etc. unterschiedliche sensorisch wahrnehmbare Merkmale handlungszielrelevant und somit begriffsbestimmend.

Experimentelle Überprüfung Die hier gemachten Annahmen über Begriffe und Begriffsstrukturen könnten Ausgangspunkt für eine experimentelle Überprüfung sein. Es sind Experimente vorstellbar, die Hypothesen zur Klassifizierungsrelevanz und zu partonomischen und taxonomischen Strukturen testen.

9.3 Differenzierung der Prognose zur semantischen Ähnlichkeit von Objekten

Anhand der oben dargestellten, neuen Annahmen über Begriffe und Begriffsstrukturen lassen sich auch neue, differenziertere Hypothesen bzw. Prognosen darüber bilden, wie NutzerInnen die semantische Ähnlichkeit von Begriffsgegenständen beurteilen. Diese Hypothesen bzw. Prognosen lauten:

Die semantische Ähnlichkeit von Objekten bestimmt sich durch ...

- a) die Verwendung in einem gemeinsamen Handlungskontext hinsichtlich des Handlungsziels: d.h. Objekte, die in einem gemeinsamen Handlungskontext hinsichtlich eines Handlungsziels verwendet werden, werden semantisch sehr ähnlich beurteilt. Objekte, die nicht in einem Handlungskontext verwendet werden, werden entsprechend semantisch unähnlich beurteilt.
- b) die Gleichwertigkeit (Äquivalenz) in der Verwendung für ein Handlungsziel: Objekte, die für ähnliche Handlungsziele verwendet werden, d.h. Objekte, die *funktional äquivalent sind*, werden semantisch sehr ähnlich beurteilt. Objekte, die nicht funktional äquivalent sind, d.h. die für sehr unterschiedliche Handlungsziele verwendet werden, werden semantisch entsprechend unähnlich beurteilt.

Diese Hypothesen differenzieren also zwischen partonomischen Kontextbegriffen und taxonomischen Nebenbegriffen. Diese Hypothese müsste in der Lage sein die Ähnlich-

keitsbeurteilungen differenzierter zu erklären und zu prognostizieren. Um diese Hypothese zu prüfen, müssten neue Untersuchungen gemacht werden.

9.4 Zusammenfassung

In diesem Kapitel wurde zunächst dargestellt, dass es wichtig wäre in der Forschung die Handlungsabhängigkeit begrifflichen Wissens stärker zu thematisieren, um den Unterschied zwischen den Gesetzmäßigkeiten der Begriffsbildung, den Inhalten begrifflichen Wissens und den Konsequenzen begrifflicher Wissensunterschiede deutlicher herauszuarbeiten. Dann wurden die Erkenntnisse aus den vier Studien verwendet, um die Theorie von Klix (1971, 1992) und Hoffmann (1983, 1986 und 1993) weiterzuentwickeln. Dies erschien nötig und möglich, da bislang der Handlungs- und Lernbezug nicht konsequent in differenzierte Annahmen über Begriffe und Begriffsstrukturen umgesetzt wurde.

Als Erstes wurden neue Basisannahmen eingeführt, die Begriffsbildung lern- und handlungstheoretisch fundieren: Begriffsbildung findet durch Lernen während des Handelns zur Erreichung des Handlungsziels statt. Darauf aufbauend wurde eine Annahme zur Klassifizierungsrelevanz von Merkmalen gebildet: Begriffsbestimmend sind invariante, handlungszielrelevante Merkmale, die die Person mit dem Gegenstand bei der Erreichung des spezifischen Handlungsziels, mit der spezifischen Handlungsabfolge, dem spezifischen Handlungskontext, bei der Überführung vom Ist- in den Soll-Zustand erlebt. Es wurden sechs Merkmalsklassen definiert. Und es wurden Annahmen über partonomische und taxonomische Begriffsstrukturen formuliert, bei denen nicht mehr der Tätigkeitsbegriff, sondern das Handlungsziel eine Ordnung stiftende Funktion hat. Partonomische Begriffsstrukturen repräsentieren Begriffe von Handlungsabfolge und -kontext zur Erreichung des Handlungsziels. Taxonomische Begriffsstrukturen repräsentieren Begriffe, die sich durch den Wechsel – hinsichtlich allgemeinerer, spezifischerer oder ähnlicher – Handlungsziele definieren. Die erweiterte Theorie erlaubt differenziertere Vorhersagen darüber, welche Merkmale begriffskonstituierend sind als die Theorie von Klix (1971, 1992) und Hoffmann (1993). Und sie erlaubt differenziertere Annahmen über die semantische Ähnlichkeit von Begriffen. Zur Überprüfung müssten Experimente folgen.

10. Leitfaden für die Entwicklung nutzerspezifischer und erwartungsgerechter Anordnungen von Objekten und Funktionen im Raum

Im *ersten* Abschnitt wird aus den Erkenntnissen dieser Arbeit ein allgemeiner Leitfaden formuliert, der angewandt arbeitenden PsychologInnen oder ErgonomInnen dazu dienen soll, Anordnungen zu entwickeln, die den Erwartungen der NutzerInnen entsprechen. Das Vorgehen ist für Input-Studien im Entwicklungsprozess geeignet, um den Gestaltern theoretisch fundierte *und* empirisch gesicherte Vorgaben für eine nutzerorientierte Anordnung der Objekte oder Funktionen zu liefern. Dieses empirische Vorgehen erfordert Aufwand, der nicht immer zu finanzieren ist. Deshalb wird zusätzlich im *zweiten* Abschnitt ein nicht empirisches, spekulatives Vorgehen skizziert. Dass so ein spekulatives Vorgehen hohe Risiken birgt und schnell falsche Lösungen liefern kann, ist selbstverständlich. Doch zunächst zu dem empirischen Leitfaden.

10.1 Leitfaden zur Entwicklung empirisch fundierter Vorgaben für nutzerspezifische, erwartungsgerechte Anordnungen

Ziel des Leitfadens Ziel des Leitfadens ist es, ein Vorgehen zu beschreiben, wie Objekte zu gruppieren und anzuordnen sind, damit die NutzerInnen schnell das finden, was sie suchen. Das Vorgehen ist theoretisch und empirisch fundiert; d.h. der Problemlösungsansatz basiert auf einer Theorie und die Kriterien, nach denen die Objekte angeordnet werden, werden empirisch an der entsprechenden Nutzerstichprobe erhoben.

Einsatzbereich Der Leitfaden setzt theoretische und methodische Kenntnisse voraus. Er sollte daher von Psychologen, Ergonomen oder Personen mit entsprechenden Kenntnissen durchgeführt werden. Er ist nicht selbsterklärend für Gestalter wie Architekten, Designer oder Konstrukteure. Der Leitfaden eignet sich für die Entwicklung nutzerorientierter Anordnungen von Objekten/Funktionen im Handel (z.B. bei Artikel- und Sortimentsanordnungen im Verkaufsraum), in der Architektur (beim Raum- und Funktions-

programm für die Grundrissplanung), im Produktdesign (bei der Anordnung von Funktionen auf Bedienoberflächen), Arbeitsplatzgestaltung (z.B. bei der Anordnung von Funktionen für Schaltpulte, Steuerungsanlagen).

Im Rahmen dieser Arbeit wurde der Leitfaden in der Architektur und im Handel durchgeführt. Auch bei Bedienoberflächen oder technischen Geräten ist eine nutzerorientierte Anordnung von Funktionen meiner Erfahrung nach ein wichtiges Usability-Kriterium, damit der Nutzer oder die Nutzerin die gesuchten Funktionen spontan und intuitiv findet. Für den Bereich Fahrzeug-Cockpits liegen Erfahrungen in unveröffentlichten und zur Veröffentlichung vorgelegten Auftragsstudien vor (Kiss et al. 2002), für die Gestaltung von Fahrkartenautomaten liegen Erfahrungen in unveröffentlichten wie veröffentlichten Diplomarbeiten vor (Kosubek 1997, Pedell 1998, Deubzer & Pedell 1999). Die Studien legen den Schluss nahe, dass sich das Vorgehen auch hierfür eignet.

Grenzen des Einsatzbereiches und Nichteignung des Vorgehens 1) Mit diesem Leitfaden ermittelt man nur, wie man eine nutzerspezifische Anordnung entwickelt, die den kognitiven Erwartungen der NutzerInnen entspricht. Ob die angeordneten Objekte körperlich gut erreichbar sind, muss also gesondert geprüft werden. Das Vorgehen ersetzt anthropometrische oder biomechanische Untersuchungen nicht.

2) Das Vorgehen nützt das Gedächtnis der NutzerInnen. Es basiert auf der Annahme, dass die NutzerInnen ihre Erwartungen, wo sich ein Objekt oder eine Funktion befindet, spontan auf der Basis ihres begrifflichen Wissens bilden. Haben die NutzerInnen keine Handlungserfahrungen mit den Objekten, haben sie auch noch kein entsprechendes begriffliches Wissen darüber ausgebildet. Das Vorgehen eignet sich also nicht für Produkte und Funktionen, mit denen der Nutzer noch keine Handlungserfahrungen gemacht hat. Inwieweit man von Erkenntnissen über bekannte Produkte und Funktionen auf unbekannte schließen darf, ist noch nicht untersucht. Es wäre aber eine lohnende Fragestellung, denn gerade bei der Einführung neuer Produkte und Funktionen werden häufig wenig nutzerfreundliche Lösungen kreiert.

3) Das Gedächtnis der NutzerInnen zu untersuchen, stellt einen Weg dar. Diese Arbeit behauptet nicht, dass dies der einzige Weg für die Entwicklung nutzerorientierter, relationaler Anordnungen ist. Aber meiner Erfahrung nach ist dieser Weg schneller und kostengünstiger als z.B. Verhaltensbeobachtungen der NutzerInnen.

Untersuchungsansatz und Prognosen Praktisches Ziel ist es, eine nutzerorientierte Anordnung von Objekten zu entwickeln, also eine Anordnung, die die Such- und Findenprozesse der NutzerInnen verbessert. NutzerInnen bilden ihre spontane Erwartung, wo etwas zu finden ist auf der Basis ihrer Begriffsstrukturen. Es wird angenommen, dass eine Anordnung dann nutzerorientiert ist, wenn Erwartungen und Realität für den Nutzer übereinstimmen, wenn also eine strukturelle Übereinstimmung vorliegt zwischen der „mentalen“ Anordnung der Begriffe in der Begriffsstruktur der NutzerInnen und der „realen“ Anordnung der Objekte im Raum.

Untersuchungsziel ist es, an der betreffenden Nutzergruppe die Begriffsstruktur bezüglich der anzuordnenden Objekte zu erheben. Die Begriffsstruktur wird definiert (operationalisiert) als begrifflich-semantische Ähnlichkeit. Die Ergebnisse der vorliegenden Untersuchungen zeigen, dass sich Sortiertechniken als Methodenklasse für die Erhebung begrifflich-semantischer Ähnlichkeit von Objekten/Funktionen eignen. Die empirisch ermittelte Ähnlichkeit zwischen den Objekten wird dann in eine räumliche Anordnung übersetzt: Semantisch ähnliche Objekte werden entsprechend räumlich nahe beieinander angeordnet.

Basierend auf der Begriffstheorie von Klix (1971, 1992), Hoffmann 1983, 1986) et al. werden Prognosen aufgestellt, welchen Kriterien die semantische Ähnlichkeitsbeurteilung folgt:

- 1. Personen mit ähnlichen Handlungsvorerfahrungen stimmen in der semantischen Ähnlichkeitsbeurteilung überein.* Für eine Nutzergruppe, d.h. für Personen mit gleichen Handlungsvorerfahrungen etc., lässt sich eine personenübergreifende, stabile semantische Ähnlichkeitsbeurteilung für die zu untersuchenden Objektbegriffe finden. Personen mit unterschiedlichen Handlungsvorerfahrungen beurteilen die begrifflich semantische Ähnlichkeit der Objekte entsprechend unterschiedlich.
- 2. Der Ähnlichkeitsbeurteilung liegen handlungsbezogene Kriterien zugrunde.* Objekte/Funktionen werden semantisch ähnlich beurteilt, die a) für ähnliche Handlungsziele verwendet werden oder b) in gemeinsamen Handlungskontexten erlebt werden.

Objekte/Funktionen werden semantisch entsprechend unähnlich beurteilt, die a) für entsprechend unterschiedliche Handlungsziele verwendet werden oder b) in entsprechend unterschiedlichen Handlungskontexten erlebt werden.

3. *Die Ähnlichkeitsbeurteilung von Objekten ist unabhängig von der aktuellen Zielorientierung der Personen.* Die Ähnlichkeitsbeurteilung, d.h. die begrifflich semantische Nähe oder Distanz zwischen den Begriffen, ist nicht von der Instruktion abhängig, sofern die Personen in der typischen Nutzungsumgebung untersucht werden.
4. *Anordnungen von Objekten im Raum, die der Begriffsstruktur der NutzerInnen entsprechen, verbessern Such- und Findeprozesse.* Eine Anordnung von Objekten ist dann nutzerorientiert, d.h. sie unterstützt die Such- und Findeprozesse, wenn die relative räumliche Nähe oder Distanz zwischen den Objekten der begrifflich semantischen Nähe oder Distanz der Objektbegriffe bei der entsprechenden Nutzergruppe entspricht.

Untersuchungsdesign Im Folgenden wird beschrieben, wie man die Begriffsstrukturen von NutzerInnen erhebt, wie die Versuchspersonen auszuwählen sind und welche Erhebungs- und Auswertungsmethoden sich hierfür eignen.

Auswahl der Versuchspersonen Als Versuchspersonen dürfen nur NutzerInnen gewählt werden, die im Alltag bereits nutzerspezifische Erfahrungen/Handlungen mit dem Untersuchungsgegenstand gemacht *hatten* und somit bereits nutzerspezifisches, begriffliches Wissen über die anzuordnenden Objekte haben. Die Ergebnisse der vorliegenden Studien zeigen, dass die Begriffsstrukturen davon abhängig sind, welche Vorerfahrungen die NutzerInnen mit den Objekten gemacht haben. Für die Entwicklung einer nutzerorientierten Anordnung ist daher genau zu klären: Welche Personen nützen die Anordnung? Für welche Nutzergruppe soll die Anordnung realisiert werden? Eventuell ist es notwendig, hier erst einmal eine Voruntersuchung zu machen, um die Nutzergruppe kennen zu lernen bzw. zu definieren.

Ein Beispiel soll dies erläutern: Soll eine nutzerorientierte Anordnung für einen Supermarkt entwickelt werden, der aus Marketingsicht für eine breite Käuferschicht geplant (positioniert) wird, dann ist die Nutzergruppe der KundInnen entsprechend breit defi-

niert und auch so zu untersuchen. Soll hingegen ein „Geronto“-Lebensmittelmarkt entwickelt werden, der spezifisch die Bedürfnisse von „Personen über 70“ erfüllen soll, dann definiert sich die Nutzergruppe dementsprechend anders.

Da bei der Auswahl der Versuchspersonen auch Fehler gemacht werden können, sei hier ein Negativbeispiel beschrieben: Will man eine nutzerorientierte Anordnung z.B. für Tasten eines Fahrkartenautomaten entwickeln, dürfen z.B. nicht Konstrukteure oder Produktentwickler als Versuchspersonen ausgewählt werden, bloß weil sie u.U. auch mit öffentlichen Verkehrsmitteln fahren und Fahrkartenautomaten bedienen. Sie haben – bedingt durch ihren Beruf – auch ganz andere Handlungsziele und -erfahrungen mit dem Fahrkartenautomaten: z.B. wie sich Münzprüfer, Wechselgeldbehälter, Papierrolle, Drucker und andere technische Komponenten platzsparend, wartungsfreundlich, vandalensicher etc. konstruktiv anordnen lassen. Diese Handlungsziele und -erfahrungen schlagen sich in ihrem begrifflichen Wissen nieder. Es sollten daher immer „echte“ NutzerInnen als Versuchspersonen gewählt werden. (Außer man will Fahrkartenautomaten bauen, die für Konstrukteure von Fahrkartenautomaten bedienerfreundlich sind.) Hiermit wird auf eine gängige, aber nicht „nutzerorientierte“ Praxis im Gestaltungs- und Produktentwicklungsprozess hingewiesen: Für Gestalter ist das Gespräch im Kollegenkreis m. E. nach üblich für die Entwicklung von Gestaltungslösungen – und natürlich ist es wichtig. Aber es sei hier ausdrücklich darauf hingewiesen, dass dies zu einer Anordnung führen kann, die zwar für die Gestalter, aber meiner Erfahrung nach noch lange nicht für die NutzerInnen erwartungsgerecht ist. Um eine nutzerorientierte Anordnung zu entwickeln ist es notwendig, wirklich „typische“ NutzerInnen zu untersuchen.

Vorversuch zur Auswahl der zu untersuchenden Objekte Der Vorversuch dient dazu, aus der Grundgesamtheit der anzuordnenden Objekte oder Funktionen eine Auswahl für das Sortierexperiment zu treffen. Die ausgewählten Objekte oder Funktionen sollten max. 20 bis 30 Objekte oder Funktionen umfassen. Die genaue Zahl ist abhängig von der Sortiertechnik. Es zeigte sich, dass es wichtig ist zu kontrollieren, dass die ausgewählten Objekte oder Funktionen sowohl inhaltlich die Grundgesamtheit repräsentieren als auch den Versuchspersonen bekannt sind. Wenn Wortmarken verwendet werden, müssen diese eindeutig verstanden werden.

Die Studien dieser Arbeit brachten dabei folgende Erkenntnis: Die Auswahl allein durch die Liste mit Zufallszahlen (Bortz 1984, siehe Süßgebäck-Studie) ist nicht geeignet, da hier nicht kontrolliert werden kann, ob die ausgewählten Objekte tatsächlich die Grundgesamtheit der Objekte abdecken. Die Auswahl in den Hauptversuch zu integrieren (siehe Flughafenstudie) erwies sich auch nicht als optimal, weil sie eine hohe Belastung für die Versuchspersonen bedeutet und eine einfache Sortiertechnik (Free Card Sorting) fordert. Eine Auswahl alleine mittels restringierter Assoziation birgt gleiche Probleme, da die Auswertung stark subjektiv belastet ist.

Die Auswahl der zu untersuchenden Objekte empfiehlt sich in drei Schritten: Im *ersten Schritt* ist eine Auswahl mittels Zufallszahlen, restringierter Assoziation oder mittels einer Befragung von NutzerInnen zu treffen. (Eine Beschreibung der Verfahren findet sich im Kapitel 6: Entwicklung des Untersuchungsdesigns.) In einem *zweiten Schritt* sind die ausgewählten Objekte von Fachkräften bzw. Experten zu beurteilen, ob sie repräsentativ für die Gesamtmenge der anzuordnenden Objekte oder Funktionen sind. So ist sichergestellt, dass später Erkenntnisse für alle Objekte oder Funktionen, die nutzerorientiert angeordnet werden sollen, vorliegen. Die Fachkräfte bzw. Experten sind daher zu befragen, ob die ausgewählten Objekte oder Funktionen stellvertretend die Gesamtmenge abdecken. Hierfür empfiehlt sich eine offene, nondirektive Befragung und inhaltliche Auswertung. In einem *dritten Schritt* muss die Auswahl noch einmal von den NutzerInnen auf Bekanntheit beurteilt werden: Die NutzerInnen sind zu befragen, ob sie die ausgewählten Objekte oder Funktionen kennen. Unabhängig davon, ob Worte, Abbildungen oder die Objekte selbst verwendet werden, muss sichergestellt werden, dass sie den NutzerInnen bekannt sind. Hierfür reicht eine standardisierte Befragung und quantitative Auswertung.

Erhebung der Begriffsstruktur im Hauptversuch Im Hauptversuch wird die Begriffsstruktur der Versuchspersonen erhoben. Die Begriffsstruktur wird als semantische Ähnlichkeit operationalisiert. Ziel ist es herauszufinden, wie ähnlich oder unähnlich die Versuchspersonen die Begriffe beurteilen, in welcher Gruppierung und in welchem Beziehungsgefüge sie diese sehen und welche Klassifizierungskriterien sie den Urteilen zugrunde legen.

Erhebungsmethoden Für die Ähnlichkeitsbeurteilung eignen sich Sortiertechniken. Es darf dabei den Versuchspersonen kein Sortierkriterium vorgegeben werden, sondern es soll ja untersucht werden, welches sie verwenden. Als Sortiertechniken empfehlen sich Hierarchisches Sortieren, Free Card Sorting und Freies Sortieren. Eine ausführliche Darstellung dieser Sortiertechniken findet sich in Kapitel 6: Entwicklung des Untersuchungsdesigns. Alle drei Methoden erwiesen sich als gut geeignet in der Durchführung. Das Hierarchische Sortieren ist am aufwändigsten für die Versuchspersonen, es liefert aber laut Literatur die differenziertesten Ergebnisse. Free Card Sorting ist weniger aufwändig. Freies Sortieren ist am einfachsten für die Versuchspersonen.

Hierarchisches Sortieren Bei dieser Methode soll die Vp die Begriffe in zwei Gruppen teilen, diese Gruppen wieder in zwei Gruppen usw. Dieser Prozess setzt sich so lange fort, bis die Gruppen nur noch jeweils einen Begriff enthalten. Die Versuchsperson wird im Sinne einer "forced choice" gezwungen, sich bei jedem Sortierschritt auf ein Sortierkriterium zu beschränken. Sie kann aber bei jedem Sortierschritt ein anderes wählen. Dies hat den Vorteil, dass die Versuchsperson jeweils das ihr wichtigste Kriterium wählen kann. Diese Methode liefert besonders differenzierte Ähnlichkeitsbeurteilungen. Es ist aber eine Begrenzung der zu untersuchenden Objekte auf ca. 25 notwendig, da die Versuchspersonen sonst überlastet werden.

Free Card Sorting Bei dieser Methode soll die Vp die Begriffe zu Gruppen sortieren, die ihr vernünftig erscheinen. Sie kann Begriffe auch mehrfach zuordnen. Diese Methode erlaubt es, eine große Anzahl von Objekten beurteilen zu lassen (bis zu 100), was allerdings die Auswertung schwierig macht. Es empfiehlt sich, die Anzahl auf ca. 50 zu begrenzen.

Freies Sortieren Bei dieser Methode soll die Vp diejenigen Begriffe, die sie für ähnlich hält, nebeneinander legen wie bei einem Dominospiel, bis alle Begriffe in einer Reihe liegen. Doppelnennungen sind nicht möglich. Diese Methode liefert nicht so differenzierte Ähnlichkeitsbeurteilungen wie die beiden oben genannten Methoden, sie gilt als die schwächste Methode und benötigt in der Durchführung nur unwesentlich weniger Zeit als z.B. das hierarchische Sortieren. Daher wird sie hier auch nur bedingt empfohlen.

Es scheint den Aufwand nicht zu lohnen, die Versuchspersonen während des Sortierens nach ihren Sortierkriterien zu befragen. (Die Ergebnisse der Zimmerpflanzenstudie zeigten, dass das Sortierverhalten und die Nennungen nicht übereinstimmten. Die Sortierkriteriums-Nennungen halfen nicht, die Ergebnisse besser zu interpretieren.)

Auswertungsmethoden Die Sortiererergebnisse werden für jede Versuchsperson dokumentiert und in eine Distanzmatrix übertragen. Die Ähnlichkeitsurteile werden dann mit raum-analogen, dimensionsanalytischen Verfahren wie der Nonmetrischen Multidimensionalen Skalierung (NMDS) und der Hierarchischen Clusteranalyse (HCA) ausgewertet. Diese Auswertungsmethoden zeigten sich, wie erwartet, inhaltlich in Bezug auf die Untersuchungshypothesen als gut geeignet, weil sie der Fragestellung: Wie muss was im Raum plaziert sein? intuitiv entsprechen. Die NMDS wandelt psychologische Ähnlichkeitsurteile in räumliche Distanzmaße um. Sie bildet somit einen begrifflich, psychologischen Raum in einen geometrischen Raum ab. Die HCA wandelt psychologische Ähnlichkeitsurteile in Cluster um. Die Ergebnisse liegen in Form einer NMDS-Konfiguration und eines hierarchischen Cluster-Dendrogramms vor. Diejenigen Objekte, die als ähnlich beurteilt werden, liegen in der NMDS-Konfiguration nahe beieinander oder sind im Dendrogramm in einem Cluster. Entgegengesetztes gilt für unähnliche Begriffe. Es empfiehlt sich, die Daten mit beiden Methoden auswerten zu lassen: der zeitliche Mehraufwand ist unerheblich. Dafür liefern beide Methoden sich ergänzende Ergebnisse, weil beide Methoden mit unterschiedlichen mathematischen Ansätzen die Originaldaten verarbeiten. Die ausführliche Darstellung der beiden Methoden ist im Kapitel 6: Entwicklung des Untersuchungsdesigns nachzulesen.

Interpretation Ziel der Interpretation ist es, zu verstehen, nach welchen Kriterien die Nutzer bestimmte Objekte oder Funktionen als ähnlich beurteilt haben und andere nicht. Dies ist vor allem wichtig, wenn nicht alle anzuordnenden Objekte oder Funktionen untersucht wurden, sondern nur eine bestimmte Auswahl und der Rest noch „richtig“ zugeordnet werden muss.

Beim Interpretieren der NMDS-Konfigurationen und HCA-Cluster besteht die Gefahr, eigene Vorurteile und Interessen einfließen zu lassen und somit die Ergebnisse der NutzerInnen zu verzerren. Die Interpretation der Ergebnisse sollte aus den Erfahrungen

der vorgestellten Studien über zwei Wege erfolgen: a) entsprechend der theoretisch fundierten Hypothesen und b) in einer Nachbefragung der NutzerInnen. Was bedeutet dies konkret?

a) Interpretation entsprechend der theoretisch fundierten Prognosen: Basierend auf der Begriffstheorie von Klix, Hoffmann et al. wurde die Prognose gestellt, dass die Ähnlichkeitsbeurteilung nach handlungsbezogenen Kriterien erfolgen wird: Objekte oder Funktionen, die die NutzerInnen für das gleiche Handlungsziel oder in einem gemeinsamen Handlungskontext verwenden, werden als ähnlich beurteilt. Objekte oder Funktionen, die die NutzerInnen für unterschiedliche Handlungsziele oder in unterschiedlichen Handlungskontexten verwenden, werden entsprechend unähnlich beurteilt.

Bei der Interpretation der nahe beieinander liegenden Funktionen oder Objekte in der NMDS-Konfigurationen und der HCA-Cluster ist daher zu überlegen, welches das gemeinsame Handlungsziel oder der gemeinsame Handlungskontext ist.

b) Interpretation durch eine Nachbefragung der NutzerInnen Will man die eigene Interpretation erhärten oder prüfen, ob sie tatsächlich die Handlungsziele bzw. Handlungskontexte wiedergeben, die die NutzerInnen oder KundInnen erleben, ist es sinnvoll, einigen NutzerInnen die NMDS-Konfigurationen vorzulegen. Die Befragung sollte nondirektiv starten mit der Einstiegsfrage: „Warum liegen diese Funktionen oder Objekte nahe beieinander?“. Im Anschluss an den nondirektiven Teil sollten dann die beiden Fragenkomplexe folgen: „Wofür verwenden Sie diese Funktionen oder Objekte? Was ist das Gemeinsame?“ „In welcher Situation verwenden Sie diese Funktionen oder Objekte? Was ist das Gemeinsame?“

Eine Interpretation mittels eines Property Fittings wird nicht empfohlen, da es ein aufwändiges und inhaltlich wenig überzeugendes Verfahren ist (siehe Zimmerpflanzen-Studie). Es basiert auf den verbalen Sortierkriterien-Nennungen und wie sich zeigte stimmen die verbalen Nennungen, die während des Sortierens gegeben werden, nicht mit dem tatsächlichen Sortierverhalten überein.

Umsetzung in ein Gestaltungsbriefing Damit die Gestalter (z.B. Architekten, Regalbeleger, Konstrukteure, Designer) für die Objekte oder Funktionen eine räumliche und nutzerorientierte Anordnung realisieren können, benötigen sie ein Gestaltungsbriefing.

Um dieses Gestaltungsbriefing zu entwickeln, müssen als Erstes die Erkenntnisse, die an den 20-40 ausgewählten und untersuchten Objekten oder Funktionen gewonnen wurden, auf die Gesamtmenge übertragen werden. Das bedeutet, dass die nicht untersuchten Objekte und Funktionen den ermittelten Clustern und Gruppen zuzuordnen sind. Zur Absicherung kann die Anordnung noch einmal NutzerInnen zur Beurteilung vorgelegt werden. Es empfiehlt sich, eine offene Befragung z.B. mit den Fragen: „Sind diejenigen Objekte/Funktionen nahe beieinander, die für Sie zusammengehören?“ „Sind die Objekte/Funktionen für Sie logisch angeordnet?“ Es empfiehlt sich m.E. nicht, die NutzerInnen zu fragen, ob sie die Objekte oder Funktionen entsprechend angeordnet haben wollen. Diese Frage ist zu komplex, regt sie dazu an, Realisierungsprobleme zu bedenken und überfordert sie meiner Erfahrung nach.

In einem zweiten Schritt muss die neue „begriffsstruktur-orientierte“ Anordnung für die Gestalter aufbereitet werden. Es ist bereits viel darüber geschrieben worden, warum Gestalter ergonomische Erkenntnisse nicht nützen. Und es ist mindestes genauso viel darüber geschrieben worden, warum Ergonomen ihre Erkenntnisse nicht in einer Art und Weise präsentieren, die für die Gestalter lesbar, verstehbar und damit nutzbar ist (vgl. Green & Jordan 1999). Um es kurz zu machen: Es empfiehlt sich als Erstes zu klären, welcher Disziplin die Gestalter angehören und welche „Sprache und Darstellungsform“ sie gewöhnt sind. Schließlich sollte auch das Gestaltungsbriefing „nutzerorientiert“ sein. Kaufleute benötigen Text und Tabellen, Ingenieure und Konstrukteure bevorzugen m.E. das Lesen von Tabellen und schematischen Zeichnungen, Designer und Architekten bevorzugen m.E. Skizzen und Visualisierungen und möglichst wenig Text.

Neben einem schriftlichen Briefing benötigen die Gestalter auch eine mündliche Erläuterung im Rahmen eines Arbeitsmeetings oder eines Workshops. Eine persönliche Erläuterung übermittelt die Erkenntnisse nachhaltiger, erlaubt Fragen und gemeinsames Diskutieren.

Evaluation von Gestaltungsentwürfen in einer Expertenbeurteilung und in einer Nutzerstudie Nachdem die Gestalter erste Gestaltungsentwürfe entwickelt haben, sollte eine Überprüfung stattfinden. Die Überprüfung kann zunächst einmal in Form einer Expertenbeurteilung durch die Nutzerforscher (z.B. Psychologen, Ergonomen) stattfinden. Hierfür erhalten die Psychologen oder Nutzerforscher Entwürfe in Form von Skizzen,

Plänen oder ersten Modellen. Ziel dabei ist es zu beurteilen, ob die Gestalter die Vorgaben des Gestaltungsbriefings passend umgesetzt haben. Bei mehreren Lösungsvorschlägen muss ein Ranking durchgeführt werden, um die beste Gestaltung aus der Sicht der Nutzerforscher zu ermitteln.

Als nächster Schritt – der eventuell erst dann stattfindet, wenn die Ergebnisse der Expertenbeurteilung erneut umgesetzt wurden – bietet sich eine Überprüfung mittels einer empirischen Evaluationsstudie an. Ziel ist es, an „echten“ NutzerInnen diejenige Lösung zu ermitteln, die das Nutzerverhalten am besten unterstützt. So eine empirische Evaluation ist besonders wichtig, bevor die Anordnung festgeschrieben und multipliziert wird oder gar in Serie geht. Die Entwicklung eines Untersuchungsdesigns für eine Evaluationsstudie wird hier nicht näher beschrieben, denn sie hängt vom Anwendungsbereich ab: Im Anwendungsbereich Handel ist es möglich, eine neue Artikelanordnung in einem Testmarkt zu überprüfen. In der Produktentwicklung von Bedienungsoberflächen ist es möglich, NutzerInnen Aufgabensimulationen zu geben und über Reaktionsinhalt, Reaktionszeit und anderen Verhaltenskriterien die Funktionsanordnungen in puncto Nutzer-orientiertheit zu vergleichen und zu testen. In der Architektur ist es meist nicht möglich, unterschiedliche Grundrissvarianten für Untersuchungen mit NutzerInnen zu realisieren. Hier empfiehlt sich z.B. ein hypothesenprüfendes Verfahren; d.h. es werden auf der Basis der empirischen Ergebnisse Hypothesen aufgestellt, welche Anordnung Probleme bringen könnte oder welche gut funktioniert. Diese Hypothesen könnten an bestehenden Gebäuden überprüft werden.

Die Arbeit ist beendet, wenn eine Anordnung empirisch ermittelt wird, in der die NutzerInnen die besten Such- und Findeleistungen zeigen oder das Such- und Findeverhalten die vorher definierten Sollwerte zeigt.

Tabelle 10.1 Leitfaden zur Entwicklung erwartungsgerechter, nutzerspezifischer Anordnungen

Untersuchungsziel und Einsatzbereich

Ziel	Entwicklung erwartungsgerechter, nutzerspezifischer Anordnungen von Objekten/Funktionen im Raum, um die Such- und Findeprozesse der NutzerInnen zu verbessern.
Einsatzbereich	<p>Anordnung von Objekten/ Funktionen im Handel, in der Architektur, bei Bedienoberflächen, Arbeitsplätzen etc.</p> <p>Voraussetzung: Die NutzerInnen müssen Handlungsvorerfahrungen mit den Objekten oder Funktionen haben.</p> <p>Nicht geeignet für: Objekte und Funktionen, mit denen NutzerInnen keine Erfahrungen haben, sowie körperliche (anthropometrische/biomechanische) Aspekte der Erreichbarkeit.</p>

Theoretisches Fundament

Theoretische Annahmen	<p>NutzerInnen bilden auf der Basis ihrer Begriffsstrukturen Erwartungen darüber, wo ein Objekt/eine Funktion zu finden ist. Begriffsstrukturen repräsentieren Wissen über Handlungsvorerfahrungen, die die NutzerInnen gemacht haben.</p> <p>Eine Anordnung ist dann nutzerorientiert, wenn die räumliche Nähe der Objekte oder Funktionen der semantischen Nähe der entsprechenden Begriffe in der Begriffsstruktur entspricht. Oder anders formuliert: Was in der Begriffsstruktur nahe beieinander ist, sollte auch in der Realität entsprechend nahe beieinander sein, um leicht gefunden zu werden.</p>
Prognosen	<p>Basierend auf der Begriffstheorie von Klix, Hoffmann et al. werden folgende Prognosen für die Begriffsstruktur – operationalisiert als semantische Ähnlichkeitsbeurteilung – formuliert:</p> <ul style="list-style-type: none"> • Für Personen einer Nutzergruppe lässt sich eine personenübergreifende Ähnlichkeitsbeurteilung für die untersuchten Objekte/Funktionen finden. Personen unterschiedlicher Nutzergruppen unterscheiden sich in der Ähnlichkeitsbeurteilung. • Objekte/Funktionen werden semantisch sehr ähnlich beurteilt, die <ul style="list-style-type: none"> a) für sehr ähnliche Handlungsziele verwendet werden oder b) in einem Handlungskontext erlebt werden. <p>Objekte/Funktionen werden semantisch sehr unähnlich beurteilt, die</p> <ul style="list-style-type: none"> a) für unterschiedliche Handlungsziele verwendet werden oder b) nicht in einem gemeinsamen Handlungskontext erlebt werden • Die Ähnlichkeitsbeurteilung ist unabhängig von der Instruktion, sofern die Personen in der für die untersuchten Objekte/Funktionen typischen Umgebung untersucht werden. • Anordnungen von Objekten/Funktionen, die der Begriffsstruktur der NutzerInnen entsprechen, verbessern deren Such- und Findeprozesse.

Tabelle 10.1 Leitfaden zur Entwicklung erwartungsgerechter, nutzerspezifischer Anordnungen

Untersuchungsdesign für die Input-Studie

Auswahl der Versuchspersonen	Ziel: Es müssen echte NutzerInnen für die Untersuchungsstichprobe ausgewählt werden, die "nutzer-typische" Handlungsvorerfahrungen mit den Untersuchungsgegenständen gemacht haben.	
Vorversuch	Ziel: Bestimmung einer repräsentativen <u>und</u> bekannten Auswahl aus der Grundgesamtheit der zu untersuchenden Objekte/Funktionen mit NutzerInnen und Fachkräften	
	Erhebungsmethode	Auswertungsmethode
	bei NutzerInnen: - Zufallsauswahl - restringierte Assoziation - standardisierte Befragung - nondirektives Interview bei Fachkräften/ Experten: - nondirektives Interview	bei NutzerInnen: - quantitatives Auszählen bei Fachkräften/ Experten: - inhaltliche Auswertung
Hauptversuch	Ziel: Erhebung der begrifflich-semantischen Ähnlichkeit	
	Erhebungsmethode	Auswertungsmethode
	Sortiertechniken - Hierarchisches Sortieren - Free Card Sorting - Paarvergleich	deskriptive, raum-analoge, dimensions-analytische Verfahren - NMDS - Hierarchische Clusteranalyse
Interpretation	Ziel: Interpretation in Hinblick auf die Prognosen/Hypothesen Benennen der Klassifizierungskriterien Übertragen der Ergebnisse auf die Gesamtmenge der anzuordnenden Objekte oder Funktionen Evaluationsgespräche mit NutzerInnen und Fachkräften	
Umsetzung der Ergebnisse in ein Gestaltungsbriefing	Ziel: Aufbereiten der Ergebnisse als Vorgaben für das Gestaltungsbriefing unter Beachtung der spezifischen Sprache, Darstellungsgewohnheit etc. der jeweiligen Disziplin (Produktmanager, Marketingfachleute, Ingenieure und Konstrukteure, Designer, Architekten)	

Tabelle 10.1 Leitfaden zur Entwicklung erwartungsgerechter, nutzerspezifischer Anordnungen
Fortsetzung

Evaluationsstudie

Evaluation von Gestaltungsentwürfen	Ziel: Überprüfung der Gestaltungsentwürfe in einer Expertenbeurteilung oder in einer Nutzer-Evaluationstudie	
	Erhebungsmethode	Auswertungsmethode
	Verhaltensbeobachtung Fehleranalyse Befragung	quantitative und inferenzstatistische Verfahren

10.2 Speklatives Vorgehen zur Skizzierung nutzerspezifischer, erwartungsgerechter Anordnungen

Manchmal scheint es nicht möglich oder sinnvoll, eine an der Nutzergruppe empirisch fundierte Anordnung zu entwickeln, z.B. wenn die Anordnung von so wenigen Personen genutzt wird, dass sich eine Untersuchung wirtschaftlich nicht rechnet. Hierfür gibt es die Möglichkeit eines spekulativen Vorgehens (vgl. Tabelle 11.2), wobei mit aller Schärfe darauf hingewiesen wird, dass die erschlossene Anordnung auch alles andere als nutzerorientiert sein kann. Deshalb muss wirklich gut abgewogen werden, ob eine empirische Studie nicht doch Marketing- und Ertragsvorteile bringt. Kernsatz ist die theoretische Annahme, dass bei einer nutzerorientierten Anordnung die räumliche Nähe der Objekte der semantischen Ähnlichkeit der Objektbegriffe entsprechen sollte. Die semantische Ähnlichkeit von Begriffen bestimmt sich durch die Handlungsziele und Handlungen, die die NutzerInnen mit den Objekten haben oder ausgeführt haben. Es bietet sich daher an, die Handlungsziele und Handlungskontexte, welche die NutzerInnen mit den Objekten typischerweise vollziehen, entweder empirisch zu erfragen oder aufgrund von Erfahrungen mit der Nutzergruppe zu spekulieren. Es empfehlen sich folgende Fragen an die NutzerInnen: „Wozu verwenden Sie das Objekt...?“ „Welche Objekte verwenden Sie zusammen?“ Ein rein spekulatives Formulieren der Nutzerhandlungen und –handlungsziele beinhaltet die größten Gefahren. Es empfiehlt sich also eine

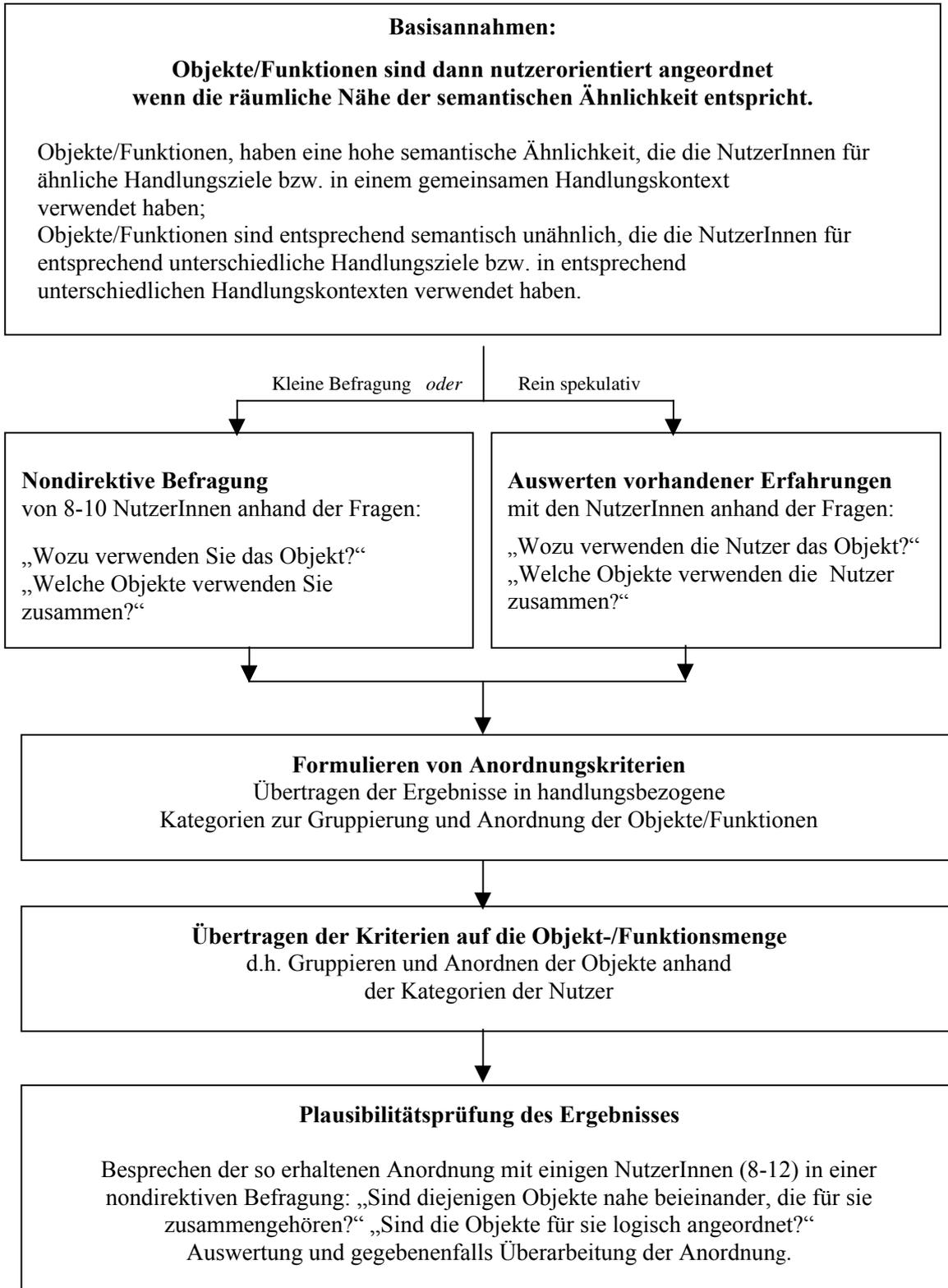
nicht repräsentative, qualitative Erkundungsstudie an 8-10 NutzerInnen. Die NutzerInnen sollten dabei nur nondirektiv im Interview befragt werden. Standardisierte Fragebögen geben nur die Kategorien wieder, die der Fragebogenentwickler implizit oder explizit unterstellt oder selbst im Kopf hat.

Aus den Handlungszielen bzw. Handlungen sind dann die Klassifizierungskriterien abzuleiten. Die Gesamtmenge der Objekte kann dann anhand dieser Klassifizierungskriterien in eine Anordnung übertragen werden. Es empfiehlt sich – zur Plausibilitätsprüfung – diese Anordnung dann wieder mit NutzerInnen – mindestens 8 NutzerInnen – einzeln zu diskutieren. Dabei sollten die NutzerInnen nicht gefragt werden, ob sie die Objekte entsprechend angeordnet haben wollen. Wenn man die NutzerInnen fragt, wie sie es angeordnet haben wollen, sind sie meiner Erfahrung nach überfordert. Sie wechseln zwischen Ähnlichkeits- und Machbarkeitsüberlegungen oder versuchen sich an das zu erinnern, was sie bereits kennen gelernt haben. Deshalb sollten die Fragen einfach sein. Man sollte sie fragen, ob die Anordnung für sie stimmig ist, ob die Objekte beieinander sind, die für sie zusammengehören, bzw. ob die Objekte für sie logisch angeordnet sind. Die NutzerInnen antworten meist spontan und wissen intuitiv ohne große weitere Erläuterung, worum es geht.

Dieses spekulative Vorgehen ist nicht überprüft worden. Es liegen keine Erkenntnisse bezüglich der Inter-Rater-Reliabilität oder Ähnlichem vor. Dieses Vorgehen wurde praktiziert für Food- und Nonfood-Sortimente in Verbrauchermärkten von DiplompsychologInnen, die bereits Erfahrungen mit dem empirischen Leitfaden hatten. Sie übertrugen die allgemeinen Erkenntnisse von Sortimenten wie Süßgebäck, Milchprodukten und kleinen elektrischen Haushaltsgeräten (unveröffentlichte Studien) auf andere Sortimente. Die Ergebnisse zeigten eben, dass nicht kategoriale Kriterien (z.B. Marke/Hersteller), sondern Handlungsziele und Verwendungszweck der Ähnlichkeitsbeurteilung zugrunde liegen. Da nicht für alle Sortimente aufwändige Sortierexperimente durchgeführt werden konnten, wurden die untersuchten Sortimente nach dem hier dargestellten spekulativen Vorgehen gruppiert und angeordnet. Tabelle 11.1 visualisiert das spekulativen Vorgehens.

Tabelle: 10.2. Visualisierung des spekulatives Vorgehens

Spekulatives Vorgehen zur Ermittlung einer nutzerorientierten Anordnung



10.3 Zusammenfassung

In diesem Kapitel wurde ein praxisorientierter Leitfaden zur theoretisch und empirisch fundierten Entwicklung von nutzerorientierten Anordnungen dargestellt. Der Leitfaden basiert auf den Erkenntnissen der vorgestellten Studien aus dem Bereich Handel und Architektur sowie unveröffentlichten Studien für Bedienoberflächen.

Das Vorgehen ist für Input- und Evaluations-Studien im Entwicklungsprozess geeignet, um den Gestaltern theoretisch fundierte und empirisch gesicherte Vorgaben für eine nutzerorientierte Anordnung der Objekte oder Funktionen zu liefern. Der Leitfaden beschreibt detailliert Untersuchungsziel, Möglichkeiten und Grenzen des Einsatzbereiches, das theoretische Fundament mit seinen Annahmen und Prognosen, die Entwicklung eines Untersuchungsdesign mit Auswahl der Versuchspersonen, Erhebungs- und Auswertungsmethoden für Vor- und Hauptversuch, das Vorgehen bei der Interpretation, die Umsetzung für Gestaltungsvorgaben und schließlich die Evaluation von Gestaltungsvarianten.

Zusätzlich zu dem empirischen Leitfaden wurde auch ein spekulatives Vorgehen skizziert. Es eignet sich – mit Vorbehalt – für Problemstellungen, bei denen eine empirische Studie z.B. aus Kostengründen nicht möglich ist. Es ist nicht überprüft.

11. Ausblick

Zu welchen Fragen und Untersuchungen regt diese Arbeit an? Meiner Meinung nach zeichnen sich fünf unterschiedliche Themen- und Untersuchungsbereiche ab, die es lohnen würde, weiter zu verfolgen.

Den Leitfaden zur Entwicklung erwartungsgerechter Anordnungen in weiteren Untersuchungen überprüfen und überarbeiten. Es bräuchte weitere Studien, um zu prüfen, ob Begriffsstruktur-orientierte Anordnungen das Such- und Findeverhalten unterstützen. Es liegen bereits einige – hier nicht dargestellte – empirische Untersuchungen vor, z.B. mit Fahrzeugcockpits (Kiss et al. 2002), Fahrkartenautomaten (Pedell 1998, Kosubek 1997) mit der Platzierung von Hausgeräten im Verkaufsraum (unveröffentlichte Studie für die Firma Bosch-Siemens-Hausgeräte). Aber es bräuchte weitere Felduntersuchungen und vor allem kontrollierte Laborexperimente. Sie fehlen bislang gänzlich. Auch sollten weitere Untersuchungen durchgeführt werden, um den Leitfaden zu überarbeiten und auch den Einsatzbereich zu klären.

Den Bezug zu den Gestaltgesetzen „Ähnlichkeit“ und „Nähe“ klären. In dieser Arbeit wurde der Bezug zum Gesetz der „Ähnlichkeit“ und zum Gesetz der „Nähe“ bewusst nicht thematisiert. Er hätte den Rahmen dieser Arbeit gesprengt! Doch es gibt ihn meiner Ansicht nach ganz eindeutig. Die Gestaltgesetze zeigen, dass die Wahrnehmung darauf ausgelegt ist, Strukturen zu suchen, zu erkennen und inhaltlich zu interpretieren. Eine Klärung des Bezuges zwischen den wahrnehmungsbezogenen Gesetzen der Ähnlichkeit und Nähe und den kognitionsbezogenen Gesetzen der begrifflichen Ähnlichkeit wäre m. E. wichtig, sowohl für die Anwendungsforschung als auch für die Grundlagenforschung.

Designer und Architekten lernen in ihrer Ausbildung (meist) die Gestaltgesetze kennen. In der Praxis versuchen sie dann diese anzuwenden. Ergebnis sind z.B. Bedienoberflächen, auf denen Funktionen zu Gruppen zusammengefasst sind. Doch nach welchen Kriterien fassen sie Funktionen zu Gruppen zusammen bzw. trennen sie? Dies machen sie m. E. nach eigenen Überlegungen. Das kann aber z.B. zur Folge haben, dass

Funktionen räumlich nahe platziert sind, die „inhaltlich“ nichts mit einander zu tun haben. Beim Nutzer entstehen aber durch die räumliche Nähe von Funktionen Erwartungen in Bezug auf eine inhaltliche Ähnlichkeit. Ein Beispiel soll dies verdeutlichen: Bei der Untersuchung eines Autoradios zeigte sich, dass die schwierigste Aufgabe für die Nutzer die Einstellung des Verkehrsfunks war. Bei der Analyse der Bedienoberfläche zeigte sich, dass die – für die Nutzer sehr wichtige – Verkehrsfunktaste mit anderen – ziemlich unwichtigen – Funktionstasten wie z.B. der Scanfunktion für CDs und der Mittelwellenfunktion gruppiert war und auch die gleiche Form, Größe, Farbe hatte. Die Gruppierung und Gestaltung ließ die Nutzer die Taste zum Einstellen des Verkehrsfunks nicht finden.

Damit Designer und Architekten die Gestaltgesetze der Ähnlichkeit und der Nähe sinnvoll realisieren können, brauchen sie empirisch fundierte Vorgaben, welche Funktionen aus Nutzersicht inhaltlich ähnlich bzw. unähnlich sind. Dann erst sind sie in der Lage die Funktionen räumlich „richtig“ zu platzieren und sie in Form, Größe und Farbe entsprechend ähnlich zu gestalten. Das Gesetz der Ähnlichkeit und das Gesetz der Nähe sind m. E. ohne empirisch fundierte Vorgaben (Input-Studien) nicht nutzerorientiert umsetzbar.

Für die Grundlagenforschung wäre es interessant, zu untersuchen, inwieweit die wahrnehmungsbasierten Gestaltgesetze der Ähnlichkeit und Nähe und Gesetze der Begriffsbildung miteinander korrespondieren. Ich habe die Vermutung, dass es sich um die gleichen Gesetzmäßigkeiten handelt, dass Wahrnehmungs- als auch Denkprozesse hier den gleichen Basisgesetzen folgen. Ich denke, es wäre lohnenswert, dies zu untersuchen.

Zentrale Kernprobleme der Gestaltung klären und dafür theoretisch und empirisch fundierte Lösungswege entwickeln. Diese Arbeit untersucht das Problem der *erwartungsgerechten Anordnung* von Funktionen im Raum. Ich denke, dass es sich bei diesem Problem um ein zentrales Kernproblem der nutzerorientierten Gestaltung handelt. Wann immer ein Gestalter (Designer, Architekt etc.) etwas gestalten muss, was mehr als drei Funktionen (Gegenstände, Handlungsziele) hat, ist er mit dem Problem konfrontiert: wo soll er die Funktion platzieren? Wo erwartet der Nutzer die Funktion?

Neben dem Anordnungsproblem gibt es m. E. aber auch noch andere zentrale Kernprobleme, z.B. die *nutzerorientierte Auswahl der Funktion*: Welche Funktionen braucht der Nutzer wirklich? Welche Funktionen unterstützen seine Handlungsziele? oder die *identifizierungsleichte Gestaltung der Funktion*: Woran erkennt der Nutzer die Funktion? Welches sind die identifizierungsrelevanten Hinweisreize? Ich denke, es würde sich lohnen 1) auf die Suche nach den gestalterischen Kernaufgaben zu gehen, diese zu definieren und 2) hierfür theoretisch und empirisch fundierte Vorgehensweisen zu entwickeln, wie man diese Kernaufgaben zur Zufriedenheit der Nutzer löst. Damit würden Praktiker eine grundlegend bessere Unterstützung erhalten als sie sie heute durch ISO-Normen und Design-Guidelines bekommen.

Begriffliches Wissens für andere Anwendungsprobleme nutzen. In dieser Arbeit werden Begriffsstrukturen untersucht und die Erkenntnisse werden zur Entwicklung erwartungsgerechter Anordnungen verwendet. Wenn begriffliches Wissen wesentliches, invariantes Handlungswissen repräsentiert, dann müsste man auch noch andere zentrale Gestaltungsprobleme mit Hilfe von Untersuchungen begrifflichen Wissens lösen können.

Da ist z.B. das meiner Ansicht nach zentrale Gestaltungsproblem der identifizierungsrelevanten Hinweisreize: wie muss die Funktion gestaltet sein (Form, Größe, Farbe, Oberfläche etc.), welche Gestaltungsmerkmale muss sie „tragen“, damit der Nutzer spontan erkennt, um welche Funktion es sich handelt. Wenn Begriffe bereits handlungs(ziel)relevante, invariante Merkmale einer Gegenstandsklasse repräsentieren, dann müsste man diese Merkmale „nur“ erheben und könnte damit eine Antwort auf die Frage bekommen, welche die identifizierungsrelevanten Hinweisreize der Funktionen sind.

Die Untersuchung der Begriffe bzw. der Begriffsstruktur von Nutzern ist m.E. weit aus ökonomischer und effektiver als Verhaltensbeobachtungen, Interviews oder Experimente. Geht man davon aus, dass im begrifflichen Wissen bereits wesentliches Handlungswissen abstrahiert, generalisiert und repräsentiert vorliegt, dann erscheint mir der direkte Weg der Beste, nämlich diese „Informationsquelle“ auch zu nützen und zu untersuchen.

Das methodische Vorgehen zur Lösung angewandter Probleme überprüfen und weiterentwickeln. In dieser Arbeit wurde in Kapitel 2 ein methodisches Vorgehen entwickelt, das vorschlägt, in welchen Schritten der Problemlösungsprozess durchgeführt werden und welchen Kriterien er folgen sollte. Es wäre wichtig, dieses methodische Vorgehen systematisch auf unterschiedlichste Anwendungsproblemen anzuwenden. Die Ergebnisse und Erfahrungen müssten ausgewertet werden und das vorhandene Vorgehen müsste entsprechend überarbeitet werden. Meiner Erfahrung nach brauchen angewandte arbeitende Psychologen, Nutzerforscher o.Ä. eine explizite Methodologie für ihre Problemlösungsprozesse. Und die muss weit über das hinausgehen, was Methodenkataloge, Design-Guidelines und ISO-Normen im Moment liefern.

Literaturverzeichnis

- Ach, N. (1921). Über Begriffsbildung. *Kongress Experimentelle Psychologie*, Bericht 7, S. 99
- Abelson, R.P. & Schank, R.C. (1977). *Scripts, plans, goals and understanding; An inquiry into human knowledge structures*. Hillsdale, New Jersey: Lawrence Erlbaum Associates
- Aebli, H. (1988). Begriffliches Denken. In: Mandl, H & Spada H (Hg.), *Wissenspsychologie* (S. 227-246). München: Psychologie Verlags Union
- Alisch, L.-M. (1991). Zur Systematisierung des Verhältnisses von Theorie und Praxis in der angewandten Psychologie, In: Witte, E.H. (Hg.), *Angewandte Sozialpsychologie: Das Verhältnis von Theorie und Forschung*. Braunschweig: Braunschweiger Studien zur Erziehungs- und Sozialarbeitswissenschaft
- Anderson, J.R. (1996). *Kognitive Psychologie*. Heidelberg: Spektrum Akademischer Verlag
- Anderson, J.R. (1993). *Rules of the Mind*. Hillsdale, New Jersey: Lawrence Erlbaum Associates
- Anderson, J.R. & Bower, G.H. (1973). *Human associative memory*. Washington, DC: Winston
- Appleyard, D. (1977). Understanding professional media; In: Altman, I., Wohlwill, J.F. (Eds.); *Human behavior and environment: Advances in Theory and Research*; New York, Plenum
- Banard, P. (1991). The Contribution of Applied Cognitive Psychology to the Study of Human-Computer Interaction. In: Shaker B. & Richardson S. (1991). *Human Factors for Informatics Usability* (pp. 151-182). Cambridge: University Press
- Barber, P. (1988). *Applied cognitive psychology* (pp. 7). London, New York: Methuen
- Barsalou, L.W. (1989) Intraconcept similarity and its implication for interconcept similarity. In: S. Vosniadou & A. Ortony (Eds.) *Similarity and analogical reasoning*; Cambridge: MA University Press, S. 76-121
- Bartlett, F.C. (1932). *Remembering: A Study in experimental and social psychology*. New York, London: Cambridge University Press
- Berker, C. (1989). *Kognitive Karten von Einkaufszentren – Die räumliche Beziehung verschiedener Einzelhandelsgeschäfte in einem Einkaufszentrum*. München: unveröffentl. Diplomarbeit
- Borg, I. (1981). *Anwendungsorientierte multidimensionale Skalierung*. Berlin, Heidelberg, New York: Springer, 1. Auflage
- Bortz, J. (1984). *Lehrbuch der empirischen Forschung*. 1. Auflage, Berlin, Heidelberg, New York: Springer
- Bortz, J. (1985). *Lehrbuch der Statistik*. 2. Auflage, Berlin, Heidelberg, New York: Springer

- Bortz, J. & Döring, N. (1995). *Forschungsmethoden und Evaluation für Sozialwissenschaftler*. 2. Auflage; Berlin, Heidelberg, New York: Springer
- Bredenkamp, J. & Erdfelder, E. (1996). Methoden der Gedächtnispsychologie. In: D. Albert & K.-H. Stapf (Hg.), *Gedächtnis – Enzyklopädie der Psychologie*. Serie I, Band 4, S. 337-431, Göttingen: Hogrefe
- Brewer, W.F. & Treyens, J.C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, 13, S. 207-230
- Brickell, C. (1999) *Dumonts große Pflanzen Enzyklopädie von A-Z*. Köln Dumont monte Verlag
- Brockhaus (2000). *Der Brockhaus in Text und Bild – Lexikon in der PC-Bibliothek*. Leipzig/Mannheim: Bibliographisches Institut & F.A. Brockhaus AG
- Bubb, H. (1992). *Menschliche Zuverlässigkeit*. Landsberg: Ecomed
- Bullinger, H.-J. (1994). *Ergonomie: Produkt und Arbeitsplatzgestaltung*. Stuttgart: Teubner (S. 272)
- Bungard, W. & Herrmann, T. Hrsg. (1993) *Arbeits- und Organisationspsychologie im Spannungsfeld zwischen Grundlagenorientierung und Anwendungen*. Bern: Huber
- Burmester, M., Hassenzahl, M. & Koller, F. (2002) Usability ist nicht alles – Wege zu attraktiven Produkten. In: *i-com – Zeitschrift für interaktive und kooperative Medien*, 1, S. 32-40
- Butler, D.L., Acquino A.L., Hissong A.A. & Scott R.A. (1993). Wayfinding by newcomers in a complex building. *Human Factors*, 35, S. 159-173
- Carroll, J.M. (1991). *Designing Interaction*. Cambridge: University Press
- Carroll, J.M. & Thomas, J.C. (1988). Fun. *SIGCHI Bulletin*, 19 (3), S. 21-24
- Chaffin, R. & Herrmann, D.J. (1984). The similarity and diversity of semantic relations. *Memory and Cognition*, 12, 134-141
- Chalmers, A.F. (1994). *Wege der Wissenschaft, Einführung in die Wissenschaftstheorie*. Berlin, Heidelberg, New York: Springer
- Chi, M.T.H. & Glaser, R. (1980). The measurement of expertise. In: Baker, E. L. & Quellmalz, E.S. (Hg.), *Educational testing and evaluation*. Beverly Hills
- Cohen, R., Baldwin, L.M. & Sherman, R.C. (1978). Cognitive maps of a naturalistic setting. *Child Development*, 49, S. 1216-1218
- Collins, A.M. & Quillians, M.R. (1972). Experiments on semantic memory and language comprehension. In: L.W. Gregg (Hg.), *Cognition and learning*. New York: John Wiley & Sons
- Collins, A.M. & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, S. 407-428
- Cooper, L.A. & Shepard, R.N. (1985). Rotationen in der räumlichen Vorstellung, In: *Spektrum der Wissenschaften*, 2
- Daly-Jones, O., Bevan, N. & Thomas, C. (1997). *Handbook of User-centered Design Information Engineering*. Usability Support Centre

- Deubzer, E. M. (1997). The Benefit from cognitive-psychological Modells for constructing userfriendly Human-Machine Interfaces. In: *Proceedings of the 16. European Annual Conference on Human Decision Making and Manual Control*, 9-11. 12. S. 55-63
- Deubzer, E. M. & Aschersleben, G. (2001). Wie muss eine nutzerorientierte Anordnung von Objekten im Raum aussehen? In: A. Zimmer et al. (Hg.): *Experimentelle Psychologie, Abstracts der 43. Tagung experimentell arbeitender Psychologen* (S. 25). Lengerich: Pabst Science Publishers
- Deubzer, E. M. & Pedell, S. (1999). Der Beitrag der kognitiven Psychologie zur benutzerfreundlichen Umsetzung der Gestaltgesetze auf Bedienoberflächen. In: *Experimentelle Psychologie, Abstracts der 41. Tagung experimentell arbeitender Psychologen* (S. 21). Lengerich: Pabst Science Publisher
- Dörner, D. (1989). *Die Logik des Mißlingens*. Hamburg: Rowohlt
- Dörner, D. & E. van der Meer (1995). *Das Gedächtnis: Probleme, Trends, Perspektiven*. Göttingen: Hogrefe
- Downs, R.M. & Stea, D. (1982). *Kognitive Karten: Die Welt in unseren Köpfen*. New York: Harper & Row
- Eason, K. & Harker, S. (1991) Human Factors contribution to the design process. In: Shaker, B. & Richardson, S. (Eds.). *Human Factors for Informatics Usability* (pp. 73-96). Cambridge: University Press
- Ebbinghaus, H. (1971). *Über das Gedächtnis*. Darmstadt: Wissenschaftliche Buchgesellschaft
- Eckes, T. (1996). Begriffsbildung. In: D. Albert & K.-H. Stapf (Hg.), *Gedächtnis – Enzyklopädie der Psychologie*. Serie I, Band 4, S. 273-319, Göttingen: Hogrefe
- Einstein, A. (1993). Foreword. In: M. Jammer, *Concepts of Space*. (pp. xi-xv). New York: Dover
- Engelkamp, J. (1990). *Das menschliche Gedächtnis*. Göttingen: Hogrefe
- Engelkamp, J. & Pechmann, T. (Hg.), (1993). *Mentale Repräsentation*. Bern: Hans Huber
- Flade, A. (1990). Kunde und Umwelt. In: Kruse, L., Graumann, C.F. & Lantermann, E.D. (Hg.), *Ökologische Psychologie*. München: Psychologie Verlags Union
- Flechtner, H.-J. (1972). *Grundbegriffe der Kybernetik*. Stuttgart: S. Hirzel
- Franks, J.J. & Bransford, F.D. (1971). Abstraction of visual patterns. *Journal of Experimental Psychology*, 76: S. 407-412
- Frese, M. & Zapf, D. (1991). Fehlersystematik und Fehlerentstehung – eine theoretische Einführung. In: *Fehler bei der Arbeit mit dem Computer*. Bern, Göttingen Toronto: Hans Huber
- Gaensslen, H. & Schubö, W. (1976). *Einfache und komplexe statistische Analyse*. München: Reinhardt utb
- Galton, F. (1879). Psychometric Experiments. *Brain*, 1, 149-162
- Gardiner, M.M. & Christie, B. (1987). *Applying cognitive psychology to user-interface design*. New York: John Wiley & Sons

- Gardner, H. (1989). *Dem Denken auf der Spur: der Weg der Kognitionswissenschaften*. Stuttgart: Klett-Cotta
- Gentner, D. (1989). Mechanisms of analogical learning. In: S. Vosniadou & A. Ortonoy (Hg.), *Similarity and analogical reasoning*. S. 197-241. Cambridge: Cambridge University Press
- Gibson, J.J. (1973). *Die Wahrnehmung der visuellen Welt*. Weinheim: Beltz
- Gibson, J.J. (1982). *Wahrnehmung und Umwelt. Der ökologische Ansatz in der visuellen Wahrnehmung*. München: Urban & Schwarzenberg
- Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie*. München: Reinhardt
- Graesser, A.C., Robertson, S.P. & Anderson, P.A. 1981. Incorporating inferences in narrative representations a study of how and why. *Cognitive Psychology*, 13, S. 1-26
- Grandjean (1979). *Physiologische Arbeitsgestaltung – Leitfaden der Ergonomie*. Thun: Ott-Verlag
- Green, W.S. & Jordan, P.W. (1999). Ergonomics, Usability and Product Development. In: Green, W.S. & Jordan, P.W. *Human factors in product design*. London: Taylor & Francis
- Greif, S. (1993). Mensch-Computer-Interaktion, psychologische Theorien und innovatives Problemlösen; In: Bungard W. & Hermann, T. (Hg.) *Arbeits- und Organisationspsychologie im Spannungsfeld zwischen Grundlagenorientierung und Anwendung*. Bern: Hans Huber
- Guski, R. (1989). *Wahrnehmung*. Stuttgart: Kohlhammer
- Hacker, W. (1986). *Arbeitspsychologie*. Bern: Hans Huber
- Hejj, A. & Strube, G. (1988) Wortfeld im Wandel. In: Marx *Verbales Gedächtnis und Informationsverarbeitung*. Göttingen: Hogrefe
- Henley, N.M. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior*, 8, 176-184
- Herrmann T. (1993) Zum Grundlagenwissenschaftsproblem der A&O Psychologie; In: Bungard W. & Hermann, T. (Hg.) *Arbeits- und Organisationspsychologie im Spannungsfeld zwischen Grundlagenorientierung und Anwendung*. Bern: Hans Huber
- Herrmann, T. (1993). Mentale Repräsentation – ein erläuterungswürdiger Begriff. In: Engelkamp, J. & Pechmann, T. (Hg.), *Mentale Repräsentation*. Bern: Hans Huber
- Herrmann, T. (1982). Über begriffliche Schwächen kognitivistischer Kognitionstheorie: Begriffsinflation und Akteur-System-Kontamination. *Sprache und Kognition*, 1, S. 3-14
- Herrmann, T. (1979). *Psychologie als Problem*. Stuttgart: Klett-Cotta
- Herrmann, T. & Schweizer, K. (1998). *Sprechen über Raum: sprachliches Lokalisieren und seine kognitiven Grundlagen*. Bern: Hans Huber
- Hoffmann, J (1996). Die Genese von Begriffen, Bedeutungen und Wörtern. In: Grabowski, J. Harras G. Herrmann T. (Hg.), (1996). *Bedeutungen, Konzepte, Bedeutungskonzepte*. S. 88-120, Opladen: Westdeutscher Verlag

- Hoffmann, J. (1996). Die visuelle Identifikation von Objekten. In: W. Prinz & B. Bridgeman (Eds.). *Handbook of Perception and Action*. Vol I: Perception. London, San Diego, New York: Academic Press
- Hoffmann, J. (1995). Gedächtnis und Verhaltensorganisation. In: D. Dörner & E. van der Meer (Eds.). *Gedächtnis*. Berlin, Heidelberg, New York: Springer
- Hoffmann, J. (1993). *Vorhersage und Erkenntnis*. Göttingen: Hogrefe
- Hoffmann, J. (1993). Unbewußtes Lernen – eine besondere Lernform? *Psychologische Rundschau* 44, S. 75-89
- Hoffmann, J. (1986). *Die Welt der Begriffe*. Berlin: Deutscher Verlag der Wissenschaften
- Hoffmann, J. (1983). *Das aktive Gedächtnis*. Berlin, Heidelberg, New York: Springer
- Hoffmann, J. (1980). Semantische Relationen und Gedächtnisstrukturen. In: A. Kosakowski (Hg.), *Psychologie im Sozialismus*. Berlin: Deutscher Verlag der Wissenschaften
- Hoffmann, J. & Ziebler, M. (1982). Begriffe und ihre Merkmale. *Zeitschrift für Psychologie*; 190, S. 46-77 zitiert nach Hoffmann, J. (1986). *Die Welt der Begriffe*. Berlin: Deutscher Verlag der Wissenschaften
- Hoffman, J. Ziebler, M. Grosser, U. (1984). Psychologische Gesetzmäßigkeiten der begrifflichen Klassifikation von Objekten In: Klix, F. (Hg.) *Wissensrepräsentation und Wissensnutzung*. Berlin: Deutscher Verlag der Wissenschaften
- Hoffman, J. Ziebler, M. Grosser, U. Kämpf, U. (1985). Struktur und Prozeßkomponenten in begrifflichen Identifikationsleistungen. In: *Zeitschrift für Psychologie*, 193, S. 51-69
- Hoffmann, J. & Grosser, U. (1985). Automatismen begrifflicher Klassifikation. In: *Sprache & Kognition*, 4, S. 28-48
- Hoffmann, J. & Klein, R. (1988). Kontexteffekte bei der Benennung und Entdeckung von Objekten. In: *Sprache & Kognition*, 7, S. 25-39
- Hütte (1991). *Grundlagen der Ingenieurwissenschaften*. 29. Auflage; Berlin: Springer
- Hull, C.L. (1920). Quantitative aspects of the evolution of concepts: An experimental study. In: *Psychocological Monogr.* 28, 123
- ISO 13407 (1999, E). Human-centered design processes for interactive systems. International Organization for Standardization
- ISO 9241 1-17 (1998, E). Ergonomic requirements for office work with visual display terminals (VDTs) International Organization for Standardization
- Jammer M. (1994). Gesetz. In: Seiffert, H., Radnity, G. (1994). *Handlexikon zur Wissenschaftstheorie*. 2. Auflage: München: dtv
- Johannsen G.(1993). *Mensch-Maschine-Systeme*. Berlin, Heidelberg, New York: Springer
- Johnson, S.C. (1967). Hierarchical Clustering Schemes. *Psychometrica*, 32, S. 241-154

- Johnson-Laird, P.N. (1989). Analogy and the exercise of creativity. In: S. Vosniadou & A. Ortonoy (Hg.), *Similarity and analogical reasoning* (pp. 313-331). Cambridge: Cambridge University Press
- Johnson-Laird, P.N. (1983). *Mental Models – Towards a Cognitive Science of Language, Inferences, and Consciousness*. Cambridge Massachusetts: Harvard University Press
- Kaplan R. & Kaplan S. (1982). *Cognition and Environment: Functioning in an uncertain world*. New York: Praeger Publishers
- Kanis H. (2000). Research for Usage oriented Design: Quantitative? Qualitative? In: *Proceedings of the IEA 2000 /HFES 2000*. Congress 6 (pp. 925-928)
- Kanis H. (1999). Design Centred Research into User Activities. In: Green, W.S., Jordan, P. W. (Eds.), *Human Factors in Product Design* (pp. 36 – 46). London: Taylor & Francis
- Kanis. H & Arisz, H.J. (2000). How many participants: a simple means for concurrent monitoring. In: *Proceedings of the IEA/HFES 2000 Congress*, 6, S. 637-640
- Kanis, H, Weegels, M.F. & Steenbekkers, L.P.A (1999). The uninformative-ness of quantitative research for usability focused desing of consumer products. In: *Proceedings of the Human Factors and Ergonomics Society*, 43 Annual Meeting, S. 481-485
- Kanis, H., Weegels, M.F. & Green, W.S. (1999). Scientific Research in a Design Context. In: *Contemporaray Ergonomics*, S. 374-378
- Kinateder, P. (1987). *Klassifizierungsleistungen von Erwachsenen bei gerichteten Gegenständen am Beispiel Süßgebäck*. München: unveröffentl. Diplomarbeit
- Kintsch, W. (1974). *The representation of meaning in memory*. New York
- Kiss, M., Wittmann, M., Gugg, P., Neidhart, M., Pöppel, E., Steffen, A., Deubzer, E., & Hiroyuki. K. (2002). Cognitive Map as a way to access the mental representation of the Car Cockpit (Manuskript in Vorbereitung)
- Kirchner, A., Kirchner, J.H., Kliem, M. & Müller J.M. (1990). *Räumlich-ergonomische Gestaltung*. Dortmund: Schriftenreihe der Bundesanstalt für Arbeitsschutz, Wirtschaftsverlag NW
- Klimesch, W. (1988). *Struktur und Aktivierung des Gedächtnisses*. Bern: Huber
- Klimesch, W. & Pfurtscheller, G. (1991). Zur Lokalisation episodischer und semantischer Information. In: Klix, F., Roth, E. & van der Meer, E. (Hg.), *Kognitive Prozesse und geistige Leistung* (pp. 30-41). Berlin: Deutscher Verlag der Wissenschaften
- Klix, F. (1998). Begriffliches Wissen – episodisches Wissen. In: Klix, F. & H. Spada (Hg.), *Kognition – Enzyklopädie der Psychologie*. Serie II, Band 6, S. 167-211 Göttingen: Hogrefe
- Klix, F. (1992). *Die Natur des Verstandes*. Göttingen: Hogrefe
- Klix, F. (1991). Über notwendige Kriterien eines psychologisch relevanten Modells der Wissensrepräsentation. In: *Zeitschrift für Psychologie*. Suppl.-Band 11. Leipzig: Barth-Verlag

- Klix, F. (1988). Gedächtnis und Wissen. In: H. Mandl & H. Spada (Hg.), *Wissenspsychologie*, S. 19-54, München: PVU
- Klix, F. (1984). *Gedächtnis, Wissen, Wissensnutzung*. Berlin: Deutscher Verlag der Wissenschaften
- Klix, F. (1976). Über Grundstrukturen und Funktionsprinzipien kognitiver Prozesse. In: F. Klix (Hg.), *Psychologische Beiträge zur Analyse kognitiver Prozesse*. Berlin: Deutscher Verlag der Wissenschaften
- Klix, F. (1971). *Information und Verhalten*. Bern: Hans Huber
- Klix, F., van der Meer, E., Preuss, M. (1984). Semantische Relationen: Erkennungsaufwand und psychophysiologische Reaktionstendenzen. In: Klix, F. (Hg.) *Gedächtnis, Wissen, Wissensnutzung*. Berlin: Deutscher Verlag der Wissenschaften
- Klix, F., E. Roth, E. van der Meer (1991). *Kognitive Prozesse und geistige Leistung*. Berlin: Deutscher Verlag der Wissenschaften
- Kluwe, R. H. (1995). Methoden der Psychologie zu Gewinnung von Daten über menschliches Wissen: In: D. Dörner & E. van der Meer: (Hg.), *Das Gedächtnis* (S. 359-385). Göttingen: Hogrefe
- Kluwe, R. (1991). Zum Problem der Wissensvoraussetzung für Prozess- und Systemsteuerung. In: *Zeitschrift für Psychologie*, Suppl-Band 11, Leipzig: Barth-Verlag
- Kluwe, R.H. (1988). Begriffliches Denken. In: Mandl, H. & Spada, H. (Hg.), *Wissenspsychologie*. S. 359-385; München: Urban und Schwarzenberg
- Knauff, M. (1997). *Räumliches Wissen und Gedächtnis: zur Wissenspsychologie des Raums*. Wiesbaden: Dt. Universitäts-Verlag
- Knauff, M., Rauh, R., Schlieder, C. & Strube, G. (1996). Analogizität und Perspektive in räumlichen mentalen Modellen. In: C. Umbach, M. Grabski & R. Hornig (Hg.), *Perspektive in Sprache und Raum*. Wiesbaden: Deutscher Universitätsverlag
- Knauff, M. Hörnig, R. Rauh, R. & Strube G. (1992). Kontextsensitivität von Ähnlichkeitsurteilen: Empirie und Modellierung. *Proceedings of Workshop: Ähnlichkeit von Fällen in Systemen fallbasierten Schließens*. Universität Kaiserslautern: SEKI Report
- Koller, R. (1994). *Konstruktionslehre für den Maschinenbau*. 3. Auflage; Heidelberg, New York, Berlin: Springer
- Kosslyn, H. (1975). Information representation in visual images. *Cognitive Psychology*, 7, S. 341-370
- Kosubek, K. (1997). *Untersuchung der Hypothesenbildung mit anschließender Modellentwicklung am Beispiel Fahrkartenkauf am Automaten*, unveröffentlichte Diplomarbeit an der Humboldt-Universität, Berlin
- Kröber-Riehl W. & Weinberg, P. (1999). *Konsumentenverhalten*. München: Vahlen
- Kruse, L., Graumann, C.F. & Lantermann, E.D. (1990). *Ökologische Psychologie, Ein Handbuch in Schlüsselbegriffen*. München: Psychologie Verlags Union

- Küchler, E. (1991). Über elementare Vorhersagemechanismen in der menschlichen Informationsverarbeitung. In: F. Klix, E. Roth & E. van der Meer (Hg.), *Kognitive Prozesse und geistige Leistungen* (pp. 148-169). Berlin: Deutscher Verlag der Wissenschaften
- Küttner, M. (1994). Prognose, Voraussage, In: Seiffert, H., Radnitzky, G. (1994). *Handlexikon zur Wissenschaftstheorie*. 2. Auflage; München: dtv
- Kukla, F. & Sydow, H. (1970). Zur mehrdimensionalen Skalierung phänomenaler Ähnlichkeiten bei komplexem Reizmaterial. *Zeitschrift für Psychologie*, 178
- Labov, W. (1973). The boundaries of words and their meaning. In: C.J.N. Bailey & R. W. Shuy (Hg.), *New ways of analyzing variations in English*. Washington, DC: Georgetown University Press
- Landauer, T.K. (1985). Psychological research as a mother of invention. *Proceedings of CHI'85: Human Factors in Computer Systems*. New York: ACM 44.
- Lantermann, E.D. (1980). *Interaktionen – Person, Situation, Handlung*. München: Urban & Schwarzenberg
- Larkin, J.H. et al. (1979). Experts and novice performance in solving physics problems. *Science*, 208, S. 1335-1342
- Leontjev, A.N (1977). *Tätigkeit, Bewußtsein, Persönlichkeit*. Stuttgart: Klett
- Levine, M., Marchon I. & Hanley G. (1984). The placement and displacement of you-are-here-maps; *Environment and Behavior*, 16, 139-157
- Lewis, C. (1991) Inner and outer theory in human-computer interaction. In: Carroll J.M. (1991) *Designing Interaction*. Cambridge, Massachusetts: MIT Press
- Lindsay, P.H. & Norman, D.A. (1972). *Human information processing*. New York: Academic Press
- Luczak, H. (1993). *Arbeitswissenschaft*. Berlin, Heidelberg, New York: Springer
- Lynch, K. (1960). *Image of the city*. Cambridge, Massachusetts: MIT Press
- Maichle, U. (1985). *Wissen, Verstehen und Problemlösen im Bereich der Physik*. Frankfurt: Verlag Peter Lang Europäische Hochschulschriften, Reihe 6 Bd. 150
- Mandl, H. & Spada, H. (1988). *Wissenspsychologie*. München: Urban & Schwarzenberg
- Marx, W. (1988). *Verbales Gedächtnis und Informationsverarbeitung*. Göttingen, Toronto, Zürich: Verlag für Psychologie Dr. C.J. Hogrefe
- Marx, W. (1982). Das Wortfeld der Gefühlsbegriffe. In: *Zeitschrift für experimentelle und angewandte Psychologie*, 29, S. 137-146
- Marx, W. & Hejj, A. (1989). *Subjektive Strukturen: Ergebnisse aus der Gedächtnis-, Sprach- und Einstellungsforschung*. Göttingen, Toronto, Zürich: Hogrefe
- Mayhew, D. J. (1999). *The Usability Engineering Lifecycle – A Practitioner's Handbook for User Interface Design*. San Francisco: Morgan Kaufmann Publishers
- Meer, E. van der (1998). Inferenzen in Wissenskörpern. In: Klix, F. & H. Spada (Hg.), *Kognition – Enzyklopädie der Psychologie*. Serie II, Band 6, S. 213-247, Göttingen: Hogrefe

- Meer, E. van der (1995). Gedächtnis und Inferenzen. In: D. Dörner & E. van der Meer: (Hg.), *Das Gedächtnis* (S. 341-38). Göttingen: Hogrefe
- Meer(E. van der (1991). Die dynamische Struktur von Ereigniswissen (pp. 41.64). In: Klix, E. Roth, E. van der Meer, *Kognitive Prozesse und geistige Leistung*. Berlin: Deutscher Verlag der Wissenschaften
- Mehrtens, A. (1999). Methode/Methodologie, In: Sandkühler, H.J. (Hg.), *Enzyklopädie der Philosophie*. Bd. 1, Hamburg: Felix Meiner
- Metzger.W. (1975). *Gesetze des Sehens*. Frankfurt: Waldemar Kramer
- Miller. G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing informations. In: *Psychological Review*, 63, S. 81-97
- Minsky, A. (1975). A framework for representing knowledge. In: Winston P.H. (Ed.). *The Psychology of Computer Vision*. New York, McGraw-Hill: S. 211-277
- Murdock, B.B.(1974). *Human Memory. Theory and Data*. New York: Academic Press
- Neumann, O. (1996). Theorien der Aufmerksamkeit. In: Bierbaumer, N. (Hg.). *Kognition -Enzyklopädie der Psychologie*. Serie II, Bd. 2 S. 559-643 Göttingen: Hogrefe
- Nielsen, J (1993). *Usability Engineering*. London: AP Professional
- Nisbett, R. E. (Ed.). (1993). *Rules for Reasoning*. Hillsdale, NJ: Erlbaum
- Norman, D. A (1989). *The Design of Everyday Things*. New York: Doubleday
- Norman, D. A. & Rumelhart, D.E. (1975). *Explorations in cognition*. San Francisco: Freeman
- Opwis, K. (1992). *Kognitive Modellierung*. Bern: Oswald
- Opwis, K. (1988). Produktionssysteme. In: H. Mandl & H. Spada (Hg.), *Wissenspsychologie*. S. 74-98; München: PVU
- Opwis, K. & Lürer, G. (1998). Modelle der Repräsentation von Wissen. In: D. Albert & K.-H. Stapf (Hg.), *Gedächtnis – Enzyklopädie der Psychologie*. Serie I, Band 4, S. 337-431 Göttingen: Hogrefe
- Packard, V. (1957). *Die geheimen Verführer*. Frankfurt am Main: Ullstein
- Pahl, G. & Beelich, K.H. (1981). Lagebericht – Erfahrungen mit dem methodischen Konstruieren, In: *Zeitschrift für Maschinenbau, Konstruktion und Fertigung*. 114. 11, 773-840
- Pahl, G. & Beitz, W. (1997). *Konstruktionslehre – Methoden und Anwendung*. 4. Auflage; Berlin: Springer
- Pedell, S. (1998). *Die Bedeutung der Gestaltgesetze Ähnlichkeit und Nähe für die Gestaltung von Benutzeroberflächen*, unveröffentlichte Diplomarbeit an der Technischen Universität Berlin
- Petermann, F. (1980). *Einstellungsmessung und -forschung*. Göttingen: Hogrefe
- Piaget, J. & Inhelder, B. (1971). *Die Entwicklung des räumlichen Denkens beim Kinde*. Stuttgart: Ernst Klett
- Popper, K.R. (1976). *Die Logik der Forschung*, Tübingen: J.C.B. Mohr

- Popper, K.R. (1974). *Die offene Gesellschaft und ihre Feinde*. Vol. II, München: UTB 4. Auflage
- Popper, K. R. (1973). *Objektive Erkenntnis*. Hamburg: Kiepenheuer & Witsch
- Preuß, M. (1991). Experimente über Relationserkennungen im menschlichen Gedächtnis. Jena: Friedrich Schiller Universität. zitiert nach F. Klix (1992). *Die Natur des Verstandes*. Göttingen: Hogrefe
- Puffe, H. (1980). Untersuchungen zu Unterscheidbarkeit semantischer Relationen im menschlichen Langzeitgedächtnis. Dipl.-Arbeit Humboldt-Universität In: F. Klix (1992). *Die Natur des Verstandes*. Göttingen: Hogrefe
- Purcell, A.T. (1986). The organisation of the experience of the built environment; *Environment and Planning*, B 11, 193-212
- Pylyshyn Z.W. (1991) Some Remarks on the Theory-Practice Gap. In: Carroll F.M. (1991) *Designing Interaction – Psychology at the Human-Computer Interface*. Cambridge, New York: Cambridge University Press
- Radnitzky, G. (1994). Wissenschaftstheorie, Methodologie. In: Seiffert, H., Radnitzky, G. (1994). *Handlexikon zur Wissenschaftstheorie*. 2. Auflage; München: dtv
- Rips, L.J. (1995). The current status of research on concept. *Mind and Language*. 10, S. 72-104
- Rips, L.J. (1989). Similarity, typicality, and categorization; In: S. Vosniadou & A. Ortonoy (Hg.), *Similarity and analogical reasoning*. Cambridge: Cambridge University Press
- Rips, L.J., Shoben, E.J. & Smith, E.E. (1973). Semantic Distance and the Verification of Semantic Relations. *Journal of Verbal Learning and Verbal Behavior*. 12, S. 1-20
- Rosch, E. (1977). Human categorization. In: N. Warren (Ed.). *Advances in cross cultural Psychology*, London: Academic Press, Vol. I., S. 3-49
- Rosch, E. & Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 8, S. 382-439
- Ryle, G. (1949). *The concept of mind*. Harmondsworth: Penguin
- Sarris, V. (1990). *Methodologische Grundlagen der Experimentalpsychologie*. München, Basel: UTB Reinhardt
- Schank, R.C. (1982). *Dynamic memory: A theory of reminding and learning in computers and people*. New York: Cambridge University Press
- Schank, R.C. & Abelson, R.P. (1977). *Scripts, goals, plans and understanding*. Hillsdale: Erlbaum
- Schmiescheck, M. (1998). Die Strukturbeschreibung von ereignisbestimmtem Wissensbesitz im menschlichen Langzeitgedächtnis. Dissertation Humboldt-Universität (unveröffentlicht), zitiert nach: F. Klix (1992). *Die Natur des Verstandes*. Göttingen: Hogrefe
- Schmidtke, H. Hrsg. (1993). *Ergonomie*; München: Hansa Verlag, 3. Auflage

- Schneider, G. (1990). Methoden der Umweltrepräsentation, zitiert in: Kruse, L., Graumann, C.-F. & Lantermann, E.-D. (1990). *Ökologische Psychologie: Ein Handbuch in Schlüsselbegriffen*. München: Psychologie Verlags Union
- Schnell, R., Hill, P.B., Esser, E. (1992). *Methoden der empirischen Sozialforschung*. 3. Auflage; München, Wien: Oldenburg
- Schöne, D. (1980). *Orientierung im Raum: Formen und Mechanismen der Lenkung des Verhaltens im Raum bei Tier und Mensch*. Stuttgart: Wissenschaftliche Verlagsgesellschaft
- Schroeder-Köstler, B. (1990). *Kognitive Repräsentation von Zimmerpflanzen*. München: unveröffentl. Diplomarbeit
- Seidel, A. (1983). Wayfinding in public spaces: The Dallas Fort-Worth Airport; In: Amedeo D., Griffin J., Potter J. (Eds.). *Proceedings of the 14th Annual Meeting of the Environmental Design Research Association*; Washington, DC
- Seiffert, H. (1994) Theorie. In: Seiffert, H., Radnity, G. (1994). *Handlexikon zur Wissenschaftstheorie*. 2. Auflage; München: dtv
- Shakel, B. & Richardson, S. (1991). *Human Factors for Informatics Usability*. Cambridge: University Press
- Shakel, B. (1991). Usability – Context Framework, Definition, Design and Evaluation. In: Shakel, B. & Richardson, S. (1991). *Human Factors for Informatics Usability*. Cambridge: University Press, (pp. 21-38)
- Shneiderman, Ben (2002). *User Interface Design*. Bonn: mitp-Verlag
- Siegel, S. (1976). *Nichtparametrische statistische Methoden*. Frankfurt: Verlagsabteilung Fachbuchhandlung für Psychologie
- Smith, E.E. & Medin, D.L. (1981). *Categories and Concepts*. London: Pergamon Press
- Smith, E.E., Shoben, E.J. & Rips, L.J. (1974). Structure and process in semantic memory. A featural model for semantic decisions. *Psychological Review*. Vol. 81, No. 3, S. 214-241
- Stern, W. (1930). *Studien zur Personenwissenschaft, Erster Teil: Personalistik als Wissenschaft*. Leipzig: Barth
- Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction-time experiments, *American Scientist*. 57, S. 421-457,
- Städtler, T. (1998). *Lexikon der Psychologie*. Stuttgart: Körner
- Stegmüller, W. (1973). *Theorie und Erfahrung*. (2.Halbb.). Berlin: Springer
- Sternberg, S. 1969). The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica*. 30, S. 276-315
- Sternberg, S. (1969). Memory Scanning: Mental processes revealed by reaction time experiments. *American Scientist*. 57, S. 421-457
- Stevens, A. & Coupe, P. (1978). Distortions in judged spatial relations; *Cognitive Psychology*, 10, S. 422-437
- Steyer, R. & Eid, M. (1993). *Messen und Testen*. Berlin, Heidelberg, New York: Springer

- Stock, A. & Hoffmann, J. (2001). Intentional fixation of behavioral learning or how R-O learning blocks S-R-learning. In: *European Journal of Cognitive Psychology*, (in press).
- Storm, C. (1980). The semantic structure of animal terms: A developmental study. *International Journal of Behavioral Development*. 3, 381-407
- Sträter, O. (1997). *Beurteilung der menschlichen Zuverlässigkeit auf der Basis von Betriebserfahrungen*. Köln: GRS-138
- Strube, G. (1984). *Assoziation: der Prozess des Erinnerns und die Struktur des Gedächtnisses*. Berlin, Heidelberg, New York: Springer
- Tetzlaff, L. & Mack, R.L. (1991). Perspectives on Methodology in HCI Research and Practice. In: Carroll F. M. (1991) *Designing Interaction – Psychology at the Human-Computer Interface*. Cambridge, New York: Cambridge University Press
- Velickovskij, B. M. (1988). *Wissen und Handeln – Kognitive Psychologie aus tätigkeitstheoretischer Sicht*. Berlin: Deutscher Verlag der Wissenschaften
- VDI-Richtlinien zur Konstruktionsmethodik (Juli 1996). *Methodisches Entwickeln von Lösungsprinzipien*. VDI 2222
- VDI Richtlinie 2424 Bl. 1-3 (1988). *Industrial Design: Grundlagen, Begriffe, Wirkungsweisen*. Düsseldorf: VDI/VDE Verlag
- Virzi, R.A. (1992). Refining the test phase of usability evaluation: How many subjects is enough? In: *Human Factors*. 34(4), S. 457-468
- Virzi, R.A. (1990). Streamlining the Design Process: Running Fewer Subjects. In: *Proceedings of the Human Factors Society 34th Annual Meeting*
- Warren, D.H. (1994). Self-localization on plan and oblique maps; *Environment and Behaviour*, 26, S. 71-98
- Wecker, B. (1994). *Kognitive Schemata in der räumlichen Orientierung am Beispiel von Flughäfen*. Regensburg: unveröffentl. Diplomarbeit
- Weisman, J. (1981). Evaluation architectural legibility – Wayfinding in the built environment, *Environment and Behaviour*. 13, S. 189-204
- Wessels, M.G. (1994). *Kognitive Psychologie*. München: Ernst Reinhardt
- Wilkening, F. (1988). *Zur Rolle des Wissens in der Wahrnehmung*. München: UTB
- Witte, E.H. (1991). Die Idee der Anwendbarkeit und die Theorie der Erkenntnisinteressen. In: Witte, E.H. (Hg.), *Angewandte Sozialpsychologie: Das Verhältnis von Theorie und Forschung*. Braunschweig: Braunschweiger Studien zur Erziehungs- und Sozialarbeitswissenschaft
- Zimmer, A. (1973). Die Bildung von Ordnungen im Gedächtnis; Münster, Dissertation Westfälische Wilhelms-Universität
- Zimmer, H.D. (1984). Blume oder Rose? *Archiv für Psychologie*. 136, S. 343-351 Bouvier Verlag