

Jochen Müsseler / Wolfgang Prinz (Hrsg.)
Sabine Maasen (Koordination)
Lehrbuch Allgemeine Psychologie
Heidelberg: Spektrum Akademischer Verlag

Kapitel 3b: Kategorisierung und Wissenserwerb

Michael R. Waldmann

Anzahl der Anschläge (mit Leerzeichen): ca. 230 000

Anschrift des Autors:

Georg-Elias-Müller-Institut für Psychologie
Universität Göttingen
Gosslerstr. 14
37073 Göttingen
Germany

E-mail: michael.waldmann@bio.uni-goettingen.de
Tel.: (0551) 393784
Fax: (0551) 393656

KAPITEL 3B: KATEGORISIERUNG UND WISSENSERWERB

Michael Waldmann

INHALT

Textbox

3b-1 EINLEITUNG

3b-1.1 Funktionen von Kategorien

3b-1.2 Konzepte und Misskonzeptionen

3b-2 DIE MENTALE REPRÄSENTATION NATÜRLICHER KATEGORIEN

3b-2.1 Ähnlichkeitsbasierte Ansätze

3b-2.1.1 Die klassische Sicht

Probleme der klassischen Sicht

3b-2.1.2 Die Prototypen-Sicht

Probleme der Prototypen-Sicht

3b-2.1.3 Die Exemplar-Sicht

*Textbox: Beispiel für eine Exemplar-Theorie: Das Generalisierte
Kontext-Modell(GCM)*

Probleme der Exemplar-Sicht

3b-2.1.4 Theorie der Entscheidungsgrenzen

Probleme der Theorie der Entscheidungsgrenzen

Textbox: Die richtige Theorie möge sich bitte erheben

3b-2.1.5 Theorien multipler Systeme

Regeln und Prototypen

Prototypen und Exemplare

Regeln und Exemplare

Explizite und implizite Regeln

Neurowissenschaftliche Evidenz für multiple Systeme

3b-2.2 Kritik ähnlichkeitsbasierter Theorien

3b-2.3 Die Theorien-Sicht

3b-2.3.1 Probleme der Theorien-Sicht

3b-3 ARTEN VON KATEGORIEN

3b-3.1 Natürliche Arten vs. Artefakte

3b-3.2 Kausale Kategorien

3b-3.3 Andere Arten von Kategorien

Textbox: Domänenspezifität von Kategorien:

Entwicklungspsychologische Evidenz

Textbox: Kategorien-spezifische Defizite

3b-4 RELATIONEN ZWISCHEN KATEGORIEN

3b-4.1 Taxonomien

3b-4.2 Nicht-hierarchische Kategorienstrukturen

3b-5 DER ERWERB VON KATEGORIEN

3b-5.1 Konnektionistische Modelle

3b-5.2 Wissensbasierte Lerntheorien

3b-6 DIE NUTZUNG VON KATEGORIEN

3b-6.1 Der Einfluss von Zielen und pragmatischen Kontexten

3b-6.2 Konzeptuelle Kombination

3b-6.3 Sprache und Kategorien

3b-6.4 Kategorien und Induktion

3b-7 AUSBLICK

3b-8 KERNSÄTZE, SCHLÜSSELBEGRIFFE, WEITERFÜHRENDE

LITERATURANGABEN, EINSCHLÄGIGE ZEITSCHRIFTEN, LITERATUR

((Anfang Textbox))

Kategorisierung gehört zu unseren basalen kognitiven Kompetenzen. Kategorien erlauben es uns beispielsweise, ein Tier, dem wir im Wald begegnen, als Hund oder als Wolf einzuschätzen, was uns in die Lage versetzt, früheres Wissen über diese beiden Tierklassen aktuell zu nutzen. Je nachdem, ob wir das Tier als Hund oder als Wolf kategorisieren, werden wir zu unterschiedlichen Vorhersagen über das Verhalten des Tieres gelangen und uns dementsprechend unterschiedlich verhalten. Ohne Kategorien würde uns jedes Objekt einzigartig erscheinen; wir wären nicht in der Lage, früheres Wissen für Vorhersagen oder für unsere Handlungsplanung einzusetzen. Kategorien fassen Objekte oder Ereignisse auf der Basis von Gemeinsamkeiten zu Klassen zusammen, was uns erlaubt, Wissen, das wir über bereits vertraute Exemplare der Kategorien angehäuft haben, für neue, aktuelle Erfahrungen zu nutzen. Kategorien spielen eine zentrale Rolle in der Wahrnehmung, der Handlungsplanung, dem Lernen, der Kommunikation und dem Denken. Das Kapitel gibt einen Einblick in die aktuelle Forschung zur Kategorisierung. Es wird sich zeigen, dass Kategorien sich nur als Produkt der Interaktion von wahrgenommenen Ähnlichkeitsbeziehungen in der Welt, Vorwissen und den Kontexten der Nutzung verstehen lassen. Neben den klassischen verhaltensbezogenen, experimentellen Studien werden auch Computermodelle des Kategorisierens und neurowissenschaftliche und anthropologische Untersuchungen diskutiert werden. **((Ende Textbox))**

((1)) 3b-1 EINLEITUNG

Wir benutzen Kategorien, um Gruppen von Objekten oder Ereignissen wie etwa Hunde, Vögel, Junggesellen, Autos, Computer, Geburtstage, Kriege, aufgrund ihrer Gemeinsamkeiten zusammenzufassen. Kategorien können konkrete Objekte wie etwa Pflanzen bezeichnen, sie können sich aber auch auf abstrakte Gebilde wie zum Beispiel Demokratie beziehen. Wir können auch Kategorien kombinieren, um so zu einer schier endlosen Zahl neuer Kategorien zu kommen wie etwa Tennisbälle, Autoreparaturen, Haustiere oder Internetkaufhäuser. Wozu benötigen wir aber Kategorien?

((2)) 3b-1.1 Funktionen von Kategorien

Eine der wichtigsten Funktionen von Kategorien besteht darin, dass sie uns ermöglichen, neue Erfahrungen mit bereits vorhandenem Wissen in Kontakt zu bringen. Ohne Kategorien wären wir nicht in der Lage, uns in der Welt zu orientieren und von Erfahrungen zu profitieren. Wäre jedes Ereignis einzigartig, wüssten wir nicht, was wir tun sollten und welche Vorhersagen wir machen könnten. Erst die Zuordnung zu einer Klasse ähnlicher Objekte oder Ereignisse erlaubt es uns, angemessen zu reagieren. Eine Maus, die ein Tier als Katze kategorisiert, wird sich anders verhalten, als wenn sie es einer anderen, weniger gefährlichen Tierkategorie zuordnet. Ein Arzt, der einen Patienten untersucht, wird in Abhängigkeit davon, welcher Krankheitskategorie er diesen Patienten zuordnet, unterschiedliche Verlaufsprognosen machen und verschiedene Therapiestrategien initiieren.

Lange Zeit wurde *Klassifikation*, also die Zuordnung von Objekten und Ereignissen zu mentalen Repräsentationen von Kategorien, als die wichtigste, wenn nicht die einzige Funktion von Kategorien angesehen (vgl. Medin & Heit, 1999; Medin, Ross & Markman, 2001). Wie die genannten Beispiele von Kategoriennutzung allerdings deutlich machen, leisten Kategorien wesentlich mehr als bloße Klassenbildung.

Kategorisierung erlaubt es, bereits vorhandenes Wissen auf neue Erfahrungen anzuwenden. Klassifiziert man ein wahrgenommenes Objekt beispielsweise als Telefon, dann versetzt uns dies in die Lage, die Bestandteile dieses Objekts zu verstehen und

angemessen mit ihm zu interagieren. Kategorien sind also auch die Grundlage von *Verstehen*.

Eine weitere wichtige Funktion von Kategorien ist ihre Rolle, die sie beim *Lernen* spielen. Neue Erfahrungen werden auf bereits vorhandenes Wissen bezogen, sie dienen aber auch dazu, die Wissensbasis zu modifizieren, die der Kategorisierung zugrunde liegt. Kategorien werden nicht nur auf Erfahrungen angewendet, sie werden aufgrund von Erfahrungen auch erworben oder modifiziert.

Eine vierte Funktion von Kategorien betrifft die Unterstützung von *Inferenzen*. In Abhängigkeit davon, ob wir ein Tier als Hund oder als Wolf kategorisieren, werden wir unterschiedliche Erwartungen über das Verhalten des Tieres haben. Ein anderes Beispiel sind Krankheitsdiagnosen, die unterschiedliche Erwartungen über die Wirksamkeit von Therapien erzeugen.

Kategorien spielen auch in *Erklärungen* eine wichtige Rolle. Die Kategorisierung eines Symptommusters als AIDS vermittelt gleichzeitig Wissen über die Ursachen der beobachteten Merkmale und über weitere kausale Effekte dieser Erkrankung, die in der Zukunft zu erwarten sind.

Auch unser *Denken* beruht in hohem Maße auf Kategorien. Sowohl deduktive als auch induktive Schlüsse beziehen sich auf Kategorien von Objekten und Ereignissen. Bei deduktiven Schlüssen inferieren wir von Eigenschaften einer Kategorie auf einen Einzelfall (z.B. „Alle Raben sind schwarz. Also ist der weiße Vogel, den du gefunden hast, kein Rabe.“), während induktive Schlüsse hypothetische Folgerungen von Einzelfällen auf die vermutlich zugrunde liegende Kategorie beinhalten (z.B. „Dieses Säugetier hat eine Milz. Also haben vermutlich alle Säugetiere eine Milz.“).

Kategorien sind auch bei *Planungen* und Prozessen der *Handlungssteuerung* beteiligt. Planen wir etwa eine Campingreise, müssen wir uns über die Dinge Gedanken machen, die wir bei einer solchen Reise mitnehmen sollten (Barsalou, 1983). Analog werden wir

bei der Handlungsvorbereitung Kategorien aktivieren, die die Ziele zusammenfassen, die man mit den zur Wahl stehenden Handlungen erreichen kann.

Schließlich spielen Kategorien auch eine zentrale Rolle in der *Kommunikation* und in der *Sprache*. Sprache erlaubt es uns, unser über gemeinsame Kategorien repräsentiertes Wissen zu kommunizieren. Sprache ermöglicht es aber auch, vorhandenes kategoriales Wissen durch indirektes Lernen zu modifizieren. Der Hinweis eines Freundes, dass man sich in der Gegenwart von Klapperschlangen besser nicht bewegt, ermöglicht es uns, unser Verhalten zu verändern, auch wenn wir zuvor nie einer Klapperschlange begegnet sind.

In manchen Texten wird zwischen Kategorien, die sich auf Klassen in der Welt beziehen, und Konzepten, die die mentale Repräsentation dieser Klassen bezeichnen, unterschieden (z.B. Smith & Medin, 1981). In diesem Kapitel wird hingegen zwischen diesen beiden Begriffen nicht differenziert, da sich zunehmend die Sicht durchsetzt, dass Kategorien Produkte der Interaktion von Strukturen in der Welt und informationsverarbeitenden Systemen sind (vgl. Abschnitt 6).

((2)) 3b-1.2 Konzepte und Misskonzeptionen

Obgleich Kategorien zentral sind für menschliche Informationsverarbeitung, können sie auch zu Problemen führen. Es ist möglich, dass wir falsche oder zu umfassende Kategorien bilden und damit die Ähnlichkeit bzw. Verschiedenheit von Dingen oder Personen übertreiben (vgl. Medin et al., 2001).

Abb. 3b-1

Ein klassisches Beispiel für die Auswirkungen von Kategorien stammt von Tajfel und Wilkes (1963). In einem Experiment haben sie Versuchsteilnehmern vier kürzere Linien, die sie als „A“, und vier längere Linien, die sie als „B“ bezeichnet haben, gezeigt (vgl. Abb. 3b-1). Im Vergleich zu einer Kontrollgruppe, die die acht Linien ohne die Kategorienbezeichnungen A und B erhielten, tendierte die Experimentalgruppe dazu, die Linien innerhalb der beiden Kategorien A und B als ähnlicher in ihrer Länge

wahrzunehmen als die Kontrollgruppe. Sie übertrieben also die objektiven Unterschiede der Kategorienmitglieder.

Diese Tendenz kann von großer Bedeutung im Bereich sozialer Kategorienbildung sein (vgl. auch Devine, 1989). In einer neueren Studie hat Hirschfeld (1994, 1996) die Entwicklung der Kategorie der Rasse untersucht. Er hat einer jüngeren (7- und 8-Jährige) und einer älteren Gruppe (11- und 12-Jährige) von Kindern aus der Mittelschicht in den USA Bilder von vier Paaren gezeigt, bei denen die Partner entweder beide schwarz oder beide weiß oder gemischt waren. Die Kinder wuchsen in einer Umgebung auf, in der es wenige Schwarze gab. Jedes Kind sollte dann aus einer Reihe von Bildern von Babys dasjenige heraussuchen, das vermutlich zu dem jeweiligen Paar gehörte. Alle Kinder wählten das schwarze Baby für das schwarze Paar und das weiße Baby für das weiße Paar. Interessant waren nun aber die Befunde zu dem gemischten Paar. Während die jüngeren Kinder keine klare Präferenz zeigten, tendierten die älteren Kinder stark dazu, das schwarze Baby diesem Paar zuzuschreiben. In einer Kontrollbedingung mit Tieren mit verschiedener Haar- und Hautfarbe machten sie diesen Fehler nicht. Die Befunde sprechen also dafür, dass die Jugendlichen das in ihrer Kultur übliche Rassenstereotyp übernommen haben, demnach Kinder von gemischten Paaren als Schwarze angesehen werden. Hirschfeld befragte auch Großstadtkinder aus einer Schule mit Schwarzen und Weißen. Bei diesen Kindern wurde bevorzugt die korrekte Wahl getroffen. Die Befunde sind also abhängig vom kulturellen Umfeld.

Ogleich diese Beispiele eher die negativen Seiten von Kategorien betonen, sollte klar sein, dass Kategorien eine unabdingbare kognitive Funktion haben, ohne die Informationsverarbeitung unmöglich wäre. Manche Studien haben ein optimistischeres Bild gezeichnet, indem sie zeigten, dass wir Stereotypen (z.B. des Geschlechts) besonders stark verwenden, wenn wir über keine anderen, spezifischeren Informationen verfügen (vgl. Locksley, Borgida, Brekke & Hepburn, 1980). In den letzten Jahren gibt es ein wachsendes Interesse von Kognitions- und Sozialpsychologen an der Frage, wann wir Informationen auf der Ebene von Individuen im Unterschied zu Gruppen aktivieren (z.B. Barsalou, Huttenlocher & Lamberts, 1998; Fiske, Neuberg, Beattie & Milberg, 1987).

Zusammenfassend dienen Kategorien einer Vielzahl von ganz unterschiedlichen Funktionen, die in mannigfacher Weise interagieren können. Kategorien sind nicht nur die Basis von Klassifikation, sie spielen auch eine zentrale Rolle in einer Reihe wichtiger kognitiver Prozesse. Neben ihren vielfältigen adaptiven Funktionen können Kategorien allerdings auch zu Verzerrungen der Übereinstimmungen und Unterschiede der von ihnen zusammengefassten Objekte führen.

((1)) 3b-2 DIE MENTALE REPRÄSENTATION NATÜRLICHER KATEGORIEN

Das zahlenmäßig dominierende Forschungsfeld im Bereich der Kategorisierungsforschung beschäftigt sich mit der Frage, wie Kategorien kognitiv repräsentiert sind und wie neue Objekte bereits vorhandenen Kategorien zugeordnet werden (vgl. auch Eckes, 1991; Medin et al., 2001; Heydemann, 1998; Hoffmann, 1986, 1993; Komatsu, 1992; Smith & Medin, 1981; Sternberg & Ben-Zeev, 2001, für Überblicke). Dabei ging man, ähnlich wie in anderen Gebieten der Kognitionspsychologie, lange Zeit davon aus, dass die semantischen Inhalte kognitiver Repräsentationen, also die Inhalte von Kategorien, psychologisch weniger wichtig sind als die Struktur der Repräsentationsmedien. Hinter dieser Forschungsstrategie verbirgt sich die Annahme, dass es allgemeine Repräsentationsstrukturen und –prozesse gibt, die unabhängig von den Inhalten sind, die sie vermitteln.

((2)) 3b-2.1 Ähnlichkeitsbasierte Ansätze

Wie ordnen wir nun konkrete Objekte oder Ereignisse Kategorien zu? Die Antwort einer großen Zahl ansonsten konkurrierender Ansätze klingt einfach. Wir fassen diejenigen Objekte oder Ereignisse zu einer Kategorie zusammen, die sich ähnlich sind. Dieses Verfahren erzeugt Klassen, bei denen die Ähnlichkeit innerhalb der Kategoriengrenzen maximiert und die Ähnlichkeit zwischen Kategorien minimiert wird (Rosch, 1978). Eine für die Kategorienforschung hochrelevante Frage ist nun, wie Ähnlichkeit kognitiv bestimmt wird. Auch wenn das Konzept der Ähnlichkeit den Kern vieler Theorien zur Kategorisierung darstellt, gibt es dennoch ganz unterschiedliche Sichtweisen darüber, wie Kategorien repräsentiert werden und wie Ähnlichkeit bestimmt wird.

((3)) 3b-2.1.1 Die klassische Sicht

Die klassische Sicht, die bis in die 60er Jahre des letzten Jahrhunderts dominierte, geht davon aus, dass Konzepte durch *definierende* Merkmale ausgezeichnet sind, die die notwendigen und hinreichenden Bedingungen der Kategorienzugehörigkeit spezifizieren (vgl. z.B. Bourne, 1970; Bruner, Goodnow & Austin, 1956; Hull, 1920). Ein Dreieck lässt sich beispielsweise als geschlossene geometrische Form mit drei Seiten, die drei Winkel umschließen, deren Summe 180 Grad beträgt, definieren. Gemäß dieser Sicht legen Konzepte auf der Basis solcher Definitionen fest, ob ein Gegenstand zu dem Konzept gehört oder nicht. Alle Mitglieder sind gleich gute Exemplare für die jeweilige Kategorie und Kategorienlernen besteht darin, die definitorischen Merkmale der Kategorie zu entdecken. Diese Lernprozesse wurden häufig mit künstlichem Versuchsmaterial (z.B. geometrischen Figuren) untersucht, da dies ein höheres Maß der experimentellen Kontrolle ermöglichte (z.B. Bruner et al., 1956). Abbildung 3b-2 zeigt ein einfaches Beispiel für Kategorien, die sich auf der Basis einer definitorischen Regel trennen lassen. Objekte mit dunklem Hintergrund gehören zu Kategorie A (oben) und solchen mit hellem Hintergrund zu Kategorie B (unten). Regeln können auch komplizierter sein. Eine relativ schwierig zu erlernende Regel könnte etwa sein, dass Objekte in einer Kategorie entweder rund oder dunkel sind, aber nicht beides (disjunktive Regel).

Abb. 3b-2

In der Forschung, die im Rahmen der klassischen Sicht durchgeführt wurde, wurde eine Vielzahl solcher Regeln untersucht. Darüber wurde aber lange Zeit versäumt, die Frage zu stellen, ob natürliche, im Alltag gebräuchliche Kategorien (z.B. Lebewesen, Fahrzeuge, Personentypen usw.) tatsächlich auf der Basis von definitorischen Merkmalen repräsentiert werden. Untersuchungen, die sich dieser Frage widmeten, kamen nämlich nahezu einhellig zu dem Ergebnis, dass es unplausibel ist anzunehmen, dass Konzepte gemäß der klassischen Sicht repräsentiert werden.

((4)) *Probleme der klassischen Sicht*. So konnten etwa McCloskey und Glucksberg (1978) zeigen, dass eine große Uneinigkeit bei ihren Probanden über die Frage bestand, ob bestimmte Objekte einer bestimmten Kategorie angehören oder nicht. Dieser Befund legt nahe, dass die Kategoriengrenzen eher unscharf sind. Ebenso problematisch für die klassische Sicht ist der Sachverhalt, dass es große interindividuelle Variabilität bei der Festlegung der für bestimmte Kategorien relevanten Merkmale gibt und dass viele der genannten Merkmale nicht notwendig und hinreichend für die Kategorien waren (vgl. Barsalou, 1989; Hampton, 1979).

Ein weiteres Problem für die klassische Sicht waren Befunde, die zeigten, dass Kategorien nicht diskret sind, sondern dass verschiedene Exemplare der Kategorie als mehr oder weniger stark der Kategorie zugehörig empfunden werden. Mit einer Reihe von Methoden konnte so gezeigt werden, dass es Unterschiede im Ausmaß der *Typikalität* von Exemplaren gibt. So werden beispielsweise ein Tisch oder ein Sofa als typischeres Möbelstück empfunden als ein Teppich oder Telefon. Ein Rotkehlchen gilt als typischerer Vogel als ein Huhn und so weiter. Unterschiedliche Typikalität kann nicht nur mit Hilfe verbaler Einschätzungen auf Rating-Skalen gemessen werden, sie drückt sich auch in Reaktionszeitmaßen aus (z.B. die gemessene Latenz der Antwort auf Satzverifikationsfragen wie „Ist ein Huhn ein Vogel?“ oder Priming-Verfahren) oder in der Reihenfolge der Nennung von Kategorienexemplaren (vgl. z.B. Barsalou & Sewell, 1985; Hampton, 1979; McCloskey & Glucksberg, 1979; Rosch, Simpson & Miller, 1976). Typische Exemplare werden generell schneller beurteilt und früher genannt als atypische (vgl. zu Priming-Verfahren auch das Kapitel von Zwitserlood und Bölte in diesem Band) .

Unter Beschuss kam die klassische Sicht aber auch aufgrund philosophischer Analysen von Konzepten (vgl. Fodor, 1998; Quine, 1960; Putnam, 1975). So ist fragwürdig, ob Konzepte, die *natürliche Arten* (z.B. Lebewesen, Pflanzen) bezeichnen, definatorische Merkmale haben, die Bestandteil unserer Repräsentation sind. In der Regel verwenden wir Kategorien im Alltag, ohne die genaue wissenschaftliche Basis für die Kategorisierung zu kennen (man denke etwa an Fische oder Pflanzen). Und selbst Experten sind sich darüber im Klaren, dass es möglich ist, dass künftige Forschung zu

einer Veränderung des Wissens über einzelne natürliche Arten und damit auch zu veränderten Kriterien für die Kategorisierung führen kann. Man geht hier also eher davon aus, dass Konzepte theorienähnliche Gebilde sind, die Hypothesen und Annahmen ausdrücken, ohne sich ein für alle Mal auf bestimmte definitorische Relationen festzulegen. Dies gilt in besonderem Maße auch für wissenschaftliche Konzepte (z.B. „Spezies“), die in der Regel nicht definitorisch festgelegt werden, sondern in Abhängigkeit vom gerade aktuellen wissenschaftlichen Paradigma und von wissenschaftlichen Entdeckungen sukzessive präzisiert und verändert werden (vgl. Sokal, 1974; Kuhl & Waldmann, 1985). Selbst ein scheinbar problemlos zu definierendes Konzept wie „Junggeselle“ zeigt, dass es gar nicht so leicht ist, alle definitorischen Merkmale zu nennen. Schlägt man als Definition beispielsweise „unverheirateter, erwachsener Mann“ vor, dann müsste auch der Papst als Junggeselle bezeichnet werden.

((3)) 3b-2.1.2 Die Prototypen-Sicht

Eine Alternative zur klassischen Sicht bieten Theorien, die von Smith und Medin (1981) unter dem Oberbegriff „probabilistische Sicht“ oder „Prototypen-Sicht“ zusammengefasst wurden. Gemäß dieser Sicht werden Kategorien als Bündel von Merkmalen repräsentiert, die typisch oder charakteristisch für die Kategorie sind, aber nicht unbedingt notwendig oder hinreichend. So mag die Kategorie „Vogel“ bei den meisten Menschen Merkmale wie Nest bauen oder fliegen beinhalten, auch wenn nicht alle Vögel all diese Merkmale haben (z.B. Pinguine, Straußenvögel).

Abb 3b-3 ODER IST DAS EINE TABLE?

Die Grundidee des Prototypen-Ansatzes soll anhand eines konkreten Beispiels skizziert werden. Abbildung 3b-3 listet die Symptome von drei Personen auf, die an einem grippalen Infekt erkrankt sind. Wie man sieht, gibt es kein einzelnes Merkmal, das von allen Personen geteilt wird. Es ist also nicht möglich, definitorische Merkmale zu finden, die die Kategorie grippaler Infekt beschreiben. Es gibt aber eine Reihe charakteristischer Merkmale, an der die Mehrzahl der Patienten leiden. Diese charakteristischen Merkmale sind Bestandteil des Prototyps (rechts). Der Prototyp fasst

die typischen Merkmale einer Kategorie zusammen, ohne dass es ein Exemplar geben muss, das ihm vollständig entspricht. Gemäß der Prototypen-Sicht werden beim Erwerb der Kategorie die charakteristischen Merkmale der Exemplare registriert. Daraus wird der Prototyp gebildet, der die typischen Merkmale der Kategorie zusammenfasst. Neue Exemplare werden durch einen Ähnlichkeitsvergleich mit den Prototypen der einzelnen Kategorien klassifiziert. Die Symptome eines neuen Patienten würde man dieser Sicht zufolge mit den Prototypen unterschiedlicher Krankheiten vergleichen und diejenige Krankheit diagnostizieren, zu dessen Prototyp die größte Ähnlichkeit vorliegt.

Prototypen-Strukturen wurden nicht nur mit verbalem, sondern auch mit visuellem Material untersucht. Abbildung 3b-4 zeigt eine beliebte Aufgabe mit abstrakten Stimuli, die von Posner und Keele (1968) entwickelt wurde. Dabei wird eine beliebige Anordnung von Punkten als Prototyp ausgewählt und die Lernitems werden dadurch erzeugt, dass die Positionen nach Zufall räumlich leicht verschoben werden (Kategorie A). Andere Stimuli fungieren dann als Gegenkategorie B. Wie man sieht, sind sich die Punktwolken innerhalb von Kategorie A untereinander ähnlicher als sie zu den Punktwolken der Kategorie B sind.

ABB. 3B-4

Die Aufnahme charakteristischer Merkmale erklärt auch, warum manche Exemplare als typischer empfunden werden als andere und warum Kategoriengrenzen unscharf sein können. Gründe dafür könnten sein, dass atypische Vertreter einer Kategorie (z.B. Straußenvögel) weniger für Vögel charakteristische Merkmale aufweisen als typische (z.B. Rotkehlchen) oder dass nicht zu einer bestimmten Kategorie (z.B. Fische) gehörige Exemplare (z.B. Wale) dennoch viele Merkmale mit dieser Kategorie teilen. Insbesondere Rosch und Mervis (1975) sind dieser Frage nachgegangen; sie konnten zeigen, dass die eingeschätzte *Typikalität* eines Exemplars exzellent durch die Anzahl der Merkmale vorhergesagt werden kann, die es mit anderen Mitgliedern der Kategorie teilt.

Gemäß der klassischen Sicht werden Kategorien als Definitionen repräsentiert. Wie kann man sich nun die Repräsentation gemäß der Prototypen-Sicht vorstellen? Eine zentrale Idee in diesem Bereich geht auf das vom Philosophen Wittgenstein (1953) entwickelte Konzept der *Familienähnlichkeit* zurück. Kategorien erlangen ihre Kohärenz dadurch, dass jedes Exemplar ein oder mehrere Merkmale mit anderen Exemplaren teilt, es aber kein einzelnes Merkmal oder Merkmalsbündel gibt, das von allen Exemplaren geteilt wird. Wittgenstein hat dies am Beispiel der Kategorie Spiele verdeutlicht, die eine Familienähnlichkeit vereint, ohne dass es Merkmale gibt, die alle Spiele besitzen.

Abbildung 3b-5 zeigt ein weiteres Beispiel für eine Prototypenorganisation von visuellen Stimuli. Auf der linken Seite sieht man vier Gesichter, die alle voneinander verschieden sind, aber aufgrund der Überlappung mehrerer Merkmale eine Familienähnlichkeit aufweisen. Obgleich drei von vier Familienmitgliedern jeweils ein Merkmal (z.B. Form des Mundes) teilen, gibt es kein einzelnes Merkmal, das alle haben. Es gibt also kein notwendiges und hinreichendes Merkmal für diese Familie im Sinne der klassischen Sicht.

ABB. 3B-5

Eine Möglichkeit, wie man sich gemäß der Prototypen-Sicht die mentale Repräsentation von Kategorien vorstellen kann, besteht darin, dass beim Lernen die charakteristischen Merkmale einer Kategorie abstrahiert werden und zu einer Repräsentation des *Prototypen* der Kategorie zusammengeführt werden. Der Prototyp beinhaltet also alle typischen Merkmale, wobei es sein kann, dass kein konkretes Exemplar diesem Prototyp vollkommen entspricht. Der Prototyp für die Gesichter in Abbildung 3b-5 wäre etwa das Gesicht auf der rechten Seite, das die charakteristische, modale Ausprägung aller Merkmale aufweist. Bei kontinuierlich variierenden Merkmalsdimensionen (z.B. Körpergröße) würde der Prototyp die mittlere Ausprägung dieser Dimension zeigen. Der Prototyp ist also eine Verkörperung der zentralen Tendenz einer Menge von sich ähnelnden Exemplaren. Es ist bei anderen Anordnungen von Exemplaren auch möglich, die einzelnen Merkmale gemäß ihrer Relevanz für die

Kategorie zusätzlich zu gewichten. Häufig werden manche Merkmale von mehr Kategorienmitgliedern geteilt als andere und sollten deshalb auch ein höheres Gewicht erhalten (vgl. Smith & Medin, 1981).

Kategorienrepräsentationen beruhen dieser Sicht zufolge also auf *Abstraktionen* über eine Reihe von konkreten Exemplaren. Kategorisierung wird dann als Ähnlichkeitsvergleich zwischen dem zu kategorisierenden Exemplar und der Prototypenrepräsentation vollzogen. Sieht man in einem Zoo beispielsweise ein neues Tier, dann könnte man die Merkmalsüberlappung zwischen diesem Tier und den Prototypenrepräsentationen für verschiedene Tiere (z.B. Vögel, Fische) vergleichen. Wird dabei ein bestimmter Schwellwert oder ein Kriterium für eine der Kategorien überschritten, dann wird das Tier dieser Kategorie zugeordnet.

Prototypisch organisierte Kategorien wurden häufig mit künstlich konstruiertem Material untersucht, bei dem sichergestellt werden kann, dass die Kategorie eine Familienähnlichkeitsstruktur hat (vgl. z.B. das in Abb. 3b-4 dargestellte Material). Viele dieser Studien haben Befunde zu Tage gefördert, die konsistent mit der Prototypen-Sicht sind. So hat sich gezeigt, dass typische Exemplare schneller gelernt und kategorisiert werden als atypische (Rosch & Mervis, 1975). Ein wichtiger Befund betrifft Transfertests, in denen nach Abschluss der Lernphase neue Exemplare zum Kategorisieren vorgelegt werden. Häufig wird hier gefunden, dass neue, bisher gar nicht beobachtete Exemplare, die dem Prototypen entsprechen, schneller und fehlerfreier der Kategorie zugewiesen und länger im Gedächtnis behalten werden als manche weniger typischen Exemplare, die in der Lernphase bereits gezeigt wurden (z.B. Homa & Vosburgh, 1976; Medin & Schaffer, 1978; Posner & Keele, 1968, 1970).

((4)) *Probleme der Prototypen-Sicht.* Ähnlich wie bei der klassischen Sicht hat der Fokus auf artifizielles Labormaterial dazu geführt, dass viele Eigenschaften von natürlichen Kategorien vernachlässigt wurden. Ein Grundproblem der Prototypensicht ist etwa, dass nur die zentrale Tendenz einer Kategorie gespeichert wird, nicht aber Informationen über die Variabilität der Kategorienexemplare oder die relative Größe der einzelnen Kategorien. Es konnte aber gezeigt werden, dass diese Informationen

berücksichtigt und bei der Kategorisierung genutzt werden (vgl. Estes, 1986; Flannagan, Fried & Holyoak, 1986; Fried & Holyoak, 1984; Medin & Schaffer, 1978).

Ein anderer Aspekt, der von der Prototypensicht vernachlässigt wird, sind Merkmalskorrelationen. So wissen wir beispielsweise, dass kleine Vögel eher singen als große oder dass Blutdruck und Herzerkrankung zusammenhängen und sind durchaus in der Lage, dieses Wissen in der Lernphase zu erwerben und bei der Kategorisierung zu nutzen (Malt & Smith, 1983; Medin, Altom, Edelson & Freko, 1982).

Kontexteffekte stellen ein weiteres Problem für die Prototypensicht dar. Wie typisch ein Exemplar für eine Kategorie ist, hängt auch vom Kontext ab, in dem es präsentiert wird. So ist Kaffee zwar das sowohl für Fernfahrer als auch Sekretärinnen typischste Getränk, danach folgt aber Tee bei Sekretärinnen, während bei Fernfahrern Milch an zweiter Stelle kommt (Roth & Shoben, 1983). Kategorienrepräsentationen, die von einer einheitlichen, am Prototypen orientierten Ähnlichkeitsstruktur ausgehen, können diese Effekte nicht erklären.

Schließlich macht die Prototypensicht Vorhersagen über die Lernschwierigkeit unterschiedlicher Kategorienstrukturen, die sich als problematisch erwiesen haben. Gemäß dieser Sicht beinhaltet der Kategorisierungsprozess eine additive Verknüpfung der gewichteten Merkmale. In der in Abbildung 3b-5 gezeigten Kategorienstruktur beispielsweise gehört ein Exemplar dann zu der Kategorie der Familie, wenn es mindestens drei der vier für die Kategorie relevanten, charakteristischen Merkmalsausprägungen (Nase, Kopf, Mund, Augen) aufweist, wobei gleichgültig ist, welche drei typischen Merkmale vorliegen. Die Merkmale werden also aufaddiert und die Summe mit der Kriteriumsregel („mindestens drei der vier Merkmale“) verglichen. Kategorien, die nach einem solchen Konstruktionsprinzip strukturiert sind, nennt man linear trennbar.

Abb. 3b-6

Abbildung 3b-6 gibt eine graphische Veranschaulichung der linearen Trennbarkeit mit Hilfe von zwei kontinuierlichen Merkmalen, die als Charakteristika von Lebensmitteln (links) oder Tieren (rechts) gesehen werden können. Geht man jeweils davon aus, dass jedes Exemplar nur mit den zwei in Abbildung 3b-3 genannten Merkmalen beschrieben wird, dann zeigt sich bei den Tieren, dass die Kategoriengrenze mit Hilfe einer Geraden festgelegt werden kann, die zwischen den Kategorien Insekten und Säugetiere trennt. Dies ist ein Beispiel für eine *linear trennbare* Kategorie. Hier ist es möglich, eine Gleichung zu formulieren, die die Kategoriengrenze (die Gerade) als lineare Funktion der beiden gewichteten Merkmale erzeugt. Bei Kategorien mit drei Merkmalen würde lineare Trennbarkeit bedeuten, dass sich eine lineare Gleichung finden lässt, die eine Fläche im dreidimensionalen Raum erzeugt, die die kontrastierenden Kategorien trennt. Dieses Prinzip lässt sich auf einen m-dimensionalen Merkmalsraum erweitern.

Bei der auf der linken Seite abgebildeten Kategorienstruktur (Früchte vs. Gemüse) ist dies nicht möglich, hier hat man es also mit einer *nicht-linear trennbaren* Kategorie zu tun. Für diese Struktur lassen sich keine Prototypen für die beiden Kategorien finden, die die Eigenschaft haben, dass alle Exemplare der jeweiligen Kategorie dem eigenen Prototypen näher sind als dem Prototypen der Gegenkategorie. Es gibt keine lineare Gleichung, mit der es möglich wäre, die beiden Kategorien voneinander zu trennen. Viele Theorien des maschinellen Lernens teilen mit der Prototypensicht die Annahme, dass nicht-linear trennbare Kategorien gar nicht oder vergleichsweise schwer zu erlernen sein sollten (Minsky & Papert, 1969).

Diese Vorhersage wurde allerdings empirisch nicht generell bestätigt. So haben Medin und Schwanenflugel (1981) keinen Unterschied in der Lernschwierigkeit bei linear und nicht-linear trennbaren Kategorien gefunden (vgl. auch Kemler-Nelson, 1984). Es gibt aber auch Gegenbefunde, die die Allgemeinheit dieses Befunds in Frage stellen (Smith, Murray & Minda, 1997).

((3)) 3b-2.1.3 Die Exemplar-Sicht

Exemplar-Theorien verkörpern einen von der Prototypen-Sicht radikal unterschiedlichen Ansatz. Exemplar-Theorien gehen im Gegensatz zu dieser Sicht *nicht*

davon aus, dass beim Lernen eine abstrakte Repräsentation des Prototypen aufgebaut wird. Vielmehr wird angenommen, dass wir uns beim Lernen einer Kategorie die einzelnen Exemplare zusammen mit der Kategorienbezeichnung einprägen. Ein neues Exemplar wird nun dadurch kategorisiert, dass es uns an diese früher gesehenen Exemplare mehr oder weniger erinnert und dass wir annehmen, dass es vermutlich der gleichen Kategorie angehört wie diejenigen Exemplare, denen es insgesamt am ähnlichsten ist. Ähnlichkeitsvergleiche werden also mit Exemplaren, nicht mit abstrakten Prototypen vollzogen. So kann es etwa sein, dass wir ein bestimmtes Tier als Nagetier klassifizieren, weil es uns an eine Maus erinnert, während wir ein anderes Tier in die gleiche Klasse einordnen, weil es Eichhörnchen ähnlich ist. Dabei muss man sich den Prozess nicht so vorstellen, dass einzelne Exemplare nacheinander und einzeln ins deklarative Gedächtnis gerufen werden. Vielmehr gehen die meisten Theorien davon aus, dass beim Kategorisieren eine Vielzahl ähnlicher Exemplare parallel im Gedächtnis aktiviert werden.

Wie würde dieser Ansatz die Kategorisierungssituation in Abbildung 3b-3 darstellen? Anstatt einen Prototypen zu abstrahieren, würden Vertreter dieser Sicht postulieren, dass man sich die Symptommuster der unterschiedlichen Personen einzeln einprägt. Der Lernprozess besteht also darin, dass man sich die Exemplare, die man beobachtet, einzeln einprägt. Begegnet einem nun ein neues Exemplar, also beispielsweise ein neuer Patient, dann wird dieses Exemplar mit allen Exemplaren der unterschiedlichen Kategorien verglichen (also beispielsweise den drei Personen in Abb. 3b-3 für die Kategorie grippaler Infekt). Diejenige Kategorie wird am Ende gewählt, zu dessen Exemplaren die Gesamtähnlichkeit des neuen Exemplars am höchsten ist.

Die meisten in der Literatur postulierten Exemplar-Theorien gehen davon aus, dass diejenigen Exemplare, die dem zu klassifizierenden Objekt am ähnlichsten sind, den größten Einfluss auf die Kategorisierungsentscheidung haben (Brooks, 1978; Medin & Schaffer, 1978; Nosofsky, 1986). Dies wurde in einer Reihe von empirischen Untersuchungen bestätigt. Ein Beispiel aus der Medizin stammt von Brooks, Norman und Allen (1991). Sie konnten zeigen, dass Diagnosen dermatologischer Erkrankungen in hohem Maße dadurch beeinflusst werden, ob die Ärzte bereits Patienten mit

ähnlichen Symptomen gesehen haben. Die Diagnosen hingen also davon ab, welche Patienten die Ärzte zuvor gesehen hatten. Interessanterweise nahm der Einfluss spezifischer ähnlicher Exemplare mit zunehmender Erfahrung der Ärzte nicht ab.

Obgleich die Grundannahme von Exemplar-Theorien, dass wir beim Lernen keine abstrakte Kategorienrepräsentation ausbilden, zunächst wenig plausibel erscheint, hat eine große Zahl von empirischen Untersuchungen zeigen können, dass diese Theorien eine Vielzahl von Befunden, die problematisch für Prototypen-Theorien sind, erklären können (z.B. Estes, 1986; Medin & Schaffer, 1978; Nosofsky, 1988, 1991; Nosofsky & Palmeri, 1997).

Ein Vorzug von Kategorisierung im Sinne der Exemplar-Theorien besteht darin, dass dabei mehr Informationen behalten werden als bei Prototypen-Ansätzen, bei denen am Ende des Lernprozesses häufig nur noch der abstrakte Prototyp übrigbleibt (vgl. aber Fried & Holyoak, 1984; Smith & Minda, 1998, 2000). Auf diese Weise erlauben es Exemplar-Theorien, Sensibilität für die Variabilität der Lernexemplare zu erklären. Auch die Befunde zur Kontextsensitivität von Kategorisierung stellen kein Problem dar, weil, wie das Beispiel der Nagetiere gezeigt hat, der Kategorisierungsprozess in hohem Maße durch das Lernexemplar und den Kontext, in dem es sich befindet, gesteuert wird. Auch die Berücksichtigung von Merkmalskorrelationen lässt sich besser mit diesem Ansatz erklären, weil die in unterschiedlichen Exemplaren verkörperten Merkmalsbeziehungen erhalten bleiben. Schließlich sagen Exemplar-Theorien nicht generell vorher, dass linear trennbare Kategorien leichter zu lernen sind als nicht-lineare, da der Lernprozess es aufgrund der getrennten Speicherung von Einzelexemplaren nicht erfordert, eine lineare Kategoriengrenze zwischen die Kategorien zu ziehen (vgl. Nosofsky, Gluck, Palmeri, McKinley & Glauthier, 1994; Medin & Schwanenflugel, 1981).

Selbst der scheinbar für Prototypentheorien überzeugendste Befund, die leichtere Klassifizierung eines neuen, dem Prototypen entsprechenden Exemplars, lässt sich mit diesem Ansatz erklären. In der Regel sind typische Exemplare einer großen Zahl von Exemplaren der eigenen Kategorie ähnlich, während sie Exemplaren anderer Kategorien

eher unähnlich sind. Dies bewirkt eine schnelle und korrekte Klassifizierung. Im Gegensatz dazu sind atypische Exemplare an den Randbereichen der Kategorie weniger Exemplaren der eigenen Kategorie, aber dafür häufig Exemplaren anderer Kategorien ähnlich. Dies führt zu höherer Unsicherheit und verlängerter Latenzzeit bei der Kategorisierungsentscheidung (vgl. Textbox *Beispiel für eine Exemplar-Theorie: Das Generalisierte Kontext-Modell (GCM)* für eine ausführlichere Darstellung derartiger Theorien).

((Anfang Textbox))

Beispiel für eine Exemplar-Theorie: Das Generalisierte Kontext-Modell

(GCM)

Alle Exemplar-Theorien vereint die Annahme, dass die Speicherung und der Zugriff auf einzelne Exemplare zentral ist für den Kategorisierungsprozess. Ausgearbeitete Theorien beinhalten aber eine Reihe von Zusatzannahmen über die Art des Zugriffs auf den Exemplar-Speicher, die über die empirische Bewährung des Modells mit entscheiden. So gibt es eine Reihe unterschiedlicher Möglichkeiten, wie man sich den Prozess des Ähnlichkeitsvergleichs vorstellen kann, ohne dass dadurch die zentrale Eigenschaft des Modells, ein Exemplar-Modell zu sein, tangiert würde. Um eine bessere Vorstellung über die Komponenten formaler Kategorisierungstheorien zu geben, wird hier das von Nosofsky (1984, 1991) entwickelte *Generalisierte Kontext-Modell (GCM)* herausgegriffen, da es das derzeit wohl am besten empirisch bewährte Exemplar-Modell ist. GCM ist als mathematisches Modell formalisiert. Im Folgenden wird dieses Modell allerdings nur qualitativ beschrieben, eine präzisere Darstellung findet sich beispielsweise in Nosofsky und Johansen (2000).

GCM geht, wie manche andere konkurrierende Theorien auch, von einem multidimensionalen Skalierungsansatz der Modellierung von Ähnlichkeit aus. Gemäß diesem Ansatz lassen sich Exemplare in einem multidimensionalen, psychologischen Raum lokalisieren, bei dem die Achsen die Merkmalsdimensionen der Kategorien repräsentieren. In dem in Abbildung 3b-6 gezeigten Beispiel sind Tiere in einem zweidimensionalen Raum platziert, der

durch die Merkmale Größe und Intelligenz aufgespannt ist. In realistischen Modellierungen muss man natürlich viele weitere Dimensionen annehmen, die einen multidimensionalen Raum bilden. Je weiter zwei Exemplare in diesem Raum voneinander entfernt sind, desto unähnlicher sind sie sich. Eine weitere wichtige Annahme ist, dass dieser multidimensionale Raum durch *selektive Aufmerksamkeit* auf bestimmte Merkmalsdimensionen modifiziert werden kann. Dieser Annahme liegt die Beobachtung zugrunde, dass Lernende in der Lage sind, bestimmten Merkmalsdimensionen besondere Aufmerksamkeit zuzuwenden, wenn dies die Klassifikationsleistung verbessert.

Abb. 3b-7

Abbildung 3b-7A gibt ein graphisches Beispiel für den Kategorisierungsprozess. In diesem zweidimensionalen Beispiel gibt es zwei Kategorien X und O. Die Exemplare X2 und O4 sind nahe beieinander, sich also ähnlich, während die Exemplare X5 und O2 sich eher unähnlich sind, da sie relativ weit voneinander entfernt sind. Möchte nun ein Beobachter Item i klassifizieren, dann würde er oder sie gemäß GCM die Ähnlichkeit von i zu allen X-Exemplaren und getrennt davon zu allen O-Exemplaren aufsummieren. Die Klassifikationsentscheidung ist von der relativen Größe dieser Summen abhängig.

Eine kritische weitere Annahme betrifft nun die Bestimmung der Ähnlichkeit. Dabei geht GCM zunächst davon aus, dass ein Beobachter den Ähnlichkeitsraum durch selektive Aufmerksamkeit verändern kann. In Abbildung 3b-7A beispielsweise ist Objekt i in etwa gleich ähnlich zu den Exemplaren der Kategorie X und O. Man kann aber sehen, dass die horizontale Dimension wesentlich wichtiger für die Separierung der beiden Kategorien ist als die vertikale. Ein aufmerksamer Beobachter wird diesen Sachverhalt bemerken und der horizontalen Dimension mehr Beachtung schenken. Dies führt dazu, dass in dem multidimensionalen Raum diejenigen Dimensionen gedehnt werden, die beachtet werden, während diejenigen Dimensionen gestaucht werden, die weniger Aufmerksamkeit auf sich ziehen. Man kann in Abbildung 3b-7B sehen, wie dieser

Prozess die Kategorisierung optimiert, indem die beiden Kategorien weiter voneinander entfernt werden. In dem durch selektive Aufmerksamkeit veränderten Raum ist nun Objekt i eindeutig näher bei den Exemplaren der Kategorie X . Formal lässt sich dies mit Hilfe von Aufmerksamkeitsparametern für die einzelnen Merkmalsdimensionen ausdrücken, die je nach Ausprägung zu einer Stauchung (d.h. Verringerung der Aufmerksamkeit) oder Streckung (d.h. Vergrößerung der selektiven Aufmerksamkeit) der Dimensionen führen.

Abb. 3b-8

Diese Distanz muss nun noch übersetzt werden in ein Ähnlichkeitsmaß. GCM stützt sich hier auf die Literatur zur Stimulusgeneralisierung, insbesondere auf das von Shepard (1958, 1987) entdeckte *Universale Gesetz der Generalisierung*. Shepard ging davon aus, dass jeder Organismus mit der Aufgabe der Generalisierung konfrontiert ist. Erfährt ein Vogel beispielsweise, dass ein bestimmter Wurm zu Übelkeit führt, dann ist die Frage kritisch, welche anderen ähnlichen Würmer zu ähnlichen Konsequenzen (der *consequential region* eines Objekts) führen. Shepard konnte mathematisch ableiten und empirisch belegen, dass der optimale Generalisierungsgradient eine negativ exponentielle Funktion der psychologischen Distanz ist. Abbildung 3b-8 gibt ein Beispiel für diese Funktion. Wie man sehen kann, führt eine solche Funktion dazu, dass nur sehr nahe Items als ähnlich empfunden werden, während die Ähnlichkeit mit zunehmender Distanz stark abnimmt.

Man kann also sehen, dass GCM sowie alle anderen Kategorisierungstheorien eine Reihe von verschiedenen, prinzipiell unabhängig voneinander variierbaren Annahmen beinhaltet, die alle zusammen zum Erfolg der Theorie beitragen.
((Ende Textbox))

((4)) *Probleme der Exemplar-Sicht*. Obgleich die Exemplar-Sicht viele der Defizite von Prototypen-Theorien überwindet und empirisch vergleichsweise gut bewährt ist, lässt sich einwenden, dass sie möglicherweise unrealistische Annahmen an unsere

Speicherkapazität stellt. Ist es wirklich plausibel, dass wir jedes einzelne erlebte Objekt speichern und wenn nein, wie entscheiden wir, welche Exemplare wir speichern und welche wir nur als Wiederholung eines bereits gespeicherten Exemplars ansehen. Dazu sagt die Theorie bisher wenig.

Ein weiteres Problem dieses Ansatzes ist, dass er im Grunde keine Aussagen darüber macht, warum es bestimmte Kategorien gibt und was diesen Kategorien ihre Kohärenz verleiht. Gemäß eines Exemplaransatzes spricht nichts dagegen, für ein Experiment zwei Exemplare von Fahrzeugen und zwei von Lebewesen zu einer gemeinsamen Kategorie A zusammenzufassen. Es ist klar, dass die Kategorien, die wir normalerweise im Alltag erwerben, mehr sind als beliebige Zusammenfassungen von Exemplaren.

((3)) 3b-2.1.4 Theorie der Entscheidungsgrenzen

Dieser vergleichsweise neue Ansatz („decision bound theory“) basiert auf einer Weiterentwicklung der Signalentdeckungstheorie (vgl. Maddox & Ashby, 1993)(vgl. auch das Kapitel von Müsseler, Textbox ???, in diesem Band). Bisher wurde diese Theorie vorwiegend anhand von wahrnehmungsnahen, visuellen Kategorien untersucht, deren Merkmalsdimensionen kontinuierlich variieren können. Typische Aufgaben sind die Klassifikation von Rechtecken, die sich in Länge und Breite unterscheiden oder die Klassifikation von Kreisen unterschiedlicher Größe, die eine radiale Linie mit variierender Orientierung beinhalten. Dabei können mehr oder weniger komplexe Kategorisierungsregeln konstruiert werden. Außerdem wurden im Rahmen dieser Theorie häufig Aufgaben untersucht, bei denen sich die Kategorien überlappen, so dass gleichartige Exemplare verschiedenen Kategorien angehören, was eine fehlerfreie Performanz unmöglich macht und die Aufgabe sehr erschwert.

Die Grundannahme der Theorie der Entscheidungsgrenzen ist, dass die perzeptuelle Information über einen Stimulus mit einer gewissen Variabilität behaftet ist, der durch das visuelle System erzeugt wird. Diese Annahme hat zur Folge, dass die perzeptuelle Identifikation eines bestimmten Stimulus ähnliche Prozesse erfordert wie die Kategorisierung von Stimuli. In beiden Aufgaben müssen variable Stimuli auf einzelne

Reaktionen (z.B. Namen von Einzelobjekten oder Kategorienbezeichnungen) abgebildet werden.

Während nun Exemplar-Theorien davon ausgehen, dass beim Kategorisieren die gesehenen, einzelnen Exemplare abgespeichert werden und neue Exemplare mit den bereits gesehenen Exemplaren verglichen werden, postuliert die Theorie der Entscheidungsgrenzen einen einfacheren Prozess. Ähnlich wie das generalisierte Kontextmodell (vgl. Textbox: *Beispiel für eine Exemplar-Theorie: Das Generalisierte Kontext-Modell (GCM)*) geht sie davon aus, dass Beobachter die Stimuli in einem geometrischen Ähnlichkeitsraum repräsentieren. Ziel des Lernprozesses ist es nun, diskriminative Funktionen zu finden, die diesen Raum in verschiedene Regionen aufspalten. Jede Teilregion ist mit einer eigenen Reaktion assoziiert, die sich beispielsweise auf verschiedene Kategorien beziehen kann. So kann z.B. jedes Mal, wenn ein Objekt in eine Region A dieses multidimensionalen Raums fällt, der Beobachter mit der Antwort, dass das Objekt zur A-Kategorie gehört, reagieren. Die Grenzen dieser Regionen nennt man Entscheidungsgrenzen („decision bounds“). Eine einfaches Beispiel für eine Entscheidungsgrenze wäre etwa die Linie in Abbildung 3b-3 (rechts). Die Theorie ist aber nicht auf linear trennbare Strukturen beschränkt, da sie nicht postuliert, dass Entscheidungsgrenzen Geraden sein oder fehlerfreie Trennbarkeit garantieren müssen. Ziel des Lernprozesses ist lediglich die Konstruktion einer möglichst optimalen Grenze (siehe unten).

Abbildung 3b-9 veranschaulicht eine typische Aufgabe, in der Linien unterschiedlicher Länge (X-Achse) und Orientierung (Y-Achse) zur Kategorisierung vorgegeben wurden. Wie man sieht, wurde eine große Zahl von verschiedenen Stimuli gezeigt, die entweder zu einer fiktiven Kategorie A (repräsentiert als blaue Kreuze) oder Kategorie B (rote Kreise) gehörten. Die Aufgabe der Probanden bestand also darin, Linien, die bezüglich ihrer Länge und Orientierung den Positionen der Kreuze entsprachen, als zu Kategorie A gehörig zu klassifizieren, und die restlichen der Kategorie B zuzuordnen. Man kann sehen, dass die beiden Kategorien getrennte Regionen bilden, die sich allerdings überschneiden. Es ist also nicht möglich, die beiden Kategorien fehlerfrei zu

klassifizieren. Die beste Diskriminierung in zwei Regionen liefert die eingezeichnete quadratische Kategoriengrenze.

Abb. 3b-9

Kategorisierung besteht dieser Theorie zufolge also nicht in einem Ähnlichkeitsvergleich mit einer Menge von Exemplaren oder einem Prototypen (dem Mittelpunkt der Entscheidungsregion), sondern darin, dass der Beobachter entscheidet, in welche Region der jeweilige Stimulus fällt, und die dementsprechende Kategorienentscheidung trifft. In gewissem Sinne kann man diese Theorie als regelbasiert bezeichnen, da es beim Lernen um den Erwerb von Entscheidungsregeln geht, die die Kategorisierungsantwort festlegen. Diese Regeln sind aber, anders als bei der klassischen Sicht, nicht notwendigerweise definitivische Festlegungen, die eindeutig und fehlerfrei die Kategorienzugehörigkeit bestimmen. Die Regeln zielen vielmehr auf *möglichst* fehlerfreie, wenn auch nicht unbedingt perfekte Diskriminierung der Kategorien. Die Befunde zu Typikalitätsunterschieden und zu unsicheren Urteilen lassen sich im Rahmen dieses Ansatzes so erklären, dass Objekte, die nahe der Entscheidungsgrenzen liegen, schwieriger zu klassifizieren sind als Objekte, die eher zentral platziert sind.

Grundsätzlich können Entscheidungsgrenzen zwischen Kategorien unendlich viele Formen annehmen, so dass weitere einschränkende Annahmen gemacht werden müssen. Eine wichtige Theorie, die von Maddox und Ashby (1993) vorgeschlagen wurde, postuliert die Annahme von quadratischen Entscheidungsgrenzen (wie etwa in Abb. 3b-9). Diese Theorie geht von rationalen Überlegungen aus. Es lässt sich nämlich zeigen, dass unter der Annahme, dass Kategorien hinsichtlich ihrer Merkmale in der Regel multivariat normal verteilt sind, die optimale Entscheidungsgrenze, die Kategorisierungen mit dem geringsten Fehler zulässt, eine quadratische Form annimmt. (In dem speziellen Fall, dass die Merkmale der Kategorien identische Varianzen und Kovarianzen haben, ist die optimale Entscheidungsgrenze linear.) Die Grundidee ist also, dass unser kognitives System von der Evolution mit einem Bias in Richtung quadratischer oder linearer Entscheidungsgrenzen ausgestattet ist.

Dabei sagt die Theorie nicht vorher, dass wir bei jeder Aufgabe die optimalen Entscheidungsgrenzen finden. Aufgrund von Limitationen unserer Informationsverarbeitungskapazität und fehlendem Wissen über Eigenschaften der beobachteten Kategorien kann es zu suboptimalem Verhalten insbesondere bei sehr komplexen, nicht-linearen Kategorien kommen. In diesem Sinne konnten Maddox und Ashby (1993) zeigen, dass ihre Probanden bei einfachen Kategorien mit einer linearen Entscheidungsgrenze die optimale lineare Funktion lernen, bei komplexeren Kategorien aber suboptimal bleiben. Dennoch zeigte sich hier, dass die Probanden versuchen, eine Entscheidungsgrenze anzuwenden, die zwar nicht exakt der optimalen Funktion entspricht, aber eine zumindest ebenfalls quadratische Form aufweist. Wurde die Komplexität noch weiter gesteigert, so dass beispielsweise vier Kategorien mit paarweise quadratischen Funktionen separiert werden mussten, dann zeigte sich, dass dies die Möglichkeiten der Versuchsteilnehmer übersteigt. In dieser Situation fallen sie auf eine einfachere Strategie der linearen Trennung der vier Kategorienzonen zurück (Ashby, Waldron, Lee & Berkman, 2001).

Die Theorie der Entscheidungsgrenzen wurde in den letzten Jahren insbesondere gegenüber Exemplar-Theorien getestet, wobei bislang kein Konsens darüber erzielt wurde, welche Theorie überlegen ist. Der Schwerpunkt empirischer Überprüfung sind dabei die Vorhersagen für Suboptimalität beim Kategorisieren, bei denen sich die konkurrierenden Ansätze am stärksten unterscheiden. Während hier Anhänger der Theorie der Entscheidungsgrenzen eindeutige Belege für ihren Ansatz sehen, gibt es zu dieser Annahme auch heftige Gegenpositionen (vgl. McKinley & Nosofsky, 1995; Nosofsky & Johansen, 2000; vgl. auch Textbox *Die richtige Theorie möge sich bitte erheben*).

((4)) *Probleme der Theorie der Entscheidungsgrenzen.* Unabhängig von diesen empirischen Tests muss man dennoch die Frage stellen, ob die Theorie der Entscheidungsgrenzen nicht zu stark auf die Art des Versuchsmaterials hin abgestimmt wurde und deshalb vielleicht primär Geltung bei elementaren visuellen Kategorien beanspruchen kann. Bislang ist zumindest unklar, wie sich diese Theorie auf

komplexere natürliche Kategorien anwenden lässt. Ein weiteres Problem dieses Ansatzes ist, dass er zwar sehr gut in der Lage sein mag, Entscheidungsreaktionen vorherzusagen, dass es aber unklar ist, wie man sich die Repräsentation von Kategorien vorstellen soll. Im Fokus des Ansatzes sind die Entscheidungskriterien und nicht so sehr die Kategorien, die relativ beliebig als Zonen in einem multidimensionalen Ähnlichkeitsraum charakterisiert werden.

((Anfang Textbox))

Die richtige Theorie möge sich bitte erheben

Angesichts der großen Zahl konkurrierender Theorien zur Kategorisierung stellt sich die Frage, wie man entscheiden kann, welche nun richtig ist. Lange Jahre ist man so vorgegangen, dass man aus den einzelnen konkurrierenden Theorien (z.B. Prototypen- vs. Exemplartheorien) divergierende Vorhersagen abgeleitet und experimentell überprüft hat. Dies ist eine am Beginn eines Forschungsparadigmas durchaus sinnvolle Strategie, aber in den letzten Jahren wurde zunehmend klar, dass diese Strategie unbefriedigend ist. So kann auf diese Weise nicht bestimmt werden, welche Aspekte der Theorie für den jeweiligen Erfolg verantwortlich sind. Am Beispiel des generalisierten Kontextmodells (GCM)(vgl. Textbox: *Beispiel für eine Exemplar-Theorie: Das Generalisierte Kontext-Modell (GCM)*) kann man beispielsweise sehen, dass Exemplarspeicherung nur eine von vielen Annahmen ist, die zum Erfolg der Theorie beitragen. Andere davon unabhängige Annahmen wie etwa zur aufmerksamkeitsbasierten Gewichtung von Dimensionen oder die Annahme der mit wachsender Distanz exponentiell sinkenden Ähnlichkeit zwischen Reizen, sind ebenso wichtig. Smith und Minda (2000) konnten zeigen, dass viele Befunde, die scheinbar eindeutig Exemplartheorien bestätigt haben, genauso gut durch ein Prototypen-Modell erklärt werden können, wenn man anstatt der in diesem Modell üblichen linearen Ähnlichkeitsregel die exponentielle Funktion von GCM integriert und die Annahme macht, dass die in der Lernphase erlebten Exemplare einen gewissen Gedächtnisvorteil gegenüber den nicht gesehenen haben.

Ein weiteres Problem beim Modellvergleich besteht darin, dass verschiedene Modelle unterschiedliche Anzahlen von freien Parametern haben. So hat GCM etwa Parameter für die Aufmerksamkeitsgewichte bei der Ähnlichkeitsbestimmung. Diese Parameter werden aus den Daten geschätzt und erzeugen aufgrund ihrer freien Bestimmbarkeit eine gewisse Flexibilität bei der Datenanpassung. Modelle, die von multiplen Systemen ausgehen, haben aufgrund ihrer erweiterten Struktur in der Regel zusätzliche freie Parameter, so dass dies allein schon erklären kann, warum sie scheinbar überlegen sind bei der Erklärung von Daten. Nosofsky und Johansen (2000) konnten allerdings zeigen, dass bestimmte Zusatzannahmen bei der Bestimmung der Parameter von GCM und ein weiterer Parameter für die Modellierung der Antwortentscheidungen ausreicht, um viele für die Theorie scheinbar kritische Befunde zu erklären. Sie betonen allerdings, dass dies noch nicht bedeutet, dass die Theorie, die nur von einem einzelnen System ausgeht, richtig ist. Es zeigt nur, dass man auch vorschnell komplexere Theorien akzeptieren kann. Neben Experimenten, die auf die Überprüfung spezifischer Annahmen von Theorien zielen, bieten in Zukunft voraussichtlich neurowissenschaftliche Studien eine wichtige weitere Datenquelle für die Theoriendiskriminierung (vgl. z.B. Ashby et al., 1998).((Ende Textbox))

((3)) 3b-2.1.5 Theorien multipler Systeme

In den letzten Jahren wurde eine Reihe von Theorien vorgeschlagen, die von multiplen Kategorisierungssystemen ausgehen. Zunehmend mehr Ansätze postulieren, dass verschiedene Systeme, beispielsweise ein regel- und ein exemplarbasiertes System, bei einer Kategorisierungsaufgabe zusammenarbeiten oder dass es interindividuelle oder auch übungsabhängige intraindividuelle Unterschiede bei der Nutzung der einzelnen Systeme gibt. Solche komplexen Modelle sind nicht nur empirisch häufig den einfachen Theorien überlegen (vgl. aber Nosofsky & Johansen, 2000), sie werden auch durch neuere neurowissenschaftliche Studien gestützt. Uneinigkeit besteht derzeit allerdings noch darüber, welche der Systeme nebeneinander bestehen.

((4)) *Regeln und Prototypen.* Die Probleme der klassischen Sicht und die der Prototypen-Sicht haben eine Reihe von Forschern dazu geführt, beide Ansätze zu

kombinieren (Armstrong, Gleitman & Gleitman, 1983; Landau, 1982; Miller & Johnson-Laird, 1976; vgl. auch Weinert & Waldmann, 1988). Gegen Prototypen-Theorien wurde eingewendet, dass sie dem Sachverhalt nicht gerecht werden, dass manche Konzepte, wie etwa Familienkategorien (z.B. „Großvater“) oder mathematische Konzepte wie „ungerade Zahl“ eindeutig definierbar sind und diese Definitionen auch von Versuchsteilnehmern generiert werden können. Auf der anderen Seite stehen aber Befunde, die zeigen, dass selbst diese Kategorien Typikalitätsgradienten aufweisen. So wird eine „7“ als ein typischerer Vertreter der Kategorie „ungerade Zahl“ angesehen als beispielsweise eine „87“. Diese Unterschiede zeigen sich auch in indirekten Maßen wie Reaktionszeiten (Armstrong, Gleitman & Gleitman, 1983).

Man hat deshalb einen dualen Ansatz vorgeschlagen, der zwei Ebenen der Kategorienrepräsentation vorsieht, einen zentralen Kernbereich, der die Definition beinhaltet und vorwiegend bei Denkprozessen eingesetzt wird, und eine oberflächliche, prototypenbasierte Identifikationsprozedur, die bei der Kategorisierung von Objekten genutzt wird. Eine Einschränkung dieses Ansatzes ist allerdings, dass er sich nur auf Konzepte anwenden lässt, die einen definitiven Kern haben, was sicherlich nur auf eine Minderzahl zutrifft. Außerdem teilt er mit den Prototypen-Theorien alle Probleme bei der Erklärung von Kategorisierungsphänomenen.

((4)) *Prototypen und Exemplare.* Dieser Ansatz geht von der nahe liegenden Annahme aus, dass Beobachter sowohl in der Lage sind, Prototypen zu abstrahieren als auch Exemplare zu speichern. Smith, Murray und Minda (1997) vermuteten, dass die Wahl der Lernstrategie von der Art der zu erlernenden Kategorien gesteuert wird (vgl. auch Medin, Altom und Murphy, 1984). In ihren Experimenten zeigte sich, dass Exemplar-Theorien immer dann überlegen waren, wenn die Kategorien klein, hochdifferenziert und wenig kohärent waren, wenn also ein hohes Maß an Unähnlichkeit innerhalb der Kategorie besteht. In solchen Situationen bilden Prototypen keine gute Repräsentation der Gesamtkategorie, so dass es den Lernenden offenbar sinnvoller erscheint, auf das Einprägen der wenigen Einzelexemplare zurückzufallen (vgl. aber Nosofsky & Johansen, 2000, für eine andere Sicht dieser Daten). Sind die Kategorien hingegen eher groß und vergleichsweise kohärent, dann zeigten sich interindividuelle Unterschiede:

Etwa die Hälfte der Versuchsteilnehmer wurde besser durch einen Prototypen-Ansatz charakterisiert (vgl. auch Textbox: *Die richtige Theorie möge sich bitte erheben*).

((4)) *Regeln und Exemplare*. Ein Beispiel für eine Theorie, die eine regelbasierte Komponente mit Exemplarspeicherung kombiniert, stammt von Nosofsky, Palmeri und McKinley (1994). Dieses Modell geht von der Annahme aus, dass Exemplartheorien zu große Ansprüche an die Speicherkapazität des Lernenden stellen. Ihr *Rule-plus-exception* Modell (RULEX) geht demgegenüber davon aus, dass Lernen ein sequentieller Prozess ist. In einem ersten Lernschritt wird nach einer einfachen, auf einer einzelnen Dimension basierenden Regel gesucht, die die Kategorien trennt. Gibt es eine solche Regel nicht, sucht RULEX nach einer *partiell* erfolgreichen einfachen Regel. (Gibt es keine hinreichend erfolgreiche partiell gültige Regel, dann sucht RULEX zusätzlich konjunktive Regeln, die Konjunktionen von Dimensionen involvieren.) RULEX könnte beispielsweise in einer Lernaufgabe zu der einfachen Regel kommen, dass die meisten Exemplare von Kategorie A rund sind und die von Kategorie B eckig. Findet es eine solche Regel, die hinreichend viele Exemplare erfolgreich klassifiziert, versucht RULEX Ausnahmen von dieser Regel zu memorieren. Nur selten werden dabei vollständige Lernexemplare gespeichert. Auch hier versucht RULEX möglichst ökonomisch zu sein und nur mit partiellen Beschreibungen der Ausnahmen auszukommen. So könnte RULEX speichern, dass eckige und blaue Objekte (mit beliebigen anderen Merkmalen) allerdings auch zu Kategorie A gehören. Neue Exemplare werden in der umgekehrten Reihenfolge dieser Schritte kategorisiert. Wird RULEX mit einem zu kategorisierenden Testexemplar konfrontiert, sucht es zunächst die (partiell) gespeicherten Ausnahmen ab, dann die konjunktiven Regeln (falls vorhanden) und schließlich die einfachen Regeln.

RULEX wurde gegenüber dem generalisierten Kontextmodell (GCM) getestet und erwies sich als gleichwertig bei der Erklärung von Gruppendaten. Als größter Vorzug wurde aber gesehen, dass es besser als GCM in der Lage ist, interindividuelle Unterschiede bei den Kategorisierungen zu erklären (vgl. aber Nosofsky & Johansen, 2000, für eine Kritik an dieser Schlussfolgerung).

Erickson und Kruschke (1998) haben ebenfalls ein Modell vorgeschlagen, das eine Regel- und eine Exemplar-Komponente verknüpft. Bei der empirischen Überprüfung ihrer Theorie machten sich diese beiden Forscher einen wichtigen Vorzug abstrakter Regeln bei der Generalisation des erlernten Wissens zunutze: Regeln erlauben die *Extrapolation* auf Exemplare, die nicht im Bereich der Lernexemplare liegen. Dies lässt sich am besten an einem artifiziellen Beispiel verdeutlichen (vgl. auch Marcus, 1998). Angenommen, wir haben gelernt, dass ein Rechteck immer zur Kategorie A gehört, wenn seine Länge doppelt so hoch ist wie seine Breite, ansonsten gehört es zu Kategorie B. Nun kann es sein, dass wir in der Trainingsphase Rechtecke zwischen 1 und 20 cm Länge und Breite gesehen haben. Exemplartheorien haben wenig Probleme mit der Klassifikation neuer Rechtecke, die innerhalb dieses Trainingsbereichs (d.h. zwischen 1 und 20 cm) liegen. Sie sind in der Lage, *Interpolationen* durchzuführen. Anders sieht es hingegen bei Rechtecken außerhalb des Trainingsbereichs aus (z.B. ein 80 cm langes und 40 cm breites Rechteck). Während wir auf der Basis der Regel eindeutig sagen können, dass dieses Rechteck zu Kategorie A gehört, würde eine Exemplar-Theorie vermutlich versagen, weil sie einfach versuchen würde, die Ähnlichkeit dieses neuen Exemplars zu den bereits gesehenen zu bestimmen. In ihren Experimenten konnten die beiden Forscher zeigen, dass sowohl regelbasierte Extrapolationen als auch exemplarbasierte Ähnlichkeitsvergleiche das Kategorisierungsverhalten steuern.

((4)) *Explizite und implizite Regeln.* Der letzte Ansatz, der hier erwähnt soll, stellt eine Erweiterung der Theorie der Entscheidungsgrenzen dar. Ashby, Alfonso-Reese, Turken und Waldron (1998) haben eine durch neuropsychologische Daten inspirierte Theorie vorgeschlagen, die von zwei parallelen Lernsystemen ausgeht, die miteinander in Wettbewerb stehen (vgl. auch Ashby et al., 2001). Das *implizite System* operiert weitgehend jenseits des Bewusstseins und lernt Kategorien im Sinne der Theorie der Entscheidungsgrenzen (siehe Abschnitt 2D). Dieses System, das von Ashby und Mitarbeitern primär im Striatum (Putamen und Nucleus caudatus), dem Input-Bereich der Basalganglien, angesiedelt wird, fasst also Stimuli zu kategorialen Zonen zusammen. Diese Entscheidungsgrenzen können komplexen (wenn auch manchmal suboptimalen) Funktionen entsprechen, die von den Lernenden zwar gelernt, aber nicht verbalisiert werden können. Sie gehen hier deshalb von implizitem (bzw. prozeduralem)

Lernen aus. Ein Beispiel für eine Aufgabe, die eher von dem impliziten Lernsystem bewältigt werden kann, findet sich in Abbildung 3b-9.

Parallel zu diesem impliziten Lernsystem wird von Ashby und Mitarbeitern ein *explizites System* postuliert. Das explizite System verwendet dem Bewusstsein zugängliche Regelhypothesen, die relativ einfach verbalisierbar sind. Dabei kommen primär einfache Hypothesen in Frage, die sich auf die trennbaren, leicht verbalisierbaren Dimensionen beziehen (also auf die Achsen im geometrischen Raum, der das Lernmaterial beschreibt, wie z.B. Länge, Breite). Eine solche Regel könnte beispielsweise lauten: „Antworte mit A, wenn das Rechteck eher eine große Länge hat.“ Explizite Regelhypothesen beziehen sich also auf leicht wahrnehmbare und sprachlich benennbare Dimensionen im Lernmaterial. Ein Beispiel für eine Aufgabe, die sich leicht durch explizites Regellernen bearbeiten lässt, findet sich in Abbildung 3b-2. Hier müssen die Probanden lediglich lernen, dass die beiden Kategorien sich in der Farbe des Hintergrunds unterscheiden, auf denen die Stimuli präsentiert wurden. Es handelt sich also um eine einfache eindimensionale Regel.

Ashby et al. (1998) vermuten, dass die Auswahl der Regelhypothesen im Wesentlichen vom präfrontalen Cortex und dem anterioren Cingulum gesteuert wird, wobei sie annehmen, dass die exakten Entscheidungsgrenzen für jede Regel wiederum im Striatum (Nucleus caudatus) bestimmt werden (vgl. Abb. 3b-10).

Ashby et al. (1998) haben diese Annahmen im Rahmen eines konnektionistischen Lernmodells (COVIS) formalisiert, dass von einem Wettbewerb zwischen dem expliziten und dem impliziten Lernsystem ausgeht. Während das implizite System optimale oder bei komplexen Aufgaben gelegentlich auch suboptimale, komplexe Entscheidungsgrenzen aus den Lerndaten abstrahiert, testet das explizite System verschiedene alternative Regelhypothesen. Diejenige explizite Hypothese setzt sich allmählich durch, die zur relativ größten Zahl korrekter Kategorisierungen führt. Die Kategorisierung der Objekte wird dann von beiden Systemen, dem expliziten und dem impliziten, beeinflusst, wobei dasjenige System bei der jeweiligen Kategorisierungsentscheidung dominiert, das eine Entscheidung mit größerer

Zuversicht zulässt. Es kann also sein, dass manche Objekte eher vom expliziten, andere eher vom impliziten System kategorisiert werden.

In dem Grenzfall, dass die expliziten Hypothesen keine zutreffenden Kategorisierungen erlauben, werden alle Entscheidungen vom impliziten System getroffen. Eine interessante Vorhersage macht das Modell allerdings für den Fall, dass sich die Entscheidungsgrenzen des impliziten und des expliziten Systems unterscheiden, beide Systeme aber relativ gute Leistungen erbringen. Angenommen beispielsweise, die Aufgabe besteht darin, Rechtecke zu klassifizieren und die optimale Entscheidungsgrenze wird durch eine Diagonale in dem zweidimensionalen Ähnlichkeitsraum festgelegt, der die Achsen Länge und Breite hat. Die dominierende verbale Hypothese, die ebenfalls viele akkurate Kategorisierungen erlaubt, liegt parallel zu einer der beiden Achsen, verwendet also eine der beiden Dimensionen allein. Für diesen Fall macht COVIS die interessante Vorhersage, dass die Entscheidungsgrenze, die vom Lernenden am Ende der Lernphase genutzt wird, eine Gerade sein sollte, die zwischen den beiden Geraden liegt, die vom impliziten und vom expliziten System generiert wurden. Diese Gerade ist also flacher als die optimale implizite Entscheidungsfunktion und damit suboptimal. Ashby et al. (1998) haben tatsächlich in einem Experiment solche suboptimalen Entscheidungsgrenzen nachgewiesen (vgl. aber Nosofsky & Johansen, 2000, für eine alternative Sicht dieser Befunde).

((4)) Neurowissenschaftliche Evidenz für multiple Systeme. Neuropsychologische Untersuchungen mit Patienten und Studien mit bildgebenden Verfahren haben in den letzten Jahren zusätzliche Hinweise auf die Existenz multipler Kategorisierungssysteme gegeben. Generell entsteht das Bild, dass bei jeder Aufgabe mehrere Gehirnbereiche zusammenwirken, dass es aber auch aufgabenspezifische Muster gibt. So haben Smith, Patalano und Jonides (1998) eine Kategorisierungsbedingung, in der eine bestimmte Regel anzuwenden war, mit einer Kontrollbedingung verglichen, in der erwartet wurde, dass die Versuchsteilnehmer eher auf Exemplarvergleiche zurückgreifen würden. Die Aufgabe wurde ursprünglich von Allen und Brooks (1991) entwickelt und untersucht. Während des Lernens wurde eine Positronen-Emissions-Tomographie (PET) durchgeführt, die zeigte, dass bei beiden Aufgaben eine große Zahl von Gehirnregionen

aktiv war. Ein interessanter Befund war, dass in der Regelbedingung 14 von 23 signifikant aktiven Regionen nur in dieser Bedingung aktiviert waren. Zu den spezifischen Regionen gehören der linke Parietallappen und der präfrontale Cortex, Bereiche, die mit selektiver Aufmerksamkeit und Arbeitsgedächtnis in Zusammenhang gebracht werden. Dieses Muster ist konsistent mit der Annahme, dass Regellernen ein bewusster, analytischer Prozess ist, während Exemplarlernen eher auf visuell-perzeptive und Gedächtnisregionen rekurriert.

Während Smith et al. (1998) zwischen Regel- und Exemplarlernen unterscheiden, verwenden andere Forscher eher die Dichotomie zwischen explizitem und implizitem Lernen (vgl. Ashby et al., 1998; Ashby & Ell, 2001). Abbildung 3b-10 veranschaulicht die neurowissenschaftlichen Grundlagen, die der Architektur des von Ashby et al. (1998) entwickelten konnektionistischen Lernmodells COVIS zugrunde gelegt wurden. Explizites Lernen entspricht weitgehend bewusstem Regellernen. Ashby und Ell (2001) kommen aufgrund eines Literaturüberblicks zur Schlussfolgerung, dass explizites Lernen wesentlich durch ein Zusammenspiel zwischen präfrontalem Cortex und den Basalganglien (insbesondere dem Nucleus caudatus) gesteuert wird.

Abb. 3b-10

Dass implizites Kategorienlernen jenseits des Bewusstseins und expliziter Regellernenprozesse möglich ist, haben neuropsychologische Studien mit Amnestikern gezeigt, die aufgrund von Schädigungen der medial-temporalen (Hippocampus u.a.) oder diencephalischen Areale keine bewusste (explizite) Erinnerung an die Lernitems hatten. Trotz dieses Defizits zeigte sich in einer Prototypenlernaufgabe mit artifiziellem Lernmaterial, dass die Amnestiker nahezu ebenso gut wie Gesunde in der Lage waren, die Kategorisierungsaufgabe zu lernen (Knowlton & Squire, 1993).

Welche Bereiche sind nun aber für das implizite Lernen verantwortlich? In den letzten Jahren zeigte eine Reihe von Befunden, dass es keinen einheitlichen Gehirnbereich gibt, der für implizites Lernen verantwortlich ist, sondern dass man eher davon ausgehen muss, dass es aufgabenabhängige Interaktionen gibt. So werden einfache motorische

Assoziationsprozesse mit dem Cerebellum in Zusammenhang gebracht, während bei der Konditionierung emotionaler Reize die Amygdala eine wichtige Rolle spielt (vgl. Beggs et al., 1999).

Was nun den Bereich des impliziten Lernens von Kategorien betrifft, mehren sich Befunde, die von einer Beteiligung des Striatum, der Input-Region der Basalganglien, ausgehen (vgl. die Übersicht in Ashby et al., 1998; Ashby & Ell, 2001). In einer neuropsychologischen Studie von Knowlton, Mangels und Squire (1996), in der die Probanden lernen sollten, probabilistische Hinweisreize (Muster auf Karten) für die Vorhersage einer fiktiven Wetterlage zu nutzen, zeigte sich, dass Amnestiker gut in der Lage waren, diese Aufgabe zu meistern. An der Parkinson-Krankheit leidende Patienten hingegen, die mit einer neuronalen Degeneration der Substantia nigra und des Striatum einhergeht, lernten die Aufgabe nicht. Sie waren im Gegensatz zur Amnestiker-Gruppe hingegen in der Lage, sich an die Items in einer Wiedererkennungsaufgabe bewusst zu erinnern. Stützende Befunde für die Beteiligung des Striatum beim Kategorisieren ergab auch eine Studie mit Patienten, die an der Huntingtonschen Erkrankung litten, die ebenfalls das Striatum schädigt. Auch diese Patienten waren nicht in der Lage, die probabilistische Wettervorhersageaufgabe zu lernen, obwohl sie eine andere implizite Lernaufgabe, den Erwerb von Regeln einer artifiziellen Grammatik, lernen konnten (Knowlton, Squire, Paulsen, Swerdlow, Swenson & Butters, 1996). Stützende Befunde für diese Schlussfolgerungen stammen auch aus einer Studie von Poldrack, Prabhakaran, Seger und Gabrieli (1999), die die Wettervorhersageaufgabe bei gesunden Probanden eingesetzt haben und während des Lernens eine funktionale Kernspintomographie (fMRI) durchgeführt haben. Diese Studie deutet auf ein Zusammenwirken des linken, dorsolateralen präfrontalen Cortex und des Striatum (insbesondere des Nucleus caudatus), wobei mit zunehmendem Lernen die Aktivierung des präfrontalen Bereichs schwächer wurde. Ashby et al. (1998) sehen dies als Beleg für ihre Hypothese, dass am Anfang des Lernens das explizite, verbale System dominiert, während später die Verarbeitung vom impliziten System übernommen wird.

Im Vergleich zur dominierenden Rolle des präfrontalen Cortex bei expliziten und den Basalganglien bei impliziten und expliziten Kategorienlernprozessen spielen andere

Gehirnbereiche, die man mit Lernen in Verbindung bringt, eine eher untergeordnete Rolle. So konnten Knowlton, Squire und Gluck (1994) zeigen, dass die medialen Temporallappen, die eine bedeutende Rolle beim expliziten Erinnern spielen, in späten Phasen des Kategorienlernens involviert sein können, wenn die Probanden versuchen, sich die Lernexemplare einzeln einzuprägen. Der visuelle Cortex ist eine weitere Region, deren Aktivierung man bei Aufgaben nachweisen konnte, in denen Probanden visuell dargebotene Prototypenstrukturen (vgl. Abb. 3b-4) lernen, wie sie etwa von Posner und Keele (1968) untersucht wurden (vgl. Ashby & Ell, 2001).

((2)) 3b-2.2 Kritik ähnlichkeitsbasierter Theorien

Die bisher besprochenen Theorien vereint die Grundannahme, dass Kategorisierung auf Ähnlichkeit basiert. Diejenigen Objekte und Ereignisse werden zusammengefasst, die sich untereinander ähnlich sind, und Kategorisierung beruht darauf, dass man die Ähnlichkeit zwischen neuen Objekten und den bereits vorhandenen Kategorien bestimmt.

In den letzten Jahren mehren sich allerdings Stimmen, die die fundamentale Rolle von Ähnlichkeit in Frage stellen (vgl. Medin et al., 2001). Ein Problem entsteht dadurch, dass Ähnlichkeit sehr variabel und kontextabhängig sein kann. So konnten Medin und Shoben (1988) zeigen, dass im Kontext von Haar die Farbe grau als ähnlicher zu weiß als zu schwarz empfunden wird, während im Kontext von Wolken grau näher an schwarz heranrückt. Dies ist ein Beispiel von Kontextsensitivität, die durch Hintergrundtheorien über den der Haarfarbe zugrunde liegenden Alterungsprozess bzw. über die Beziehung der Wolkenfarbe zum Wetter erzeugt wird.

Ein noch radikaleres Beispiel von Kontextsensitivität stammt von Medin, Goldstone und Gentner (1993). In einer Studie wurde gezeigt, dass selbst die Art der Merkmale eines Objekts, die in einem Ähnlichkeitsvergleich berücksichtigt werden, davon abhängen kann, welche Merkmale das Vergleichsobjekt hat (vgl. Abb. 3b-11). So wird in Abbildung 3b-11 das mittlere Objekt B als dreizackig wahrgenommen, wenn es mit A verglichen wird, aber als vierzackig, wenn es mit C verglichen wird.

Abb 3b-11

Andere Beispiele von Kontextsensitivität wurden im Zusammenhang mit Tverskys (1977) einflussreichem Kontrastmodell der Ähnlichkeit untersucht. Dieses Modell geht davon aus, dass die Ähnlichkeit zweier Objekte eine gewichtete Funktion der gemeinsamen Merkmale und der für jedes Objekt distinkten, einzigartigen Merkmale ist, wobei Tversky folgende Funktion zugrunde legt:

$$S(A, B) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$$

Diese Gleichung modelliert die Ähnlichkeit S von A und B als eine Funktion der gemeinsamen Merkmale ($A \cap B$), der Merkmale von A , die B nicht hat ($A - B$) und der Merkmale von B , die A nicht hat ($B - A$), wobei f eine monoton wachsende Funktion ist und θ, α, β Gewichtungparameter für die einzelnen Komponenten bezeichnen.

Tversky konnte zeigen, dass sowohl die Gewichtung der einzelnen Merkmale als auch die Gewichtung der für die beiden Objekte distinkten Merkmale variabel und kontextabhängig ist. So werden ein Amerikaner aus Maine und einer aus Florida als unähnlicher angesehen, wenn sie sich in Washington begegnen, als wenn sie sich in Tokio treffen. Ein anderes Beispiel sind Richtungsasymmetrien, die schwer vereinbar sind mit geometrischen Ähnlichkeitsmodellen, wie sie in vielen Kategorisierungsmodellen verwendet werden. So wird die Ähnlichkeit eines Granatapfels zu einem Apfel höher eingestuft als die Ähnlichkeit eines Apfels zu einem Granatapfel. Während geometrische Modelle von einer Symmetrie der Ähnlichkeit ausgehen, kann das Kontrastmodell diesen Befund erklären, wenn es von einer unterschiedlichen Gewichtung der distinkten Merkmale von Apfel und Granatapfel ausgeht (vgl. auch Smith, 1995).

Philosophen haben auf ein weiteres Problem des Ähnlichkeitskonzepts aufmerksam gemacht (Goodman, 1972). Solange man keine Einschränkungen für die Auswahl der Merkmale macht, kann man verschiedene Objekte arbiträr ähnlich oder unähnlich machen. Die Anzahl gemeinsamer oder verschiedener Merkmale zweier Objekte ist grundsätzlich unendlich groß. Dies haben Murphy und Medin (1985) am Beispiel von

Pflaumen und Rasenmäher gezeigt. Beide wiegen weniger als 1000 kg, hören schlecht, lassen sich auf der Erde finden, haben Teile und können von Elefanten getragen werden. Solche Merkmale werden natürlich von Versuchsteilnehmern nie genannt, was zeigt, dass die in Befragungsstudien (z.B. Rosch & Mervis, 1975) ermittelten Merkmalslisten von Vorwissen und weiteren Faktoren abhängig sind und nicht einfach Auflistungen der objektiv möglichen Merkmale darstellen. Keil (1981) konnte zudem zeigen, dass Versuchsteilnehmer wichtige Merkmale oft nicht nennen, obwohl sie ihr Kategorisierungsverhalten bestimmen. So finden sich die Merkmale „hat ein Herz“, „atmet“, „ist ein physikalisches Objekt“ selten in den frei generierten Listen von Versuchsteilnehmern, möglicherweise weil diese Merkmale als selbstverständlich angesehen werden.

Nicht alle Ähnlichkeitsbasierten Theorien gehen davon aus, dass Kategorisierungen dieselbe Art von Ähnlichkeitsvergleich zugrunde liegt wie freien Ähnlichkeitsvergleichen, wie sie etwa Tversky (1977) modelliert hat (vgl. Nosofsky & Johansen, 2000). Wir haben bereits gesehen (Textbox: *Beispiel für eine Exemplar-Theorie: Das Generalisierte Kontext-Modell (GCM)*), dass das generalisierte Kontextmodell (GCM) beispielsweise postuliert, dass Kategorien zu einer selektiven Aufmerksamkeit auf relevante Merkmale führt, was zu einer Verzerrung des zugrunde gelegten geometrischen Ähnlichkeitsraums und damit zu einer Veränderung der Ähnlichkeitsbeziehungen führt (vgl. auch Goldstone, 1994b, für empirische Evidenz). Dieses Modell geht aber dennoch von stabilen Merkmalen aus, die mehr oder weniger stark beachtet werden.

Die Annahme stabiler Merkmale wurde in den letzten Jahren allerdings ebenfalls in Frage gestellt. In einer Reihe von Experimenten konnte gezeigt werden, dass es auch lernabhängige („top down“) Einflüsse auf die Art der wahrgenommenen Merkmale gibt (vgl. die Übersicht in Schyns, Goldstone & Thibaut, 1998). So konnten Goldstone und Steyvers (2001) zeigen, dass in einer Aufgabe, in der die Versuchsteilnehmer lernten, computergenerierte Gesichter in vom Versuchsleiter vorgegebene Klassen einzuteilen, während des Lernprozesses neue, vorher nicht genutzte Gesichtermerkmale abstrahiert wurden, die den Lernprozess erleichterten (vgl. auch Schyns & Rodet, 1997).

Kategorien können also nicht nur die selektive Aufmerksamkeit steuern, sondern auch zu einer Neustrukturierung des repräsentierten Merkmalsraums führen.

Abb. 3b-12

Abbildung 3b-12 zeigt das Material von Goldstone und Steyvers, das bezüglich der zwei künstlich generierten Dimensionen A und B angeordnet ist. Die Pole der beiden Dimensionen sind durch die Gesichter 1 bis 4 veranschaulicht. Die Stimuli in der Mitte sind gemorphte Mischungen dieser vier Gesichter. Für die Lernaufgabe wurden den Probanden in einem Experiment acht dieser Gesichter präsentiert, die in zwei Kategorien A und B klassifiziert werden sollten. Dabei verlief für verschiedene Versuchsgruppen die Kategoriengrenze entweder horizontal, so dass nur Dimension B relevant war, oder vertikal, so dass nur Dimension A relevant war. Die Forscher konnten zeigen, dass die Probanden nach etwa 200 Lerndurchgängen die künstlich erzeugten Dimensionen A und B gelernt hatten und bei neuen Kategorisierungen in Abhängigkeit von der Lernbedingung nutzten.

Selbst wenn es einmal möglich sein sollte, das Ähnlichkeitskonzept befriedigend zu beschreiben, bedeutet dies noch nicht, dass dadurch Kategorisierung vollständig beschrieben ist. Ein Beispiel für die mögliche Entkoppelung von Ähnlichkeit und Kategorisierung stammt von Rips (1989). Er hat Versuchsteilnehmer gebeten, sich ein rundes Objekt vorzustellen, das einen Durchmesser von 7,5 cm hat. Eine Gruppe wurde gebeten, die Ähnlichkeit dieses Objekts zu einem Geldstück (ein 25 Cent Stück) und zu einer Pizza einzuschätzen, während eine andere Gruppe, die Kategorisierungsgruppe, gefragt wurde, ob das Objekt eher eine Pizza oder ein Geldstück sei. Die Befunde zeigten eine interessante Dissoziation. Während das Objekt als dem Geldstück ähnlicher empfunden wurde, gab die Kategorisierungsgruppe dennoch an, dass es vermutlich eher eine Pizza sei. Rips erklärt dies damit, dass die Versuchsteilnehmer Vorwissen nutzten, das ihnen sagte, dass die Größe einer Pizza flexibler sei als die von Geldstücken (vgl. auch Smith & Sloman, 1994).

Abb. 3b-13

Ein weiteres Beispiel für die mögliche Entkoppelung von Ähnlichkeit und Kategorisierung stammt von Gelman und Markman (1986). Die beiden Forscherinnen zeigten Kindern Bilder von drei Tieren, einem Flamingo, einer Amsel und einer Fledermaus (vgl. Abb. 3b-13). Wie man in Abbildung 3b-13 sehen kann, sahen sich die Amsel und die Fledermaus oberflächlich ähnlicher als die Amsel und der Flamingo. Wurde den Kindern gesagt, dass der Flamingo und die Amsel Vögel seien, also zur selben Kategorie gehören, dann schlossen selbst 4-Jährige, dass diese beiden Exemplare eher weitere, nicht-sichtbare Merkmale (z.B. über die Fütterungsmethode) teilen als die beiden oberflächlich ähnlichen Tiere, die nicht als zur gleichen Kategorie gehörig bezeichnet wurden. Diese Studie zeigt, dass selbst junge Kinder in der Lage sind, Oberflächenähnlichkeit zugunsten anderer Grundlagen der Kategorisierung zu ignorieren.

((4)) Zusammenfassung

Die diskutierten Argumente zeigen, dass das Ähnlichkeitskonzept bislang keineswegs befriedigend theoretisch aufgeklärt ist. Ähnlichkeit ist nicht nur eine mögliche Ursache von Kategorien, es gibt auch Belege, die zeigen, dass sie ein Produkt von Kategorisierungsprozessen sein kann. Dies muss aber nicht bedeuten, dass man das Ähnlichkeitskonzept vollständig fallen lassen sollte. Viele Forscher gehen davon aus, dass basale, perzeptive Ähnlichkeitsvergleiche eine zentrale Rolle beim Kategorisieren spielen, selbst wenn wir noch keine vollständige Theorie dessen haben, wie diese Vergleiche vollzogen werden. Es ist aber klar, dass eine befriedigende theoretische Klärung des Konzepts der Ähnlichkeit eine schwierige, bisher nur ansatzweise gelöste Aufgabe für die aktuelle Forschung darstellt (vgl. auch Goldstone, 1994a; Goldstone & Barsalou, 1998; Medin et al., 1993; Smith & Samuelson, 1997).

((2)) 3b-2.3 Die Theorien-Sicht

Die Probleme ähnlichkeitsbasierter Theorien haben eine Reihe von Kognitions- und auch kognitiven Entwicklungspsychologen zu einer alternativen Sichtweise gebracht, die davon ausgeht, dass Konzepte wissensbasiert sind und von intuitiven Theorien über

die Welt beeinflusst werden (vgl. Carey, 1985; Gopnik & Meltzoff, 1997; Medin et al. 2001; Murphy & Medin, 1985; Waldmann, 1996a, b). Diese Beziehung wurde in den ähnlichkeitsbasierten Theorien, deren empirische Forschung häufig auf artifiziellem Lernmaterial beruhte, häufig nicht gesehen oder bewusst ausgeblendet. Sieht man sich aber natürliche Kategorien wie z.B. das Konzept „Lebewesen“ oder „Fahrzeug“ an, dann wird klar, dass man sich die Repräsentation solcher Kategorien nicht einfach als Bündel von unabhängigen Merkmalen vorstellen kann, sondern dass die Merkmale häufig durch kausale oder funktionale Relationen verknüpft sind. Die Merkmale „kann fliegen“ und „hat Flügel“ beim Konzept „Rotkehlchen“ stehen in einer relationalen Beziehung, die einen Teil unseres Konzeptwissens ausmacht.

Murphy und Medin (1985) gehen deshalb davon aus, dass Kategorisierung ein Prozess ist, der der Beziehung von wissenschaftlicher Theorie zu empirischen Daten eher entspricht als einfachen Ähnlichkeitsvergleichen. Kategorien spielen dabei die Rolle von intuitiven Theorien und die Lernexemplare die Rolle von Daten, für die man entscheiden muss, ob sie zu der Theorie passen. Dabei geht es nicht nur darum zu bestimmen, ob in den Daten die von der Theorie geforderten Merkmale vorhanden sind, sondern auch darum zu entscheiden, ob die von der Theorie postulierten explanatorischen Relationen vorliegen. So sind wir beispielsweise in der Lage, das Konzept „Betrunkener“ auf eine Person anzuwenden, die auf einer Party in einen Swimmingpool springt, weniger weil dies ein charakteristisches Merkmal von Betrunkenen ist, sondern weil wir eine abstrakte Theorie haben, die solches Verhalten als mögliche Konsequenz des Alkoholpegels erklärt (vgl. Medin et al., 2001).

Eine wichtige Besonderheit der Beziehung von Theorien und Daten ist, dass diese beiden Komponenten nicht unabhängig voneinander sind. Die Art und Weise, wie Daten beschrieben werden, ist zum Teil abhängig von den Theorien, die man an sie heranträgt. Dies hat eine Studie von Wisniewski und Medin (1994) gezeigt, in der Versuchsteilnehmer Kinderzeichnungen kategorisieren sollten (vgl. Abb. 3b-14). Dabei zeigte sich, dass die Probanden andere Merkmale in den Zeichnungen fanden, wenn sie dachten, sie wären zum Teil von hochintelligenten Kindern produziert worden, als wenn sie die Hypothese hatten, dass sie von intellektuell normalen Kindern gezeichnet

wurden. Merkmale wie „ungewöhnliche und kreative Perspektive“, die von manchen Versuchsteilnehmern verwendet wurden, die dachten, hochbegabte Kinder hätten die Bilder gezeichnet, ist ein hochtheoretischer Begriff, dessen Anwendung auf eine Kinderzeichnung eine Reihe von Inferenzen erfordert. Dieses Beispiel belegt erneut, dass Merkmale nicht objektiv vorfindliche Eigenschaften der Daten sind, sondern eine Konstruktion, die aus der Interaktion von Theorie („top down“) und Daten („bottom up“) resultiert.

Abb. 3b-14

Eine wichtige Funktion von Theorien ist, dass sie eine Begründung dafür liefern, warum wir bestimmte Kategorien haben und was ihre Kohärenz ausmacht. Eine Reihe von Untersuchungen hat gezeigt, dass man durch die Aktivierung von Vorwissen Lernende dazu bringen kann, eine kohärente Kategorie zu sehen, die sie ohne dieses Wissen eher nicht wahrgenommen hätten. Dabei stellte sich heraus, dass das Ausmaß von Kohärenz einen wichtigeren Einfluss auf die Lernschwierigkeit von Kategorien hat als abstrakte Merkmale des Lernmaterials wie etwa dessen lineare Trennbarkeit. Ein Beispiel dafür ist eine Studie von Wattenmaker, Dewey, Murphy und Medin (1986). In dieser Untersuchung wurde Versuchsteilnehmern Lernmaterial vorgelegt, das nicht automatisch ein bestimmtes relevantes Vorwissen aktiviert. So sollten die Versuchsteilnehmer etwa Personen klassifizieren, die mit Hilfe von mehreren Verhaltensmerkmalen beschrieben wurden (z.B. „ging nie nach Einbruch der Dunkelheit in den Park“). Die Personenkategorien waren so konstruiert, dass sie entweder linear trennbar oder nicht-linear trennbar waren. Es zeigte sich, dass ohne weitere Instruktion die nicht-linearen Strukturen etwas leichter zu lernen waren. Wurde den Versuchsteilnehmern allerdings gesagt, bei einer der beiden Klassen handle es sich um Personen, die einen bestimmten Charakterzug hatten (z.B. „Ehrlichkeit“), dann war die linear trennbare Struktur leichter zu lernen. Diese Struktur war so konstruiert, dass sich alle Merkmale in einer Kategorie häuften, die typisch waren für den erwähnten Charakterzug. Das Aufsummieren von Evidenz ist eine Strategie, die zu linear trennbaren Kategorien führt. In diesem Experiment führte das Vorwissen also dazu,

dass zunächst unverbunden erscheinende Merkmale einen gemeinsamen thematischen oder funktionalen Bezug erhielten.

Nicht jedes Wissen ist so strukturiert, dass es das Aufsummieren von Evidenz (und damit lineare Trennbarkeit) nahe legt. In einem weiteren Experiment haben Wattenmaker et al. (1986) Versuchsteilnehmern wiederum Beschreibungen unterschiedlicher Personen vorgelegt. Als alternative Merkmale wurden benutzt: „arbeitet gerne innen“ versus „arbeitet gerne außen“, „arbeitet in einer kleinen Gruppe“ versus „arbeitet in einer großen Gruppe“, „arbeitet das ganze Jahr“ versus „arbeitet nicht im Winter“. Die Aufgabe bestand darin zu lernen, welche der Personen Maler und welche Bauarbeiter sind. Das erste Merkmal jedes Merkmalspaars ist typisch für Maler, das zweite ist typisch für Bauarbeiter. Linear trennbare Kategorien, die in den beiden Klassen die typischen Merkmale der beiden Berufe enthielten, sollten deshalb vergleichsweise leicht zu lernen sein. Dies war auch einer der in diesem Experiment erzielten Befunde. Das Muster kehrte sich allerdings um, wenn man den Versuchsteilnehmern den Hinweis gab, dass es zwei Klassen von Malern gibt: Maler, die im Inneren arbeiten und solche, die Außenarbeiten durchführen. Dieser Hinweis sensitivierte die Versuchsteilnehmer auf Korrelationen zwischen dem ersten und dritten Merkmal. Maler, die im Inneren arbeiten, haben typischerweise die beiden Merkmale „arbeitet gerne innen“ und „arbeitet das ganze Jahr“, während die andere Klasse eher die Merkmale „arbeitet gerne außen“ und „arbeitet nicht im Winter“ hat. Für Versuchsteilnehmer, die den Hinweis auf die beiden Klassen von Malern erhielten, war nun die nicht-linear trennbare Struktur einfacher zu lernen, da diese Struktur die nahe gelegte korrelative Beziehung zwischen den beiden Merkmalen verkörperte (vgl. auch Pazzani, 1991).

Das eben geschilderte Experiment hat gezeigt, dass die relative Schwierigkeit verschiedener Kategorienstrukturen mit dem Vorwissen interagiert. Wattenmaker (1995) hat nun die Hypothese formuliert, dass es allgemeine bereichsspezifische Unterschiede der Kategorienstruktur gibt, die mit stabilen Strukturunterschieden zwischen verschiedenen Domänen zu tun hat. Er hat eine Reihe von Experimenten durchgeführt, die nahe legen, dass soziale Kategorien wie etwa Personenkategorien eher

Theorien beinhalten, die mit linear trennbaren Kategorien verträglich sind. Kategorien wie „Wissenschaftler“ oder „Extravertierter“ sind häufig dadurch charakterisiert, dass man charakteristische Merkmale aufsummiert und umso eher die Kategorie anwendet, je mehr dieser Merkmale vorliegen. Anders sieht es aus in Objektdomänen. Eine Kategorie wie „Fahrzeug“ ist eher durch komplexe Relationen zwischen den Merkmalen charakterisiert, was dazu führt, dass lineare Trennbarkeit bei diesen Kategorien weniger häufig zu finden ist. Wattenmaker geht also davon aus, dass wir allgemeines Wissen über strukturelle Domänenunterschiede haben, die wir beim Lernen nutzen.

((3)) 3b-2.3.1 Probleme der Theorien-Sicht

Ein grundlegendes Defizit aktueller Untersuchungen zur Theorien-Sicht ist, dass nur ansatzweise klar ist, wie die zugrunde liegenden Annahmen in eine formal präzierte Theorie zu übersetzen sind, die sich empirisch testen lässt. Im Gegensatz zu den formal stringenten Theorien ähnlichkeitsbasierter Ansätze liegen in diesem Bereich häufig nur qualitative Überlegungen vor (vgl. aber Abschnitt 5 für erste Versuche in Richtung einer Theoriebildung).

Eine zweite offene Frage ist, wie man sich den Erwerb des Vorwissens vorstellen soll, das die Kategorisierungsprozesse beeinflusst. Ist es nicht die plausibelste Annahme, dass am Anfang des Lernprozesses ähnlichkeitsbasierte Konzepte stehen, die dann die Basis für komplexere theoretische Strukturen bilden? Manche Forscher sind deshalb der Meinung, dass die Theorien-Sicht vorschnell die Bedeutung ähnlichkeitsbasierter oder wahrnehmungsnaher Kategorien in Frage gestellt hat (vgl. Goldstone, 1994a; Smith & Samuelson, 1997). Es gibt aber auch Ansätze, vorwiegend im Bereich der Entwicklungspsychologie (z.B. Carey, 1985; Gopnik & Meltzoff, 1997; Keil, 1989), die davon ausgehen, dass die Annahme plausibler ist, dass quasi-theoretisches Wissen den Lernprozess von Anfang an in unserer Ontogenese beeinflusst (vgl. Textbox *Domänenspezifität von Kategorien: Entwicklungspsychologische Evidenz*).

Schließlich wurde auch kritisch eingewandt, dass das Ähnlichkeitskonzept im Rahmen der Theorien-Sicht vorschnell aufgegeben wurde (Goldstone, 1994a; Hahn & Chater,

1997). Es wird allerdings gefordert, dass man die klassischen, auf einfachen unabhängigen Merkmalen basierten Ähnlichkeitstheorien zumindest erweitern muss. Die Studien zur Theorien-Sicht haben gezeigt, dass relationale Merkmale eine wichtige Rolle bei der Repräsentation von Konzepten spielen, die bei der Bestimmung von Ähnlichkeit eine große Bedeutung haben (vgl. Medin et al., 1993). Manche Konzepte wie „Barriere“, „Falle“ oder „Sieb“ sind geradezu dadurch charakterisiert, dass der Kern ihrer Bedeutung durch Relationen zu anderen Objekten charakterisiert ist (vgl. Barr & Caplan, 1987; Goldstone, 1996).

In manchen Untersuchungen wurde gezeigt, dass Relationen beim Ähnlichkeitsvergleich gelegentlich wichtiger sein können als die Merkmale, die durch diese Relationen verknüpft werden. Dies bringt den Ähnlichkeitsbegriff in Richtung des Konzepts von Analogien (vgl. auch das Kapitel von Knoblich in diesem Band). Und tatsächlich finden sich in der Literatur Vorschläge, Theorien des analogen Transfers auf Konzepte anzuwenden (vgl. Gentner & Markman, 1997).

((1)) 3b-3 ARTEN VON KATEGORIEN

Ein auffälliges Merkmal ähnlichkeitsbasierter Theorien ist, dass sie sich in der Regel nicht darum bemühen, zwischen unterschiedlichen Typen von Kategorien zu unterscheiden. Die meisten Theorien, die im Rahmen dieses Ansatzes entwickelt wurden, gehen davon aus, dass es für die Modellierung von Kategorisierungsprozessen gleichgültig ist, worauf sich Kategorien beziehen. Konsequenterweise wurden häufig artifizielle Stimuli als Untersuchungsmaterial verwendet. Im Rahmen der Theorien-Sicht hingegen wurde zunehmend Interesse daran entwickelt, sich unterschiedliche Kategorien, die wir in unserem Alltag verwenden, genauer anzusehen, um die Frage zu untersuchen, ob verschiedene Inhaltsbereiche auch einhergehen mit Unterschieden in der Struktur und der Verarbeitung von Kategorien. In diesem Kapitel werden nur einige wichtige Beispiele für Typen von Kategorien diskutiert (vgl. Medin, Lynch & Solomon, 2000, für einen Überblick).

((2)) 3b-3.1 Natürliche Arten vs. Artefakte

Eine wichtige Unterscheidung, die im Rahmen der Theorien-Sicht gemacht wurde, ist die zwischen natürlichen Arten („natural kinds“), die sich auf in der Natur vorfindliche Objekte wie Tiere, Mineralien, Pflanzen, Flüssigkeiten usw. beziehen und Artefakten, die Objekte bezeichnen, die von Menschen geschaffen wurden (z.B. Auto, Regenschirm, Vase). Eine große Zahl von Untersuchungen hat sich mit der Frage auseinander gesetzt, ob diese beiden Objektbereiche unterschiedlich repräsentiert werden.

Natürliche Arten. Viele Untersuchungen zu natürlichen Arten wurden von der Position des *psychologischen Essentialismus* (Medin & Ortony, 1989) inspiriert, die davon ausgeht, dass Menschen sich so verhalten, als hätten die Dinge, die sie umgeben, eine tiefer liegende Essenz, die sie zu den Dingen macht, die sie sind (vgl. auch Putnam, 1975). Unabhängig davon, ob der Essentialismus eine wissenschaftlich haltbare Position ist, lässt sich zeigen, dass Menschen in vielen Kulturen daran glauben, dass es tiefer liegende wesentliche Eigenschaften gibt, die die wahrnehmbaren Oberflächenmerkmale beeinflussen (Atran, 1990, 1998). In unserer Kultur werden biologische Kategorien (z.B. Geschlecht) mit genetischen Merkmalen zusammengebracht oder Objekte mit chemischen Eigenschaften, die stärker als Oberflächenmerkmale über Kategorienzugehörigkeit entscheiden.

Belege für psychologischen Essentialismus wurden mit Studien beigebracht, die die Versuchsteilnehmer mit fiktiven Objekttransformationen konfrontierten. So konnte Rips (1989) zeigen, dass Probanden einen fiktiven Vogel, der durch Umweltgifte bedingt das Aussehen eines Insekts angenommen hat, zwar als dem Insekt ähnlicher einschätzen, ihn aber immer noch als Vogel bezeichnen würden. Ähnlich konnte Keil (1989) zeigen, dass ein Waschbär, dessen Äußeres so verändert wurde, dass er wie ein Stinktier aussieht, dennoch weiterhin als Waschbär klassifiziert wird. Die Versuchsteilnehmer glaubten anscheinend, dass die inneren Merkmale wichtiger wären als das Äußere des Tieres. Schließlich lässt sich auch die bereits erwähnte Studie von Gelman und Markman (1986) im Sinne eines psychologischen Essentialismus interpretieren, in der gezeigt wurde, dass bereits 4-Jährige Inferenzen eher auf der Basis der

Kategorienbezeichnung und nicht auf der der äußerlichen Ähnlichkeit machen (vgl. Abb. 3b-13).

Obwohl diese Studien eindeutig die Bedeutung tiefer liegender Merkmale belegen, ist die Frage, welche Rolle daneben die Oberflächenmerkmale spielen und wie homogen die angenommene Essenz ist, eine aktuelle Forschungsfrage (vgl. Rips, 2001). Malt (1994) ist beispielsweise der Frage nachgegangen, welche Flüssigkeiten Amerikaner im Alltag als Wasser bezeichnen, als wie typisch die unterschiedlichen Exemplare eingeschätzt werden, und ob die Befunde durch eine essentialistische Heuristik, die Wasser mit H_2O in Verbindung bringt, erklärbar sind. Der Hauptbefund dieser Studie ist, dass diese unterschiedlichen Maße für Kategorienzugehörigkeit nicht konvergieren. Tee wird beispielsweise nicht als Wasser bezeichnet, obgleich dieses Getränk von den Probanden mit einem höheren H_2O -Gehalt in Verbindung gebracht wird als Meerwasser, das als Wasser bezeichnet wird. Umgekehrt wird Mineralwasser, dem ein hoher Gehalt an H_2O zugeschrieben wird, eher als untypisches Wasser eingeschätzt. Diese Studie zeigt, dass andere Faktoren wie der Gebrauch, der typische Ort und die Quelle der Flüssigkeit einen ebenso wichtigen Einfluss auf die Kategorisierung haben wie die chemische Zusammensetzung.

Artefakte. Im Gegensatz zu natürlichen Arten werden Artefakte weniger mit einer inneren Essenz in Verbindung gebracht, sondern eher mit funktionalen Merkmalen. Ein Stuhl beispielsweise ist eher durch seine Funktion charakterisiert, eine Sitzgelegenheit zu schaffen, als durch die innere Zusammensetzung seines Materials. So konnten Barton und Komatsu (1989) beispielsweise zeigen, dass die Veränderung der molekularen Struktur (eine Ziege mit veränderten Chromosomen vs. ein Reifen, der nicht aus Gummi gemacht ist) eher die Kategorisierung der natürlichen Arten veränderte als die der Artefakte, während das Umgekehrte für funktionale Merkmale galt (eine Ziege, die keine Milch mehr gibt vs. ein Reifen, der nicht mehr rollen kann).

In den letzten Jahren gab es eine Debatte darüber, ob die objektive Funktion von Artefakten tatsächlich das entscheidende, charakterisierende Merkmal ist. So wurde vorgeschlagen, dass nicht die objektive Funktion, sondern die vom Macher des Objekts

intendierte Funktion entscheidend ist. Ein kaputter Stuhl wird immer noch als Stuhl bezeichnet und ein Objekt, das wie ein Regenschirm aussieht, aber von der Firma als Lampenschirm gedacht ist, wurde in Experimenten von Rips (1989) eher als Lampenschirm gesehen. Diese Theorie erklärt auch, warum wir Fußböden nicht als Stuhl oder Schatten als Uhren bezeichnen, obwohl diese die entsprechenden Funktionen unter Umständen erfüllen können.

Aber selbst diese Charakterisierung reicht nicht aus. Malt und Johnson (1992) haben gefunden, dass neben der intendierten Funktion auch die physikalischen Eigenschaften des Objekts entscheidend sind. So wird ein Gegenstand, der wie ein Boot aussieht und als Boot funktioniert, auch als Boot klassifiziert, selbst wenn man sagt, dass es nicht als solches vom Erbauer intendiert wurde. Umgekehrt wird eine Gummikugel, die an einer Gruppe von Delphinen befestigt wird, nicht als Boot angesehen, selbst wenn diese Anordnung diese Funktion erfüllt. Es sieht also so aus, dass bei der Klassifikation die Relation der intendierten Funktion zu dem Design des Objekts entscheidend ist (vgl. Bloom, 2000). Ein Objekt wird dann als bestimmtes Artefakt kategorisiert, wenn sich sein Design am besten dadurch erklären lässt, dass man dem Erschaffer des Objekts eine bestimmte Intention unterstellt. Dies ist auch eine Form von Essentialismus, aber ein Essentialismus, der sich nicht auf materielle Eigenschaften, sondern auf inferierte Intentionen bezieht.

((2)) 3b-3.2 Kausale Kategorien

Eine andere Möglichkeit, die Unterschiede zwischen verschiedenen Kategorien wie natürlichen Arten und Artefakten zu begründen, besteht darin, dass man von Unterschieden in der zugrunde liegenden kausalen Struktur ausgeht. So hat Ahn (1998) die Hypothese formuliert, dass die Zentralität von Kategorienmerkmalen von ihrem kausalen Status abhängt (vgl. Ahn & Kim, 2001, für einen Überblick). Ahn geht davon aus, dass generell Merkmale, die Ursachen bezeichnen, eine größere Bedeutung haben als Merkmale, die Effekte bezeichnen. Zur Stützung dieser Hypothese haben Ahn, Kim, Lassaline und Dennis (2000) Probanden Kategorienexemplare mit drei Merkmalen vorgelegt (z.B. Tiere, die Früchte essen, klebrige Füße haben und Nester auf Bäumen bauen). Einem Teil der Versuchsteilnehmer wurde mitgeteilt, dass die drei Merkmale

eine kausale Kette bilden. So wurde ihnen gesagt, dass das Essen von Früchten zu klebrigen Füßen führt, was es den Tieren erleichtert, Nester auf Bäumen zu bauen. Die Ergebnisse der Studie zeigen, dass die Probanden neue Exemplare von Tieren, die nur zwei der drei kritischen Merkmale haben, eher als zu der Kategorie zugehörig einschätzten, wenn der terminale Effekt der Kette fehlt als wenn die initiale Ursache fehlt, wobei die Befunde zum Fehlen des mittleren Effekts dazwischen lagen.

Ahn (1998) hat zeigen können, dass der kausale Status auch eine wichtige Rolle bei natürlichen Arten und Artefakten spielen könnte. Sie konnte an Hand von artifiziellem Material demonstrieren, dass sowohl bei Artefakten als auch bei natürlichen Arten weniger die Unterscheidung zwischen funktionalen versus kompositionalen Merkmalen entscheidend war, sondern der kausale Status. In Abhängigkeit davon, welches der beiden Merkmale als Ursache für das andere eingeführt wurde, variierte der Einfluss der beiden Merkmalstypen auf die Kategorisierung – und zwar unabhängig von der Art der Kategorie. Umstritten ist allerdings die Generalität der kausalen Status-Hypothese. Rehder und Hastie (2001) haben eine Studie vorgelegt, die gezeigt hat, dass bei Kategorien, die multiple Ursachen eines gemeinsamen Effekts verknüpfen (wie etwa bei vielen Krankheiten), der Effekt zentraler ist als die alternativen Ursachen.

Der Einfluss von Kausalwissen wirkt sich auch auf andere Maße der Kategorienstruktur aus. So haben Medin, Wattenmaker und Hampson (1987) ihren Versuchsteilnehmern Exemplare von Begriffen vorgelegt, deren Merkmale in probabilistischer Beziehung zu zwei Kategorien standen. Die Aufgabe bestand darin, die Exemplare nach Belieben in zwei Kategorien zu sortieren. Dabei zeigte sich, dass die Versuchsteilnehmer eher auf der Basis korrelierter Symptome sortierten, wenn die Symptome kausal verknüpft werden konnten (z.B. Ohrenschmerzen—Schwindel), als wenn es sich um nicht kausal aufeinander beziehbare Symptome handelte (z.B. Halsschmerzen—Hautausschlag).

Schließlich konnte auch gezeigt werden, dass die kausale Struktur von Kategorien den Lernprozess beeinflusst. Murphy und Allopenna (1994) ließen ihre Probanden Kategorien von neuen Tieren lernen, die Merkmale hatten, die entweder kausal (z.B. scharfe Zähne – Fleischesser) oder nicht kausal (spitze Ohren – Flecken) verknüpft

waren. Die Ergebnisse der Studie zeigen, dass Kategorien mit kausal verknüpften Merkmalen schneller gelernt wurden als solche mit kausal unverbundenen Merkmalen und dass das Vorliegen kausal verbundener Merkmale einen großen Einfluss auf Typikalitätseinschätzungen hatte, selbst wenn die Merkmale eher selten vorkamen (vgl. auch Pazzani, 1991).

Abb. 3b-15

Waldmann, Holyoak und Fratianne (1995) gingen einen Schritt weiter. Während etwa die Studie von Murphy und Allopenna (1994) und ähnliche Untersuchungen den Einfluss von bereits vorhandenem bereichsspezifischen Wissen belegten, interessierten sich Waldmann und Mitarbeiter dafür, ob auch *abstraktes* Wissen über die Kausalstruktur den Lernprozess beeinflussen kann (vgl. auch Waldmann & Holyoak, 1992; Waldmann, 1996a, 1997, 2000, 2001). In diesen Experimenten erhielten die Versuchsteilnehmer generell identische Lernexemplare mit identischen Merkmalen und identischen Kategorien. Der einzige Unterschied bestand darin, ob die Merkmale als Ursache der Kategorien (Gemeinsamer Effekt-Modell) oder als Effekte der Kategorien (Gemeinsame Ursache-Modell) beschrieben wurden (vgl. Abb. 3b-15).

Ein konkretes Beispiel dafür sind etwa die Untersuchungen von Waldmann (2000), in denen die Probanden Informationen über das Vorhandensein oder die Abwesenheit dreier Substanzen im Blut einzelner Patienten erhielten. Die Aufgabe bestand darin, die Patienten in solche zu klassifizieren, die an einer neuartigen Blutkrankheit („Midosis“) litten oder nicht erkrankt waren. Manipuliert wurde in den eingangs gegebenen Untersuchungen, ob die Substanzen aus der Nahrung kamen und damit potentielle Ursachen der Erkrankung waren (Abb. 3b-15A) oder ob gesagt wurde, dass die Substanzen von der Krankheit im Blut produziert würden, also Effekte seien (Abb. 3b-15B). Da weder die Substanzen näher charakterisiert noch die Krankheit bekannt waren, hatten die Probanden kein Vorwissen über die Aufgabe. Der einzige Unterschied zwischen den Bedingungen bestand in der abstrakten Struktur des instruierten Kausalmodells.

In einer Reihe von Studien konnte gezeigt werden, dass identische Lernaufgaben zu unterschiedlichem Lernen führen in Abhängigkeit von der eingangs instruierten Kausalstruktur. So haben Waldmann et al. (1995) mit Merkmalen, die kontinuierlich variierten, gezeigt, dass die Probanden leichter eine nicht-linear trennbare Struktur mit korrelierten Merkmalen lernten, wenn sie dachten, die Merkmale repräsentierten Effekte; umgekehrt war die linear trennbare Struktur leichter, wenn die Lernenden dachten, die Merkmale wären Ursachen. Dieser Befund lässt sich aus der Struktur der kontrastierten Kausalmodelle vorhersagen. Während Gemeinsame Ursache-Modelle eine Korrelation der durch die gemeinsame Ursache bedingten unabhängigen Effekte implizieren, drücken Gemeinsamer Effekt-Modelle eher die Annahme aus, dass die alternativen Ursachen unabhängig voneinander den Effekt beeinflussen.

Andere Experimente zum Erwerb kausaler Kategorien (z.B. der oben beschriebenen fiktiven Krankheit) haben weitere Belege für den Einfluss von Kausalmodellen erbracht. So konnte gezeigt werden, dass die Versuchsteilnehmer davon ausgehen, dass mehrere Ursachen im Wettbewerb in Bezug auf die Erklärung des Effekts stehen, während mehrere Effekte einer gemeinsamen Ursache nicht miteinander im Wettbewerb in Bezug auf ihre diagnostische Aussagekraft sind (Waldmann & Holyoak, 1992; Waldmann, 2000, 2001). Diese Abhängigkeit der Kompetition der Kategorienmerkmale vom kausalen Status steht im Widerspruch zu den Annahmen assoziationalistischer und konnektionistischer Lerntheorien, die in der Regel einen kompetitiven Lernmechanismus unabhängig von der Art der Merkmale annehmen (für eine nähere Erläuterung von kompetitivem Lernen, vgl. Abschnitt 5). Zusammen genommen belegen diese Untersuchungen, dass nicht nur bereichsspezifisches Wissen den Lernprozess beeinflussen kann, sondern auch vergleichsweise abstraktes Wissen über die Struktur von Kausalmodellen (vgl. auch Rips, 2001).

((2)) 3b-3.3 Andere Arten von Kategorien

Die Kategorienforschung hat viele Jahre so getan, als seien Objektkategorien, die sich durch Nomen benennen lassen, die einzige Art von Kategorien. In den letzten Jahren kommen aber zunehmend auch andere Kategorientypen in den Blick, die andere Theorien erfordern. So haben Gentner und Mitarbeiter (vgl. Gentner, 1981) ihr

Augenmerk auf die Unterscheidung zwischen Nomen und Verben gelegt. Während Nomen eher Cluster von korrelierten Merkmalen bezeichnen, die zu perzeptuellen Chunks führen, beziehen sich Verbkategorien eher auf zeitlich ausgedehnte Ereignisse, die kausale Relationen zwischen Objekten, Aktivität oder Veränderungen kodieren (vgl. auch Kersten, 1998; Zaki & Homa, 1999).

Ereignisbezogene Relationen wurden auch in der Arbeitsgruppe von Klix und van der Meer untersucht (Klix, van der Meer, Preuß & Wolf, 1987; van der Meer, 1986). In diesen Studien wurde gezeigt, dass zwischenbegriffliche Relationen häufig als „Geschehenstypen“ in Form ereignisbezogener Netzwerke repräsentiert werden. Der Geschehenstyp Schule wird etwa als Netzwerk konzeptualisiert, in denen die Konzepte „Lehrer“, „Schüler“, „Unterricht“, „Wissen“, „Pause“ usw. durch ereignisbezogene Relationen verknüpft sind (vgl. auch Barsalou & Sewell, 1985; Waldmann, 1990). Lin und Murphy (2001) haben ebenfalls in einer groß angelegten Studie gezeigt, dass die herkömmliche Sicht, dass Kategorien bei Erwachsenen vorzugsweise als taxonomische Hierarchien abgespeichert sind, falsch ist. In einer Serie von zehn Experimenten konnten sie zeigen, dass thematische, ereignisbezogene Relationen (z.B. Hund-Knochen) einen deutlichen Einfluss auf Kategorisierungsverhalten (z.B. Sortieren, induktives Denken, Kategorisierungsentscheidungen) haben. Bei vielen Probanden hatten diese Relationen einen stärkeren Einfluss als die entsprechenden taxonomischen Beziehungen (z.B. Hund-Katze).

Ein weiterer Bereich, der bisher kaum untersucht ist, sind abstrakte Kategorien wie z.B. Gerechtigkeit oder Zeit. Hierzu gibt es wenig Forschung. Ein Ansatz postuliert, dass solche Begriffe in Analogie zu konkreten, perzeptiven Kategorien verstanden werden (Barsalou, 1999; Lakoff & Johnson, 1980). So gibt es Belege dafür, dass wir in vielen Fällen das Konzept der Zeit mit Rückgriff auf räumliche Beziehungen verstehen.

Eine wichtige aktuelle Debatte betrifft die Frage, ob Kategorien nach Inhaltsbereichen differenziert werden müssen. Während die traditionellen Ansätze der Kognitionspsychologie davon ausgingen, dass die Repräsentation und Verarbeitung von Kategorien weitgehend unabhängig vom konkreten Inhalt ist, den die Kategorien

bezeichnen, findet sich seit einigen Jahren gehäuft die Position in der Literatur, dass es angeborene bereichsspezifische Constraints gibt, die den Lernprozess steuern. Dieser Ansatz wird v.a. von Entwicklungspsychologen vertreten (vgl. Textbox: *Domänenspezifität von Kategorien: Entwicklungspsychologische Evidenz*). Auch neuropsychologische Untersuchungen liegen zur Frage der Domänenspezifität kategorialen Wissens vor (Textbox: *Domänenspezifität von Kategorien: Entwicklungspsychologische Evidenz*).

((Anfang Textbox))

Domänenspezifität von Kategorien: Entwicklungspsychologische Evidenz

Während sich die Mehrzahl kognitionspsychologischer Theorien der Vorstellung von bereichsübergreifenden Lernmechanismen, die Wissen in ganz unterschiedlichen Bereichen erwerben können, verbunden fühlt, haben sich viele kognitive Entwicklungspsychologen einem anderen Trend in den Kognitionswissenschaften angeschlossen. Basierend auf evolutionspsychologischen Überlegungen wurde die Annahme solcher Allzweck-Mechanismen in Frage gestellt. Das Gehirn wird hier eher in Analogie zu einem Schweizer Taschenmesser als System gesehen, das aus spezialisierten Mechanismen besteht, die auf spezifische Domänen beschränkt sind (Tooby & Cosmides, 1992). Lernen und Konzepterwerb werden demnach nicht durch allgemeine Mechanismen gesteuert, sondern durch bereichsspezifische Prinzipien und Vorannahmen angeleitet (vgl. auch Hirschfeld & Gelman, 1994, für einen Überblick; vgl. auch Textbox: *Kategorien-spezifische Defizite* und den Beitrag von Hoffrage und Vitouch in diesem Band ???).

Ein Kriterium für die Abgrenzung von Domänen ist, ob die jeweiligen Konzepte und Kategorien in unterschiedliche Arten kausaler Erklärungsmuster eingebettet sind. Auf der Basis dieses Kriteriums hat eine Reihe von Psychologen postuliert, dass es angeborenes Wissen über allgemeine Bereiche wie Physik, Biologie oder Psychologie gibt (vgl. Carey, 1985; Gopnik & Meltzoff, 1997; Keil, 1989; Waldmann, 1996b; Wellman & Gelman, 1992). Dieses Wissen ist im Säuglingsalter zunächst nur sehr allgemein und unvollständig, bietet aber einen

konzeptuellen Rahmen für den Erwerb weiteren spezifischeren Wissens. So haben beispielsweise Säuglingsstudien gezeigt, dass bereits nach wenigen Monaten Wissen über manche Eigenschaften von physikalische Objekten nachgewiesen werden kann, wie etwa darüber, dass sich Objekte gegenseitig nicht durchdringen können, oder dass Kontakt bei der Übertragung kausaler Kräfte nötig ist (Baillargeon, 1998; Spelke, Breinlinger, Macomber & Jacobson, 1992; Gopnik & Meltzoff, 1997). Darüber hinaus können Säuglinge Lebewesen von unbelebten Objekten u.a. auf der Basis dessen unterscheiden, dass sich Lebewesen selbständig fortbewegen und sich gegenseitig über eine Distanz beeinflussen können (Spelke, Phillips & Woodward, 1995; vgl. auch Pauen, 2000).

Eine aktuelle Debatte in diesem Forschungsfeld betrifft die Frage, ob in der frühen Kindheit zwischen Psychologie und Biologie unterschieden wird. Während Carey (1985) davon ausgeht, dass junge Kinder biologische Eigenschaften im Rahmen einer intuitiven Psychologie organisieren, bei der alle Induktionen am Menschen als Prototyp ausgerichtet sind, postuliert Keil (1989), dass es neben einer frühkindlichen Psychologie auch eine, wenngleich sehr rudimentäre, Biologie gibt. So konnten Springer und Keil (1989) zeigen, dass Vorschulkinder wissen, dass sich biologische, aber nicht psychologische Merkmale vererben lassen.

Obgleich diese Positionen in der Entwicklungspsychologie eine große Bedeutung haben, gibt es auch eine Reihe von Gegenpositionen, die die Annahme von angeborenem Bereichswissen und das Herunterspielen bereichsübergreifender, auf Ähnlichkeit basierender Mechanismen kritisieren (vgl. z.B. Jones & Smith, 1993; Smith & Samuelson, 1997). So kann in der Regel selbst bei Säuglingsstudien nicht ausgeschlossen werden, dass rudimentäres Wissen in den ersten Monaten gelernt wurde (anstatt angeboren zu sein). Auch die Frage, wie Domänen voneinander abzugrenzen sind, ist angesichts vieler Ähnlichkeiten und Überlappungen der Domänen schwierig. Selbst wenn man von angeborenem Domänenwissen ausgeht, ist noch weitgehend unklar, wie Wissen und Kategorien im Rahmen von Domänen genau erworben werden, welche Form von Transfer

und Interaktion es zwischen Domänen gibt und welche Rolle domänenübergreifende Kompetenzen spielen (wie zum Beispiel kausales, logisches oder mathematisches Denken)(vgl. auch Medin et al., 2000 und den Beitrag von Knoblich in diesem Band).((Ende Textbox))

((Anfang Textbox))

Kategorien-spezifische Defizite

Beobachtungen an Patienten mit selektiven Gehirnschädigungen können Hinweise darauf geben, wie Kategorien repräsentiert werden. Eine wichtige Beobachtung ist, dass es Patienten gibt, die selektive kategorien-spezifische Ausfallerscheinungen haben. So berichteten Warrington und Shallice (1984) über Patienten, die Schwierigkeiten hatten, Lebewesen (Tiere oder Pflanzen) zu identifizieren und zu benennen, während sie kaum Probleme mit unbelebten Dingen (z.B. Artefakte wie Schere, Uhr, Stuhl) hatten (vgl. auch Sartori & Job, 1988). Auch der umgekehrte Fall mit besserer Performanz bei unbelebten Dingen als bei Lebewesen wurde berichtet, wenngleich solche Patienten deutlich seltener zu beobachten sind.

Wie kommt es zu diesen Asymmetrien? Eine Hypothese, die durch evolutionstheoretische Überlegungen motiviert ist, geht davon aus, dass die Fähigkeit, andere Lebewesen (z.B. Feinde oder Nahrungsmittel) schnell und effektiv zu erkennen, eine Möglichkeit darstellt, eine erhöhte Überlebenschance zu garantieren (vgl. auch Textbox: *Domänenspezifität von Kategorien: Entwicklungspsychologische Evidenz*). Dies könnte nach Ansicht einiger Forscher der Grund dafür sein, dass sich in unserem Gehirn in der Evolution möglicherweise neuroanatomische Subsysteme herausgebildet haben, die auf die Erkennung von Lebewesen oder Werkzeuge spezialisiert sind (Caramazza & Shelton, 1998). Diese Position postuliert also eine strikte Domänenspezifität des Kategorisierungssystems mit einer angeborenen Spezialisierung von Gehirnregionen auf inhaltliche Bereiche (vgl. Abb. 3b-16).

Abb. 3b-16

Häufiger wird in der Literatur allerdings die alternative Sicht vertreten, dass unterschiedliche Gehirnbereiche auf verschiedene Merkmalstypen spezialisiert sind. So haben Warrington und Shallice (1984) die Hypothese formuliert, dass sensorische Merkmale getrennt von funktionalen Merkmalen gespeichert werden. Ihrer Ansicht nach sind sensorisch-visuelle Merkmale (Form, Farbe usw.) wichtiger, um belebte Dinge (natürliche Arten) zu identifizieren, während funktionale Merkmale, die den Zweck eines Objekts bezeichnen, eine größere Rolle bei Artefakten spielen. Läsionen in Gehirnbereichen, die visuelle Merkmale kodieren, sollten deshalb zu einem selektiven Defizit für Lebewesen führen. Eine ähnliche Sicht wurde auch von anderen Neuropsychologen vertreten. So geht etwa Damasio (1990) davon aus, dass Artefakte häufiger manipuliert werden als natürliche Arten, weshalb kinästhetische und motorische Merkmale bei der Repräsentation von Artefakten eine größere Rolle spielen sollten als bei Lebewesen.

Eine interessante Methode zur Untersuchung von kategorien-spezifischen Defiziten haben Farah und McClelland (1991) eingesetzt (vgl. Abb. 3b-16). Sie haben ein konnektionistisches Netzwerk entwickelt, das zwei periphere Inputbereiche hat, eines, das verbale Benennungen produziert, und ein zweites, das Objekte visuell repräsentiert. Jedes Objekt und jede Benennung wurde durch eine verteilte Repräsentation auf 24 Knoten kodiert. Diese beiden Bereiche sind nun mit einem semantischen System verbunden, das aus einer Reihe von Knoten besteht, die entweder visuelle (links) oder funktionale (rechts) Merkmale repräsentieren. Beide Bereiche sind aber miteinander verknüpft. Auf der Basis einer empirischen Untersuchung gingen Farah und McClelland dabei von einem 3:1 Verhältnis für visuelle gegenüber funktionalen Merkmalen aus (60 visuelle, 20 funktionale) und von einem größeren Prozentsatz visueller Merkmale bei der Repräsentation von Lebewesen (7,7:1) im Vergleich zu Artefakten (1,4:1). In diesem Modell wurde zunächst trainiert, 20 Objekte zu benennen, indem man die Objekte auf der visuellen Input-Seite (rechts) und die Namen auf der verbalen

Input-Seite (links) als Muster von Aktivierungen nacheinander paarweise enkodierte. Für jedes der vorgegebenen Paare von Objekt und Namen wurden dann die assoziativen Verbindungen zwischen den Knoten nach einer konnektionistischen Lernregel (vgl. Abschnitt 5) so lange modifiziert, bis die 20 Objekte korrekt benannt wurden. Danach haben Farah und McClelland das Netzwerk selektiv „geschädigt“, um den Einfluss von Gehirnläsionen zu simulieren. Das Modell reproduziert eine Reihe von Befunden aus der Literatur. So zeigt das Modell Probleme bei der Identifikation von Lebewesen, wenn das visuelle System geschädigt ist und Probleme bei Artefakten, wenn das funktionale System betroffen ist. Probleme mit der Benennung von Lebewesen waren allerdings deutlich häufiger, weil Artefakte auf der Basis visueller Merkmale selbst dann noch relativ gut identifiziert werden konnten, wenn alle funktionalen Merkmale ausfielen, während die Läsion des visuellen Systems für Lebewesen nicht durch andere Merkmale kompensiert werden konnte.

Auch der Versuch, kategorienspezifische Defizite durch eine Unterscheidung zwischen sensorischen und funktionalen Merkmalen zu erklären, ist nicht unumstritten. So haben Tyler und Moss (2001) ebenfalls ein konnektionistisches Modell entwickelt, das von einem einheitlichen Speicher ausgeht. Unterschiede zwischen Domänen erklären sie durch verschiedene Strukturen der Merkmalsverteilung. Lebewesen zeichnen sich durch eine größere Zahl gemeinsamer, überlappender Merkmale aus als Artefakte. Artefakte hingegen weisen eine größere Zahl an distinkten Merkmalen auf, die sie von anderen Artefakten unterscheiden. Hinzukommt, dass bei Artefakten eine stärkere Korrelation zwischen funktionalen und sensorischen Merkmalen besteht (z.B. Form eines Messers—schneiden) als bei Lebewesen. Diese Unterschiede in der Struktur der Kategorien können selektive Defizite bedingen. So erklärt dieses Modell, warum Patienten Lebewesen häufig noch erkennen, aber nicht benennen können, weil Benennung den Zugriff auf distinkte Merkmale erfordert. Dieses Muster wird bei Artefakten weniger häufig beobachtet, da bei diesen distinkte und funktionale Merkmale stärker korreliert sind. Auch dieses Modell erklärt allerdings nicht alle Befunde. Zur Zeit zeichnet sich die Sicht ab, dass

kategorienspezifische Defizite eine Reihe ganz verschiedener Ursachen haben können, die in komplexer Weise zusammenwirken (vgl. Humphreys & Forde, 2001, für ein vergleichsweise komplexes Modell). ((Ende Textbox))

((1)) 3b-4 RELATIONEN ZWISCHEN KATEGORIEN

Kategorien existieren nicht isoliert, sondern sie stehen in Relationen zu anderen Kategorien. Schwerpunkt der bisherigen Forschung sind taxonomische Beziehungen. Es lassen sich aber auch andere Strukturen auf der Basis anderer Relationstypen postulieren. So können Begriffe kausal oder funktional verknüpft sein (vgl. Abschnitt 3), sie können in Teil-Ganzes-Beziehungen stehen oder Bestandteil typischer Ereignisfolgen sein (vgl. Hoffmann, 1993). Im Folgenden werden die am häufigsten untersuchten Strukturen diskutiert werden (vgl. auch Medin et al., 2000; Medin et al., 2001).

((2)) 3b-4.1 Taxonomien

Häufig sind Kategorien taxonomisch auf unterschiedlichen Abstraktionsebenen angeordnet. Auf jeder Ebene gehört ein Exemplar zu einer aus einer Reihe alternativer, wechselseitig sich ausschließender Kategorien (vgl. Abb. 3b-17). So kann ein Pudel als Hund, Säugetier und Lebewesen kategorisiert werden. Dies bedeutet gleichzeitig, dass der Hund kein Reptil oder keine Pflanze ist. Zwischen den Ebenen besteht Klasseninklusion. So sind alle Typen von Hunden Säugetiere, aber nicht alle Säugetiere Hunde.

Abb. 3b-17

Eine wichtige Beobachtung ist, dass es eine psychologisch privilegierte Ebene der Taxonomien gibt, bei der uns eine Reihe von Aufgaben am leichtesten fällt. Diese Ebene ist die mittlere Beschreibungsebene, die als Ebene der *Basis-Kategorien* bezeichnet wurde (Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976). Sehen wir beispielsweise ein bestimmtes vierbeiniges Tier, dann tendieren wir dazu, es primär als Hund und nicht als Tier oder als Pudel zu bezeichnen. Hund ist eine Kategorie auf der

Ebene der Basis-Kategorien, während Tier eine übergeordnete und Pudel eine untergeordnete Kategorie ist.

Eine Reihe von Studien hat die psychologische Sonderstellung der Basis-Kategorien belegt (vgl. Hoffmann, 1986; 1993; Rosch et al., 1976). So werden Objekte primär auf der Basis-Ebene benannt. Sollen Probanden entscheiden, ob ein bestimmtes Objekt zu einer Kategorie gehört, dann gelingt ihnen dies in der Regel am schnellsten, wenn Basis-Kategorien verwendet werden. Basis-Kategorien sind auch die ersten Kategorien, die in der Entwicklung gelernt werden (vgl. Mervis & Crisafi, 1982; Weinert & Waldmann, 1988).

Es gibt unterschiedliche Annahmen darüber, was Basis-Kategorien auszeichnet. Rosch et al. (1976) gehen davon aus, dass Basis-Kategorien eine besonders hohe Familienähnlichkeit aufweisen. So sind sich Hunde im Vergleich zu Tieren untereinander vergleichsweise ähnlich, während der Ähnlichkeitszuwachs, der auf der untergeordneten Ebene (Pudel) erreicht wird, relativ gering ist. Andere haben darauf hingewiesen, dass auf dieser Ebene die Vorhersagbarkeit zwischen Merkmalen und Kategorien sowohl in die Merkmals-Kategorien-Richtung („cue validity“) als auch in die Kategorien-Merkmals-Richtung („category validity“) simultan besonders groß ist (vgl. Corter & Gluck, 1992). Schließlich wurde darauf hingewiesen, dass Basis-Kategorien die höchste Ebene von Objekten darstellen, die auf der globalen Ebene gleichartig aussehen (Hoffmann & Ziebler, 1982, 1983), die relativ viele gleichartige Teile aufweisen (z.B. Autos, aber nicht alle Fahrzeuge haben Räder)(Tversky & Hemenway, 1984) und die gleichartigen Zwecken dienen (Murphy & Wisniewski, 1989).

Es gibt allerdings einige Evidenz dafür, dass die Basis-Ebene mit dem Ausmaß von Expertise in einem Bereich variieren kann. So sind Hunde die Basis-Ebene für die meisten Menschen, während sich für Experten in einem Bereich die Basis-Ebene auf die untergeordnete Ebene (z.B. Pudel) verschiebt (Tanaka & Taylor, 1991). Die Basis-Ebene scheint also auch vom Vorwissen der Kategorisierer beeinflusst zu werden. Diese Hypothese ist auch konsistent mit einer Reihe von interkulturellen Studien. Während

Rosch et al. (1976) für amerikanische Studenten gefunden haben, dass die Basis-Ebene auf der Ebene von Kategorien wie Vogel, Hund oder Baum liegt, haben Anthropologen die Entdeckung gemacht, dass in anderen Kulturen (z.B. die Itzaj Mayas) häufig die für uns untergeordnete Ebene (Rotkehlchen, Pudel, Eiche) die primäre Ebene der Kategorisierung ist (Atran, 1990; Berlin, Breedlove & Raven, 1973). Eine mögliche Erklärung für diese Divergenz mag sein, dass Menschen, die wie die Itzaj Mayas im Regenwald von Guatemala leben, Biologieexperten sind im Vergleich zu amerikanischen Studierenden.

Eine interessante Inkonsistenz zwischen diesen Überlegungen über den Einfluss von Expertise auf die Basisebene zeigt sich aber, wenn man ein anderes Maß für die Bestimmung der Basisebene verwendet. Coley, Medin und Atran (1997) haben induktive Argumente von der Sorte „Alle Forellen haben Enzym X. Wie wahrscheinlich ist es, dass alle Fische Enzym X haben?“ untersucht (vgl. auch Abschnitt 6). Die Frage war, welche Ebene induktiv privilegiert wird. Für amerikanische Versuchsteilnehmer wurde beispielsweise vorhergesagt, dass die Kategorie der Fische die Basis-Ebene darstellt, für die die Probanden davon ausgehen sollten, dass sie viele Merkmale mit untergeordneten Kategorien teilt. Sie sollten also relativ zuversichtlich beim Schluss von Forellen auf Fische sein, diese Zuversicht sollte aber überproportional abfallen beim Schluss von der untergeordneten Ebene (Forelle) auf die vergleichsweise heterogene übergeordnete Ebene (Tiere). Bei den Itzaj Mayas hingegen erwartete man, dass die untergeordnete Ebene (z.B. Forelle) privilegiert sein sollte, wo für diese Population die Basisebene liegt. Für diese Gruppe wurde erwartet, dass sie zuversichtlich sein sollten, wenn es um die Generalisierung von Merkmalen von speziellen Forellentypen auf Forellen im Allgemeinen geht, aber ihre Zuversicht überproportional stark abfallen sollte bei Schlüssen auf die übergeordnete Kategorie der Fische. Im Widerspruch zu den Erwartungen fand sich allerdings, dass für beide Personengruppen die untergeordnete Ebene eine induktive Sonderstellung hatte. Auch amerikanische Studierende hatten überproportional mehr Zutrauen zu einem Schluss von speziellen Forellen auf alle Forellen als auf die Kategorie der Fische.

Dies ist angesichts der Unterschiede der Basis-Ebene überraschend. Eine mögliche Erklärung für die Divergenz zwischen Basis-Ebene und der induktiv privilegierten Ebene bei den Studierenden mag sein, dass diese zwar kein tieferes Wissen über Bäume und Tiere, aber dennoch Metawissen darüber haben, dass unterschiedliche Bäume viele Unterschiede haben dürften, auch wenn die Studierenden diese konkret nicht benennen können. Die Studie macht aber auch deutlich, dass Eigenschaften wie der Ort der Basis-Ebene nicht aufgabeninvariant sind, ein Thema, das ausführlicher in Abschnitt 6 angesprochen werden wird.

((2)) 3b-4.2 Nicht-hierarchische Kategorienstrukturen

Viele Kategorien, die wir benutzen, lassen sich nicht in eine taxonomische Hierarchie bringen. Dies gilt vor allem für soziale Kategorien. So kann eine Person als Frau, Liberale, Psychologin und Tennisspielerin gleichzeitig kategorisiert werden, ohne dass es eine Über- oder Unterordnung gibt. Ein solches Nebeneinander möglicher Kategorien nennt man auch Kreuzklassifikationen („cross classifications“). Angesichts dieser potentiellen Kompetition zwischen Kategorisierungen ist die Frage interessant, welche Kategorie wir bevorzugen. Srull und Wyer (1979) haben gezeigt, dass Häufigkeit und Rezenz der Kategorienaktivierung eine wichtige Rolle spielen. So konnten sie zeigen, dass eine Aufgabe, in der eine bestimmte Personenkategorie („feindselig“) nahe gelegt wurde, dazu führte, dass in einer anderen Aufgabe eine Person später eher mit dieser Kategorie in Zusammenhang gebracht wurde, obwohl die beiden Aufgaben vollkommen unverbunden waren und die Informationen über diese Person auch anders hätten interpretiert werden können. Es gibt auch Hinweise darauf, dass manche der Kategorien automatisch, auch gegen den Willen der Versuchsteilnehmer, aktiviert werden (z.B. Rassen- oder Geschlechtsstereotype) (Bargh, 1994; Greenwald & Banaji, 1995) und dass die Aktivierung einer Kategorie andere mögliche konkurrierende Kategorien hemmt (Macrae, Milne & Bodenhausen, 1994)(vgl. Macrae & Bodenhausen, 2000, für eine Übersicht).

Ross und Murphy (1999) sind einem anderen Typ von Kreuzklassifikationen nachgegangen, der Konkurrenz zwischen taxonomischen und ereignisbezogenen Kategorien. So lassen sich Nahrungsmittel taxonomischen (z.B. Getränke, Früchte,

Gemüse) als auch ereignisbezogenen Script-Kategorien (z.B. Desserts, Snacks, Appetizer, Frühstück) zuweisen. Die Studie zeigte, dass beide Kategorientypen generiert und genutzt wurden, wobei allerdings die taxonomischen Kategorien häufiger aktiviert wurden als die Script-Kategorien. Beide Kategorientypen wurden auch für induktive Schlüsse genutzt, wobei die Nutzung des jeweiligen Kategorientyps vom Typ der Inferenz abhing (vgl. Abschnitt 6).

Die hier näher besprochenen Kategorienrelationen stellen nur zwei Beispiele aus der großen Zahl möglicher Relationen dar. Eine offene Frage dabei ist, wie man sich die Repräsentation und den Zugriff auf eine Wissensbasis vorstellen sollte, die alle diese gleichzeitig möglichen Relationen beinhaltet.

((1)) 3b-5 DER ERWERB VON KATEGORIEN

Bisher haben wir uns mit der Repräsentation und Nutzung bereits vorhandener Kategorien befasst. Offen bleibt dabei, wie diese Kategorien gelernt werden. Dazu wurde in den letzten Jahren eine Reihe von Theorien entwickelt, die hier nur paradigmatisch diskutiert werden können (vgl. auch Medin & Heit, 1999).

((2)) 3b-5.1 Konnektionistische Modelle

Die meisten Theorien des Erwerbs von Kategorien basieren auf der Konzeption ähnlichkeitsbasierter Theorien, die den Lernprozess unabhängig von Vorwissen modellieren (vgl. auch Heydemann, 1998). Dabei dominieren in den letzten Jahren konnektionistische Modelle, die von assoziativen Lernprozessen ausgehen, wie man sie auch bei basalen Lernprozessen des klassischen und instrumentellen Konditionierens unterstellt (vgl. aber Anderson, 1991).

Abb. 3b-18

Abbildung 3b-18 verdeutlicht einige Alternativen konnektionistischer Modelle (vgl. auch Nosofsky, Kruschke & McKinley, 1992): Modell I entspricht einer verallgemeinerten Form von Gluck und Bowers (1988a, b) Komponenten-

Merkmalsmodell (*component-cue network model*), wie es von Nosofsky et al. (1992) genannt wurde.

Auf der linken Seite befindet sich die Ebene der Input-Knoten, die die Merkmale des Lern-Inputs kodieren (A, B, C). Dieses Modell sieht nur Kodierungen der einfachen Merkmale vor. Auf der rechten Seite befinden sich die Output-Knoten R1 und R2, die beispielsweise zwei unterschiedliche Kategorien kodieren können. Schließen diese Kategorien sich gegenseitig aus, kann man eine binäre Kategorisierungsentscheidung auch mit einem Knoten modellieren. Das Netzwerk reduziert sich dann zu einer assoziativen Struktur, die mit dem Standardfall der Rescorla-Wagner-Theorie (1972) korrespondiert (vgl. auch den Beitrag von Koch in diesem Band). Ein Beispiel für eine Aufgabe sind Krankheitsklassifikationen, bei denen die Input-Knoten die jeweiligen Symptome eines Patienten kodieren (z.B. Fieber, Übelkeit, Schnupfen) und der Output die zu diagnostizierende Krankheit (z.B. Grippe).

Die Assoziationsgewichte (Linien), die Input und Output verknüpfen, werden in diesen Modellen mit der Delta-Regel gelernt, die formal der Rescorla-Wagner-Regel (1972) entspricht. Gemäß dieser Regel wird der Output-Knoten bei jedem Lerndurchgang auf der Grundlage der Summe der Assoziationsgewichte der vorhandenen Merkmale, also beispielsweise der bei dem jeweiligen Patienten vorliegenden Symptome, aktiviert. Fehlerfrei ist der Lerndurchgang, wenn die Aktivierung des Output-Knoten 1 beträgt, wenn die Rückmeldung ergibt, dass die Krankheit bei dem Patienten tatsächlich vorliegt, und wenn die Aktivierung dieses Knotens 0 beträgt, wenn der Patient gemäß der Rückmeldung gesund ist. Fehlerfrei ist die Vorhersage also immer dann, wenn der durch die Symptome vorhergesagte Wert des Output-Knotens dem in der Rückmeldung vermittelten richtigen Wert entspricht. Weichen die vorhergesagten Werte davon ab, dann werden die Assoziationsgewichte der vorhandenen Symptome in Abhängigkeit von der Größe des Vorhersagefehlers in Richtung Minimierung dieses Fehlers modifiziert. Es handelt sich also um eine Fehlerkorrektur-Regel, die in Abhängigkeit davon Lernprozesse initiiert, wie überraschend die Kategorienrückmeldung (z.B. „der Patient ist gesund“) im Vergleich zur Vorhersage auf der Basis der beobachteten

Symptome ist. Es lässt sich zeigen, dass dieser Lernprozess bei linear trennbaren Prototypenkategorien am Ende des Lernprozesses zu fehlerfreien Klassifikationen führt.

Da die Output-Aktivierung immer auf der Basis der Summe der Assoziationsgewichte der vorhandenen Merkmale vorhergesagt wird, impliziert diese Regel *Kompetition* zwischen Merkmalen um die Vorhersage der Kategorie. Erlaubt eines der Merkmale beispielsweise nahezu perfekte Klassifikationen, dann werden die Gewichte der anderen Merkmale kaum verändert, auch wenn diese individuell sehr prädiktiv für die Kategorie wären (vgl. Abschnitt 3 und Waldmann, 1996a, für eine Diskussion kritischer Befunde).

Liegen mehrere Output-Knoten vor, dann benötigt man noch eine Wahlregel, die festlegt, wie die Reaktion bestimmt wird. Typischerweise werden die Output-Aktivierungen in eine Exponentialfunktion eingespeist. Die Wahrscheinlichkeit einer bestimmten Reaktion ist dann eine Funktion der relativen Stärke des jeweiligen Outputs (vgl. Nosofsky et al., 1992). Nosofsky (1992) und Massaro und Friedman (1990) haben bewiesen, dass Komponenten-Merkmalmodelle an der Lernasymptote formal äquivalent sind mit denjenigen Prototypentheorien, die, ähnlich wie das Kontextmodell, eine exponentielle Ähnlichkeitsregel verwenden (vgl. Textbox: *Beispiel für eine Exemplar-Theorie: Das Generalisierte Kontext-Modell (GCM)*). Es gibt eine Reihe von Varianten des Komponenten-Merkmalmodells (z.B. bezüglich der Kodierung des Inputs, des Kontexts usw.), denen allen gemeinsam ist, dass sie, ähnlich wie die einfache Rescorla-Wagner-Theorie und Prototypentheorien, nicht in der Lage sind, das Lernen nicht-linear trennbarer Kategorien zu erklären. Dies liegt primär daran, dass der Input nur einfache Merkmale kodiert und nicht sensitiv ist für Merkmalskonfigurationen.

Dieses Defizit wird in Modell II behoben, dem von Gluck und Bower (1988b) vorgeschlagenen konfiguralen Merkmalsmodell („configural-cue model“). Dieses Modell ist im Grunde eine Wiederauflage des „Unique Cue“-Modells von Rescorla (1973). Die einfache Deltaregel wird beibehalten, aber der Input wird um Knoten erweitert, die Merkmalskonfigurationen kodieren (z.B. AB; BC). Ein Grundproblem dieses Modells ist die explosiv anwachsende Menge von Input-Knoten, wenn man nicht

nur Konfigurationen von Paaren, sondern alle möglichen Konfigurationen (z.B. auch ABC) zulässt. Gluck und Bower (1988b) haben deshalb viele ihrer Simulationen mit der Restriktion versehen, nur paarweise Konfigurationen zuzulassen, was allerdings den Erwerb einer Dreifach-Interaktion unmöglich macht.

Abbildung 3b-18 III repräsentiert einen Typ konnektionistischer Netzwerke, der zu den wichtigsten innerhalb der aktuellen konnektionistischen Bewegung gehört. Ähnlich wie die einfachen Netzwerke kodieren diese komplexeren Netzwerke nur einfache Merkmale (A, B, C) auf der Input-Ebene. Auch die Output-Ebene ist vergleichbar. In dem Beispiel gibt es wieder zwei Reaktionsknoten R1 und R2, die etwa zwei sich wechselseitig ausschließende Kategorien repräsentieren können. Neu ist die mittlere Ebene („hidden layer“) mit vermittelnden Knoten („hidden nodes“) H1 und H2. Diese Knoten können so trainiert werden, dass sie konfigurale Beziehungen im Input repräsentieren. Eine der einflussreichsten Lernregeln, die dafür entwickelt wurde, ist *Back-Propagation* (Rumelhart, Hinton & Williams, 1986). *Back-Propagation* kann als Verallgemeinerung der Delta-Regel angesehen werden. *Back-Propagation*-Netzwerke besitzen typischerweise vermittelnde Knoten, deren Aktivierung einer linear-sigmoiden Funktion folgt. Die Aktivierung ist abhängig vom Input und den assoziativen Gewichten zwischen Input und den vermittelnden Knoten. Die Output-Aktivierung ist eine Funktion der Aktivierung der vermittelnden Knoten und der assoziativen Gewichte zwischen den beiden Ebenen. Analog zur Delta-Regel (bzw. der Rescorla-Wagner-Regel) werden die Gewichte proportional zum Fehler am Output modifiziert. Dabei werden die Fehlerinformationen zunächst bis zur Ebene der vermittelnden Knoten und dann bis zur Input-Ebene zurückübertragen.

Dieses Modell ist in der Lage, konfigurale, nicht-lineare Aufgaben zu bewältigen (vgl. Rumelhart et al., 1986), an denen die früheren Modelle mit nur einer Input- und Output-Schicht gescheitert waren (Minsky & Papert, 1969). Gerade diese Leistung gehört zu den wichtigsten Gründen, warum konnektionistische Modelle neuerdings wieder ernst genommen werden, nachdem sie fast 20 Jahre in der kognitionspsychologischen Forschung wenig beachtet wurden. *Back-Propagation*-Netzwerke haben also das

Potential, schwierige konfigurale Aufgaben zu lernen. Aber sind sie auch ein plausibles psychologisches Modell für Lernprozesse?

Auf den ersten Blick erscheinen diese Modelle wenig vielversprechend. Eine ihrer augenfälligsten Eigenschaften ist, dass sie relativ viel Zeit benötigen, um manche konfigurale Aufgaben zu lösen, die Probanden eher leicht fallen (vgl. Nosofsky et al., 1994; Kruschke, 1992). Ein weiteres Problem von *Back-Propagation*-Modellen ist ihre Neigung zu katastrophischer Interferenz (McCloskey & Cohen, 1989). Einmal Gelerntes wird schnell wieder vergessen, wenn neue, mit dem alten Lernmaterial inkompatible Lernaufgaben präsentiert werden. Die Ursache dafür liegt wiederum in der Aktivationsfunktion der vermittelnden Knoten. Aufgrund ihrer linearen Kontur sind sie sensitiv für eine große Zahl von Input-Knoten. Dies führt sehr schnell dazu, dass eine große Zahl gelernter Gewichte in Richtung der neuen Beziehungen modifiziert werden (vgl. Kruschke, 1992).

Das vielleicht erfolgreichste und einflussreichste Modell menschlicher Kategorisierung ist schließlich Kruschkes (1992) ALCOVE-Modell („attention learning covering map“). Es lässt sich als konnektionistische Implementation eines Exemplar-Modells des Begriffslernens beschreiben, wobei es viele Annahmen des generalisierten Kontextmodells (GCM)(Nosofsky, 1986) implementiert (vgl. Textbox: *Beispiel für eine Exemplar-Theorie: Das Generalisierte Kontext-Modell (GCM)*). Dabei werden die Annahmen von Exemplartheorien (insbesondere des generalisierten Kontextmodells) mit konnektionistischen Lernmechanismen verknüpft.

Abbildung 3b-18 IV zeigt die Architektur von ALCOVE. Auf den ersten Blick handelt es sich um ein konnektionistisches Netzwerk, ähnlich wie es auch in den *Back-Propagation*-Modellen Verwendung findet. Die Input-Ebene (A, B) aktiviert über vermittelnde Knoten (E1, E2) die beiden Reaktionsknoten R1 und R2. Auch die Lernregel ist ähnlich. Wieder wird eine von der Delta-Regel abgeleitete Regel eingesetzt, bei der die assoziativen Gewichte in Abhängigkeit vom Fehler beim Output modifiziert werden. Zwischen den Gewichten besteht Konkurrenz. Der entscheidende

Unterschied besteht in der Aktivationsfunktion der vermittelnden Knoten und der Bedeutung der Gewichte.

Die vermittelnden Knoten (E1, E2) repräsentieren einzelne Lernexemplare. Für jedes Lernexemplar wird im Netzwerk ein eigener Knoten angelegt. Die Input-Knoten kodieren die Ausprägungen der zu kategorisierenden Exemplare auf unterschiedlichen Merkmalsdimensionen. Greift man wieder auf eine geometrische Veranschaulichung zurück, dann kann man die Exemplarknoten als Punkte in einem von den Dimensionen aufgespannten Raum betrachten. Die Aktivierung der Exemplarknoten ist nun eine Funktion der Ähnlichkeit mit dem Lern-Input, wobei die im generalisierten Kontextmodell verwendete exponentielle Ähnlichkeitsfunktion verwendet wird. Diese Funktion impliziert einen exponentiell abfallenden Ähnlichkeitsgradienten, was zu der diamantenen Form (in der Abbildung nur zweidimensional) der rezeptiven Felder der Exemplarknoten führt. Anders als die linearen vermittelnden Knoten bei *Back-Propagation*-Netzwerken besitzen die Exemplarknoten also nur ein eingeschränktes rezeptives Feld: Sie werden nur von Lern-Inputs aktiviert, die ihnen vergleichsweise ähnlich sind. Eine wichtige Eigenschaft ist dabei allerdings, dass das System interaktiv ist. Ein Lernexemplar aktiviert meistens nicht nur einen Exemplarknoten, sondern in Abhängigkeit von der Ähnlichkeit mehrere Knoten. Die Aktivierung der Exemplarknoten ist wiederum über assoziative Verbindungen mit den Reaktionsknoten auf der Output-Ebene verbunden.

Eine weitere Besonderheit von ALCOVE sind die an der Input-Ebene befindlichen Aufmerksamkeitsgewichte (#####90). Mit Hilfe dieser Gewichte lässt sich die psychologische Relevanz der unterschiedlichen Merkmalsdimensionen variieren. Denkt man an den geometrischen Merkmalsraum, dann bewirken diese Gewichte eine Dehnung oder Schrumpfung entlang der jeweiligen Dimensionsachse (vgl. Textbox: *Beispiel für eine Exemplar-Theorie: Das Generalisierte Kontext-Modell (GCM)*). Wird die Achse aufgrund zunehmender selektiver Aufmerksamkeit gedehnt, dann bewirkt dies, dass die Lernexemplare entlang dieser Merkmalsdimension leichter diskriminierbar sind. Der Raum lässt sich allerdings nur entlang der Achsen dehnen, nicht in der Diagonale. Entscheidend ist nun, dass in ALCOVE nicht nur die

assoziativen Gewichte zwischen Exemplarknoten und Reaktionsknoten gemäß der erweiterten Delta-Regel gelernt werden, sondern auch die Aufmerksamkeitsgewichte (#####). Ähnlich wie bei *Back-Propagation* wird hier der beim Output erzielte Fehler als Modifikationsgrundlage für die Aufmerksamkeitsgewichte herangezogen. Gerade die Fähigkeit, in Abhängigkeit vom Lernmaterial die selektive Aufmerksamkeit zu optimieren, ist ein zentraler Grund für den Erfolg von ALCOVE bei der Simulation von Kategorienlernen.

ALCOVE wurde auf eine Vielzahl von Aufgaben aus der Kategorisierungsforschung mit vergleichsweise großem Erfolg angewendet (vgl. Kruschke, 1992; Nosofsky & Kruschke, 1992; Nosofsky, Kruschke & McKinley, 1992), wenngleich es auch kritische Befunde gibt (z.B. Macho, 1997). Um weitere Befunde, wie etwa die Interaktion von regelbasiertem und exemplarbasiertem Kategorisieren zu modellieren, hat Kruschke (2001) eine Reihe von erweiterten konnektionistischen Modellen entwickelt, die eine Exemplarlernkomponente mit zusätzlichen Lernmodulen verknüpfen (vgl. auch Erickson & Kruschke, 1998)(vgl. Abschnitt 2).

((2)) 3b-5.2 Wissensbasierte Lerntheorien

Angeht die große Zahl ausgearbeiteter ähnlichkeitsbasierter Lerntheorien ist es nicht verwunderlich, dass sich einige Modellierer bemüht haben, Wissenseffekte in diese Modelle zu integrieren (vgl. Heit, 1997; Medin & Heit, 1999). So haben beispielsweise Choi, McDaniel und Busemeyer (1993) einfache konnektionistische Modelle zugrunde gelegt und dabei die nahe liegende Annahme getestet, dass im Falle von Vorwissen die Assoziationsgewichte nicht bei Null starten, sondern bei Werten, die den früheren Lernerfahrungen entsprechen. Auf diese Weise ließen sich Vorwissenseffekte über einfache artifizielle Kategorien erklären, ihr Modell ist aber im Wesentlichen darauf beschränkt, Vorwissen über assoziative Beziehungen zwischen Merkmalen und Kategorie zu modellieren. Kruschke (1993) hat analog versucht, Vorwissen auf der Basis voreingestellter Aufmerksamkeitsgewichte in sein Modell ALCOVE zu integrieren.

Ein anderes Beispiel für diese Strategie, das auf Exemplartheorien basiert, ist die *Integrationstheorie* von Heit (1994). Dieses Modell geht von der Annahme aus, dass Kategorisierungsentscheidungen nicht nur auf aktuell erlebten Exemplaren beruhen, sondern dass Exemplare aus früheren, anderen Lernkontexten ebenfalls rekrutiert werden. Zieht man beispielsweise in eine neue Stadt und möchte entscheiden, ob eine Person ein Jogger ist, dann wird man gemäß dieser Theorie diese Person nicht nur mit anderen Joggern aus dieser Stadt vergleichen, sondern auch mit Joggern aus den anderen Städten, in denen man gelebt hat. Auf diese Weise können frühere Lernerfahrungen den aktuellen Lernprozess mit beeinflussen.

Ein Problem dieser Ansätze besteht allerdings darin, dass es nicht klar ist, welche Vorerfahrungen aktuell herangezogen werden. Außerdem ist es in diesen Theorien schwierig, komplexere Formen von Vorwissen wie etwa Wissen über Kausalität auszudrücken. Es wurden deshalb auch Modelle vorgeschlagen, die eher der Theorien-Sicht zuzurechnen sind. Eine aus der Künstlichen-Intelligenz-Forschung stammende Technik („explanation-based learning“) verwendet Regeln, die theoretische Annahmen über die Kategorie ausdrücken. Kategorisierung besteht dann darin, dass man überprüft, ob das Lernexemplar die in der Theorie spezifizierten Bedingungen erfüllt (vgl. Mooney, 1993; Pazzani, 1991). Besagt meine Theorie über Becher etwa, dass sie in die Hand genommen werden können sollen, kann ich das Gewicht und das Vorhandensein eines Griffes als kritische Merkmale auszeichnen, auf die ich bei der Klassifikation achte. Bislang liegen allerdings nur wenige empirische Untersuchungen zu diesem Ansatz vor.

Ein Problem aller bisher diskutierten Ansätze besteht darin, dass sie spezifisches Vorwissen, das in früheren Lernkontexten erworben wurde, vergleichsweise gut berücksichtigen können, aber dass sie keine Möglichkeit bieten, abstraktes Vorwissen auszudrücken, wie es etwa Waldmann und Mitarbeiter für den Bereich kausaler Kategorien nachgewiesen haben (vgl. Abschnitt 3). Eine vielversprechende Möglichkeit, kausale Kategorien zu modellieren, bieten formale Methoden der Analyse von kausalen Netzwerken (vgl. Pearl, 2000). Solche Bayesianischen Netzwerke, wie sie in einfachster Form in Abbildung 3b-15 zu sehen sind, vereinfachen die Repräsentation von Kausalwissen, weil sie es ermöglichen, die statistischen Beziehungen zwischen

allen Merkmalspaaren auf der Basis von Wissen über die Stärke der Kausalpfeile und der Häufigkeiten der exogenen Ursachen abzuleiten. Da die Anzahl möglicher statistischer Relationen zwischen allen Ereignissen die Anzahl direkter Kausalbeziehungen, wie sie sich in den Pfeilen ausdrücken, und exogener Ursachen übersteigt, ergibt sich ein deutlicher computationaler Gewinn. Insbesondere ist es möglich, auch die indirekten Beziehungen abzuschätzen, wie etwa die statistische Beziehung zwischen mehreren Effekten einer gemeinsamen Ursache. Waldmann et al. (1995) und Hagmayer und Waldmann (2000) haben gezeigt, dass Probanden tatsächlich sensitiv sind für solche strukturellen Implikationen von Kausalwissen (vgl. auch Rehder & Hastie, 2001). Eine auf Bayesianischen Netzwerken basierende Lerntheorie liegt aber bislang nur in Ansätzen vor (vgl. aber Waldmann & Martignon, 1998).

Insgesamt lässt sich also sagen, dass wissensbasiertes Lernen bisher nur ansatzweise modelliert wurde und ein Desiderat zukünftiger Forschung bleiben muss. Viele wichtige Aspekte von Wissen wie etwa die von Wisniewski und Medin (1994) nachgewiesene enge Koppelung zwischen Theorien und der Beschreibung der Daten werden bislang von keiner Theorie ins Auge gefasst.

((1)) 3b-6 DIE NUTZUNG VON KATEGORIEN

Im Mittelpunkt der Kategorisierungsforschung steht von jeher die Frage, wie Kategorien gebildet werden und welche Eigenschaften sie haben. Hinter dieser Forschungsstrategie steht die Annahme, dass Kategorien unabhängig von ihren Nutzungskontexten stabil und invariant sind. Begründet wurde diese Position häufig mit dem Argument von Rosch und Mervis (1975), dass unsere Konzepte mentale Repräsentationen von Merkmalskorrelationen sind, die sich objektiv in unserer Umwelt finden lassen. Diese naive Widerspiegelungstheorie wurde in den letzten Jahren zunehmend in Frage gestellt (vgl. Malt, 1995, für einen Überblick über die konkurrierenden Positionen). So hat man zunehmend begonnen zu untersuchen, welche Funktionen Kategorien erfüllen können. Dabei hat sich herausgestellt, dass es häufig eine Wechselwirkung zwischen den Kontexten der Kategoriennutzung und der Kategorienstruktur gibt, die die traditionelle Strategie der kontextfreien Untersuchung von Kategorien in Frage stellt. Auch ethnologische und anthropologische

Untersuchungen über Kategorien in anderen Kulturen haben das Augenmerk zunehmend auf kontextuelle Einflussfaktoren gerichtet. Dabei wurde deutlich, dass Kategorisierung ein wesentlich komplexerer Prozess ist als es frühere Theorien vermutet haben (vgl. auch Solomon, Medin & Lynch, 1999).

((2)) 3b-6.1 Der Einfluss von Zielen und pragmatischen Kontexten

Barsalou (1983, 1985, 1987) war einer der ersten, der gezeigt hat, dass Kategorien ad hoc im Hinblick auf aktuelle Ziele gebildet werden können. Typische Beispiele solcher *Ad-hoc-Kategorien* sind etwa „Dinge, die man auf einen Campingtrip mitnimmt“ oder „Dinge, die man aus einem brennenden Haus trägt“. Barsalou konnte zeigen, dass solche Kategorien wie natürliche Kategorien Typikalitätsgradienten aufweisen. Die Basis solcher Einschätzungen sind allerdings nicht Ähnlichkeitsvergleiche zwischen den Exemplaren, sondern ein Vergleich mit einem Idealexemplar. Die Kategorie „Nahrung, die man bei einer Diät zu sich nimmt“ wird beispielsweise mit dem Ideal von null Kalorien verglichen und nicht mit dem Prototyp der typischerweise verzehrten Diätahrungsmittel. Ad-hoc Kategorien sind allerdings, möglicherweise aufgrund ihrer seltenen Aktivierung und der geringeren Überlappung der Merkmale der Exemplare, weniger stabil und kohärent als die herkömmlicherweise untersuchten natürlichen Kategorien. Denkt man etwa an die Dinge, die man aus einem brennenden Haus trägt, dann lassen sich wenige Gemeinsamkeiten zwischen den Objekten finden und die Kategorie würde einem spontan nicht in den Sinn kommen, wenn man diese Dinge vor Augen hätte.

Während Barsalou zielbezogene Kategorien nur als eine von mehreren Formen von Kategorisierung ansieht, vertritt Hoffmann (1993) die extreme Sicht, dass Kategorien generell funktional äquivalente Objekte zusammenfassen, die im Zusammenhang mit Handlungszielen austauschbar sind. Kategorisierung wird in diesem Ansatz also der antizipativen Handlungssteuerung untergeordnet.

Ein Beleg für den Einfluss von Handlungszielen, die mit unterschiedlichen Formen von Expertise einhergehen, sind die Studien von Medin, Lynch, Coley und Atran (1997). Diese Forscher haben Baumexperten untersucht und gefunden, dass Landschaftsgärtner

Bäume anders sortieren als Biologen. Während die Biologen zu einer wissenschaftlichen Taxonomien entsprechenden Anordnung neigten, tendierten die Landschaftsgärtner dazu, die Bäume im Hinblick auf ihre Nutzung und Anordnung in Parkanlagen zu ordnen. Lynch, Coley und Medin (2000) und Atran (1998) haben zudem gefunden, dass Baumexperten Typikalität anders einschätzen als Laien. Während Laien in der Regel den Prototypen mit mittleren Ausprägungen in allen Merkmalen als typischsten Baum ansehen, neigen Experten und auch Itzaj Mayas dazu, ein *Ideal*, etwa den größten Baum, als Vergleichsmaßstab heranzuziehen. Diese Studien zeigen, dass Ideale nicht nur bei Ad-hoc-Kategorien, sondern auch bei natürlichen Kategorien Typikalitätsurteile steuern können.

Ein Beispiel dafür, wie die Nutzung von Kategorien deren Repräsentation verändern kann, stammt aus den Untersuchungen von Ross (1997). In diesen Experimenten lernten Versuchsteilnehmer, fiktive Krankheiten auf der Basis von Symptomen zu klassifizieren. Nach der Diagnose der Erkrankung sollten sich die Probanden zusätzlich für eine Therapie entscheiden. Für diese Entscheidung waren nur einige der für die Diagnose relevanten Symptome wichtig. Sollten nun die Versuchsteilnehmer neue Patienten diagnostizieren, zeigte sich, dass sie nun den therapierelevanten Symptomen besondere Beachtung schenkten und sie für wichtiger und auch für häufiger vorkommend hielten als die für die Therapieentscheidung irrelevanten Symptome. Die Nutzung der Kategorien veränderte also die Repräsentation, obwohl Diagnose und Therapieentscheidung zunächst als unabhängig voneinander zu lernende Aufgaben eingeführt wurden. Ahn und Kim (2001) reinterpreten diese Befunde allerdings im Hinblick auf Annahmen über Kausalität. Es wäre denkbar, dass die Probanden schlossen, dass die Symptome, die über die Therapie entscheiden, eher indikativ sind für die zentralen Ursachen der Erkrankung als die anderen Symptome, da es ja Ziel jeder Therapie ist, die Krankheit zu heilen und nicht nur einige Symptome zum Verschwinden zu bringen.

Eine weitere Funktion von Kategorien besteht in ihrer Nutzbarkeit bei Kommunikationen. Markman und Makin (1998) sind der Frage nachgegangen, welchen Einfluss Kommunikation auf die Bildung von Kategorien haben kann. In dem

Experiment sollten Versuchsteilnehmer LEGO-Modelle entweder allein und zusammen mit einer anderen Person bauen. In der Bedingung mit den Personenpaaren musste immer eine Person die Anweisungen geben und die andere Person diese Anweisungen ausführen. Die Befunde zeigten, dass Personen, die zusammengearbeitet hatten, später die LEGO-Bauteile ähnlicher sortierten als die Einzelpersonen. Die Kommunikation zwischen den Versuchsteilnehmern führte also zu Gemeinsamkeiten der Kategorisierung der Umwelt.

((2)) 3b-6.2 Konzeptuelle Kombination

Konzepte können auch kombiniert werden. So ist es uns möglich, die Konzepte „Erdbeere“ und „Allergie“ zu dem Konzept „Erdbeerallergie“ zusammenzufügen und damit einen Sachverhalt ausdrücken, der durch die Elemente allein nicht bezeichnet wird. Diese Möglichkeit der Kombination ermöglicht ein hohes Maß an Produktivität bei der Nutzung kategorialen Wissens.

Bisher gibt es keine Theorie der konzeptuellen Kombination, die allen Phänomenen gerecht wird (vgl. auch Medin et al., 2001). Ursprünglich hatte man gehofft, ein Modell zu finden, dass die Bedeutung des kombinierten Konzepts allein als Funktion der Bedeutung seiner Elemente ausdrückt. Ein Beispiel für diese Art von Modell ist das *selektive Modifikationsmodell* von Smith und Osherson (1984), das sich auf Adjektiv-Nomen-Kombinationen bezieht. Gemäß diesem Modell wird ein Konzept wie „brauner Apfel“ so verstanden, dass man zunächst den Prototypen von Apfel aktiviert, dann würde die Farbdimension besonders beachtet und die prototypische Farbe („rot“) durch „braun“ ersetzt werden. Dieses Modell erlaubt es, eine Reihe von Phänomenen zur Typikalitätseinschätzung solcher komplexer Konzepte zu erklären (Smith, Osherson, Rips & Keane, 1988).

Es gibt aber auch viele Befunde, die dieses Modell nicht erklärt. So haben beispielsweise Medin und Shoben (1988) gezeigt, dass bei manchen Konzepten die Typikalität der kombinierten Konzepte sich nicht aus den Elementen vorhersagen lässt. So werden kleine Löffel als typischer eingeschätzt für die Kategorie Löffel als große Löffel. Ebenso sind metallene Löffel typischer als Holzlöffel. Dennoch halten die

meisten Urteiler große hölzerne Löffel für typischer als kleine hölzerne und große metallene Löffel. Man könnte die Theorie natürlich dadurch retten, dass man postuliert, dass wir Konzepte wie Holzlöffel getrennt und unabhängig von den Komponenten lernen. Dies würde aber angesichts der vielen Kombinationsmöglichkeiten von Konzepten als generelle Strategie zu Schwierigkeiten führen. Ein weiteres Problem besteht darin, dass Kombinationen gelegentlich Merkmale aufweisen, die keines der Elemente hat. Haustiervögel leben in der Regel in Käfigen, was weder typisch für Haustiere noch für Vögel ist (Hampton, 1996; Murphy, 1988).

Schließlich lässt sich zeigen, dass es eine Reihe unterschiedlicher Strategien gibt, um kombinierte Begriffe zu verstehen (Wisniewski & Gentner, 1991; Wisniewski & Love, 1998). Kombinationen zweier Nomen werden häufig so verstanden, dass man eine auffällige Eigenschaft eines Nomens auf das andere überträgt. Dies funktioniert am besten, wenn die beiden Nomen Objekte bezeichnen, die vergleichbare, korrespondierende Merkmale haben (z.B. das Fellmuster bei Tieren)(„alignable features“). Ein „Zebraferd“ wird deshalb spontan als Pferd mit Streifen interpretiert. Liegt eine solche Korrespondenz der Merkmale („non-alignable“) nicht vor, dann wird eher nach einer Relation gesucht, die die Merkmale verknüpfen könnte. So wird beispielsweise ein „Zebrahaus“ eher als Haus für Zebras verstanden und nicht als Haus mit Streifen.

((2)) 3b-6.3 Sprache und Kategorien

Sprache ermöglicht es Menschen, ihr Wissen untereinander zu kommunizieren. Die enge Verbindung von Wörtern und Konzepten hat viele Psychologen beschäftigt (vgl. auch das Kapitel von Zwitserlood und Bölte in diesem Band). Sind Wörter einfach sprachliche Benennungen von sprachunabhängigen Konzepten und Kategorien oder hängen unsere Konzepte und Kategorien davon ab, welche Sprache wir sprechen? Die radikalste Sichtweise, die kaum jemand vertritt, besagt, dass Denken ohne Sprache nicht möglich sei (vgl. Bloom, 2000). Diese Sicht würde die wenig plausible Implikation haben, dass man Tieren, Babys oder Aphasikern Denkfähigkeit absprechen müsste (Pinker, 1994).

Eine demgegenüber abgemilderte Sicht wurde u.a. von Whorf (1956) formuliert. Whorf hat die Doktrin des *Linguistischen Determinismus* entwickelt, demzufolge Sprache unser Denken determiniert. Er ging zwar davon aus, dass es vorsprachliches Denken gibt, war aber der Meinung, dass dieses Denken so lange unstrukturiert und chaotisch sei, bis es durch die Muttersprache strukturiert wird. Ein berühmtes Beispiel von Whorf ist der Kontrast zwischen Eskimos und Amerikanern. Während Eskimos über eine große Zahl von Wörtern für verschiedene Formen von Schnee verfügen, verwenden Amerikaner nur ein Wort (vgl. Pullum, 1991, für eine Kritik an dieser Position).

Whorfs radikale Sicht wurde mittlerweile weitgehend aufgegeben (Bloom, 2000). So konnte beispielsweise gezeigt werden, dass der Umstand, dass in einer Sprache nur wenige Farbbezeichnungen existieren, keineswegs bedeutet, dass nicht viel mehr Farben perzeptuell und kategorial diskriminiert werden können. Es konnte beobachtet werden, dass der Erwerb einer Fremdsprache mit mehr Farbbezeichnungen kaum Probleme bereitet. Ein neueres Beispiel für die potentielle Unabhängigkeit von Sprache und vorsprachlichen Kategorien stammt aus einer interkulturellen Studie von Malt, Sloman, Gennar, Shi und Wang (1999). In dieser Studie wurde untersucht, wie Behälter in unterschiedlichen Kulturen kategorisiert werden. Interessanterweise zeigte sich eine Dissoziation zwischen Ähnlichkeitsurteilen und Benennungen. Während es zwischen Englisch, Chinesisch und Spanisch deutliche Unterschiede darin gibt, welche Behälter die gleiche Bezeichnung erhalten, ergaben sich kaum Unterschiede in den Ähnlichkeitseinschätzungen der Probanden. Dabei zeigte sich auch, dass Benennungen nicht unbedingt auf der Basis großer Ähnlichkeit vollzogen werden. Pillenbehälter werden im Englischen beispielsweise als Flaschen bezeichnet, obwohl sie wenig Ähnlichkeit zu Flaschen haben. Eine mögliche Erklärung für diese Divergenzen ist, dass Benennungen auch von kommunikativen Funktionen und kulturellen Traditionen beeinflusst werden und nicht nur von perzeptiven Faktoren. Diese Befunde zeigen, dass unterschiedliche Kategorien in der sprachlichen Kommunikation und beim Objekterkennen genutzt werden können und belegen die enge Verbindung von Kategorien und deren Nutzung (vgl. auch Smith & Samuelson, 1997).

Aber auch die radikale Gegensicht zum Linguistischen Determinismus, dass unsere Kategorien vollkommen unabhängig von Sprache seien, wird nicht mehr von vielen Psychologen geteilt. Diese Sicht wurde von Fodor (1975) formuliert, der Spracherwerb praktisch als Erwerb einer zweiten Sprache ansah. Dieser Sicht zufolge kommen alle Kinder mit Konzepten („mentalese“) zur Welt, die beim Spracherwerb mit den Wörtern in Verbindung gebracht werden. Fodor argumentierte, dass nur so erklärbar ist, warum Kinder relativ schnell alle Sprachen der Welt erwerben können.

Wie kann man sich also das Wechselspiel von Sprache und Kategorien nach derzeitigem Erkenntnisstand vorstellen? Eine gemäßigtere Sicht, die in den letzten Jahren zunehmend empirisch gestützt wird, geht von einer wechselseitigen, wenngleich grundsätzlich reversiblen Beeinflussung zwischen Sprache und Denken aus. So wird vermutet, dass die linguistischen Kategorien der Muttersprache bestimmte Erfahrungsdimensionen zugunsten anderer hervorheben und damit das Denken in subtiler Weise beeinflussen (Slobin, 1996; Hunt & Agnolli, 1991). Anders als Whorf geht man allerdings davon aus, dass dieser Einfluss nicht rigide ist, und dass die Kategorien durch weitere Erfahrungen (z.B. Zweitspracherwerb) modifiziert werden können.

Ein Beispiel für diese Interaktion zwischen Sprache und Denken stammt aus der Forschung von Boroditsky (2001). Sie hat sich mit der Frage beschäftigt, wie wir in verschiedenen Kulturen über *Zeit* denken. Generell neigen wir dazu, zeitliche Relationen mit Hilfe von Wörtern auszudrücken, die auch räumliche Beziehungen ausdrücken. Im Englischen beispielsweise werden Präpositionen wie *before* und *behind* verwendet, um zukünftige von vergangenen Ereignissen zu unterscheiden. Man drückt also Zeit mit Hilfe von Begriffen aus, die auch *horizontale* räumliche Relationen benennen. Interessanterweise unterscheidet sich die Sprache Mandarin in dieser Beziehung vom Englischen. In dieser Sprache werden neben den horizontalen räumlichen Relationen auch vertikale zur Beschreibung von Zeit benutzt: *Oben* entspricht hier zukünftigen und *unten* vergangenen Ereignissen.

Abb. 3b-19

Boroditsky (2001) ist nun der Frage nachgegangen, ob diese Unterschiede in der Sprache sich auch im Denken niederschlagen. Sie hat dazu eine Aufgabe entwickelt, bei Relationen so schnell wie möglich beantworten (z.B. „März ist früher als April“), die mit nicht-räumlichen Begriffen formuliert wurden. Interessanterweise zeigen die Befunde, dass Amerikaner schneller die temporale Frage beantworten konnten, wenn sie direkt zuvor die Voraufgabe mit horizontalen Relationen bearbeitet hatten, während bei Mandarin die Antwortzeit am kürzesten war, wenn sie vorher die Voraufgabe mit vertikalen Relationen bekamen. Dies lässt sich als Effekt der Muttersprache auf die Art und Weise erklären, wie über Zeit spontan gedacht wird. Dieser Einfluss ist aber nicht absolut und irreversibel. In einer weiteren Studie konnte Boroditsky nämlich auch zeigen, dass Amerikaner nach kurzem Training ähnliche Effekte zeigten wie Sprecher des Mandarin.

Andere Beispiele für die Beziehung zwischen Muttersprache und Denken wurden in dem Bereich räumlicher Relationen gefunden. So konnte Levinson (1996) zeigen, dass Holländer (wie auch Deutsche) räumliche Relationen eher situationsspezifisch beschreiben (vgl. auch Bowerman, 1996, für eine Studie über räumliche Präpositionen im Koreanischen). Im Holländischen werden in der Regel *intrinsische* Bezugssysteme verwendet, die die Position von Objekten relativ zu anderen Objekten in der Szene beschreiben („Der Junge steht *vor* dem Lastwagen“), oder es werden Bezugssysteme zugrunde gelegt, die relativ zu der Perspektive des Betrachters konstruiert werden („Der Junge steht *links* vom Lastwagen“). Demgegenüber verwendet man im Tzeltan-Dialekt, einer Sprache, die in Mexiko gesprochen wird, *absolute* Koordinaten, die in etwa unseren Himmelsrichtungen entsprechen („Der Junge steht *südlich* vom Lastwagen“).

Levinson konnte zeigen, dass diese sprachlichen Unterschiede beeinflussten, wie Probanden aus den beiden Kulturen Objekte räumlich anordneten. Sollten etwa in einem Experiment vier Objekte, die nebeneinander auf einem Tisch lagen, auf einen räumlich anders orientierten Tisch in gleicher Weise angeordnet werden, wählten die Holländer die gleiche relative Anordnung (rechts-links), während Sprecher des Tzeltan die Objekte so platzierten, dass ihre Position der gleichen Himmelsrichtung entsprach wie

beim ersten Tisch. Gegen die Interpretation, dass es sich um einen Einfluss der Sprache handelt, lässt sich allerdings einwenden, dass auch andere Faktoren eine Rolle spielen könnten wie etwa die unterschiedliche Ökologie. Holländer halten sich mehr in kleineren Räumen auf als Sprecher des Tzeltal, so dass relative Positionierungen für Holländer sinnvoller erscheinen (vgl. Bloom, 2000). Dass es sich nicht um irreversible Unterschiede handelt, zeigt sich daran, dass die beiden Sprachgruppen natürlich auch in anderen Bezugssystemen denken können.

Die bisher diskutierten Studien legen die Sicht nahe, dass unser kognitives System flexibel genug ist, um die Semantik ganz unterschiedlicher Sprachen zu erwerben. Dabei kann die jeweilige Sprache dazu führen, dass bestimmte Unterscheidungen in den Vordergrund geschoben werden, während andere in den Hintergrund treten, und kann damit bestimmte Denkweisen nahe legen. Solche Einflüsse sind aber nicht spezifisch für Sprache, sondern findet sich auch bei anderen Lernaufgaben, die Kategorien generieren und auf diese Weise Denk- und Wahrnehmungsprozesse beeinflussen (vgl. z.B. die oben diskutierte Studie von Goldstone und Steyvers, 2001).

Als mögliche Ausnahme für diese Generalisierung wird in letzter Zeit mathematisches Denken diskutiert. Nach Dehaene (1997) gibt es zwei kognitive Systeme, die mit Zahlen operieren. Ein primitives System, das sich auch bei Tieren findet, ist in der Lage, approximative Vergleiche von Mengen mit unterschiedlicher Anzahl von Elementen zu machen. Dieses System, das sprachunabhängig ist, funktioniert allerdings nur, wenn die Zahlen vergleichsweise klein sind. Operationen mit großen Zahlen (z.B. $20-2=18$) sind hingegen nicht mit diesem System zu bewältigen. Nach Dehaene setzen solche exakten Operationen mit großen Zahlen *sprachliche* Repräsentationssysteme voraus, die nur Menschen zur Verfügung stehen. Gestützt wird diese Hypothese durch neuropsychologische Befunde, die zeigen, dass der Verlust der Fähigkeit zum mathematischen Denken häufig mit dem Verlust der Sprachfähigkeit einhergeht (vgl. Bloom, 2000). Außerdem konnten Studien mit bilingualen Sprechern (russisch und englisch) zeigen, dass exakte mathematische Operationen leichter in der Sprache durchgeführt können, in der sie trainiert wurden (unabhängig von der Muttersprache), während approximative Mengenvergleiche sprachunabhängig gleich gut zu bewältigen

waren (Spelke & Tsivkin, 2001). Diese Befunde deuten darauf hin, dass bestimmte Formen des Denkens erst durch das Medium von Sprache ermöglicht werden.

((2)) 3b-6.4 Kategorien und Induktion

Ein letztes Beispiel für die Nutzung von Kategorien soll die Rolle von Kategorien beim induktiven Lernen sein (vgl. Heit, 2000, für einen Überblick). Osherson, Smith, Wilkie, Lopez und Shafir (1990) haben ein einflussreiches Modell der kategorienbasierten Induktion (*similarity-coverage model*) entwickelt (vgl. Sloman, 1993, für ein alternatives Modell). Die Hauptidee ist, dass die Ähnlichkeit zwischen Exemplaren und Kategorien den Induktionsprozess steuert. Ein Beispiel für induktive Schlüsse ist etwa ein Argument mit der Prämisse „Rotkehlchen haben Eigenschaft X“ und entweder der Konklusion „Straußenvögel haben Eigenschaft X“ oder der Konklusion „Alle Vögel haben die Eigenschaft X“. Für X werden Eigenschaften gewählt, bei denen die Probanden nicht wissen, ob sie auf die Aussagen zutreffen oder nicht (z.B. „haben eine hohe Konzentration von Kalium im Blut“). Die Aufgabe besteht darin zu entscheiden, welches Argument stärker ist. In dem Beispiel finden Versuchsteilnehmer in der Regel den Schluss auf Straußenvögel schwächer als den Schluss auf alle Vögel (s.u.).

Die Stärke eines Arguments wird in diesem Modell von zwei Faktoren beeinflusst: der Ähnlichkeit zwischen der Kategorie in der Prämisse und der Kategorie in der Konklusion und der Ähnlichkeit der Kategorie in der Prämisse zur niedrigsten taxonomischen Kategorie, die noch die Kategorien in Prämisse und Konklusion umspannt. Die Konfidenz in ein Urteil basiert auf einer Gewichtung dieser beiden Faktoren. So wird die Argumentstärke zwischen der Prämisse über Rotkehlchen und der Konklusion über Vögel so bestimmt, dass man zunächst die Ähnlichkeit von Rotkehlchen zu Vogel bestimmt (die perfekt ist, weil Rotkehlchen Vögel sind) und dann die Ähnlichkeit zu anderen Exemplaren der Kategorie Vogel abschätzt. (Vogel ist die niedrigste Kategorie, die „Rotkehlchen“ und „Vogel“ umspannt.) Daraus folgt, dass ein Argument von Rotkehlchen auf Vögel als stärker empfunden wird als ein Argument von Straußenvögeln auf Vögel, weil Rotkehlchen typischere Vögel sind und damit eine höhere Ähnlichkeit zu anderen Exemplaren haben. Das Modell macht auch die interessante Vorhersage, dass ein Schluss von Rotkehlchen auf Straußenvögel als

stärker eingeschätzt werden sollte als ein Schluss von Straußenvögeln auf Rotkehlchen, ein Muster, das in empirischen Studien auch gefunden wurde. Wieder liegt der Grund in der unterschiedlichen Typikalität des in der Prämisse erwähnten Exemplars für die niedrigste umspannende taxonomische Kategorie (Vögel).

Viele Vorhersagen des Modells sind intuitiv plausibel, manche sind allerdings aber auch überraschend. So sagt es vorher, dass ein Argument von der Prämisse „Alle Rotkehlchen haben X“ auf „Alle Vögel haben X“ als stärker empfunden werden sollte als auf die Konklusion „Alle Straußenvögel haben X.“ Diese widerspricht normativen Prinzipien, weil aus logischen Gründen alles, was auf alle Vögel zutrifft, automatisch auch auf Straußenvögel zutreffen muss. Der Grund für die vorhergesagte Asymmetrie ist, dass Rotkehlchen aufgrund ihrer hohen Typikalität der Kategorie Vögel als ähnlicher empfunden werden als der atypischen Subkategorie Straußenvögel. Die empirischen Befunde bestätigen diese Vorhersage.

Eine andere intuitiv plausible Vorhersage des Modells betrifft Argumente mit mehreren Prämissen. Hier sagt das Modell einen Unterschiedlichkeitseffekt vorher. So wird das Argument „Jaguare haben X und Leoparden haben X“ auf „Alle Säugetiere haben X“ als schwächer empfunden als das gleiche Argument, bei dem „Leoparden“ durch „Mäuse“ ausgetauscht wurde. Der Grund dafür ist, dass im zweiten Fall unterschiedlichere Tiere genannt werden, die die umfassende Kategorie der Säugetiere besser abdecken (d.h. eine höhere Gesamtähnlichkeit zu den Exemplaren dieser Kategorie haben). Tatsächlich ließen sich die vorhergesagten Effekte bei studentischen Versuchsteilnehmern finden (vgl. Heit, 2000).

Es gibt aber auch Ausnahmen, insbesondere wenn man Experten in einer Domäne untersucht (vgl. Coley, Medin, Proffitt, Lynch & Atran, 1999; Heit, 2000, für Überblicke). Vergleicht man etwa ein Argument wie „Krankheit A findet sich bei Trauerweiden und Föhren“ mit dem Argument „Krankheit B findet sich bei der Papierbirke und der Flussbirke“, dann sagt das Modell von Osherson et al. (1990) vorher, dass man aufgrund der größeren Unterschiedlichkeit der betroffenen Bäume Krankheit A eher bei allen Bäumen vermuten würde. Tatsächlich halten studentische

Versuchsteilnehmer aus Michigan das erste Argument für stärker. Anders sieht es aber aus, wenn man Baumexperten untersucht oder wenn man (mit Material, das auf ihren ökologischen Kontext zugeschnitten ist) Mitglieder der Itzaj Maya befragt, die man ebenfalls als Experten für ihre Ökologie betrachten kann. Hier findet man eine Präferenz für die engere Kategorie (also das Argument mit Krankheit B).

Der Grund für diese überraschende Wahl ist, dass Experten eher dazu neigen, kausale, ökologische Argumente zu verwenden, während Laien mangels Vorwissen eher auf allgemeine Ähnlichkeitsbeziehungen zurückfallen müssen (vgl. auch Proffitt, Medin & Coley, 2000). So wurde bei dem Baumbeispiel von manchen Experten bei der Begründung für Krankheit B etwa gesagt, dass Birken sehr anfällig für Krankheiten und weit verbreitet seien, so dass man sich gut vorstellen könne, dass sich die Krankheit schnell auf andere Bäume ausbreitet. Dieses Beispiel zeigt die Bedeutung von Vorwissen über kausale Mechanismen, das ähnlichkeitsbasiertes Denken überlagern kann.

Beispiele für die Rolle von Kausalwissen lassen sich auch finden, wenn man sich die Rolle der Prädikate in induktiven Argumenten ansieht (also das X in den Beispielen oben). Darüber sagt das Modell von Osherson et al. (1990) nichts, da es auf die vorkommenden Objektkategorien fokussiert und mit inhaltlich unspezifischen (bzw. unvertrauten) Prädikaten operiert (vgl. Heit, 2000). Verwendet man aber vertraute Prädikate, dann zeigt sich auch hier ein Einfluss auf die Einschätzung der induktiven Stärke. So empfinden Versuchsteilnehmer das Argument von „Viele ehemalige Kriminelle sind Leibwächter“ auf „Viele Kriegsveteranen sind Leibwächter“ als überzeugender als das Argument „Viele ehemalige Kriminelle sind arbeitslos“ auf „Viele Kriegsveteranen sind arbeitslos“. Da die gleichen Personenkategorien verwendet werden, muss der Unterschied an den Prädikaten liegen. Nach Sloman (1994) besteht der Grund für den Unterschied darin, dass man sich im ersten Fall eine kausale Erklärung zurechtlegen kann („Kriminelle und Veteranen sind kräftig und kampferprobt“), die den induktiven Schluss plausibel macht. Ähnlich wie bei der Diskussion ähnlichkeitsbasierter Kategorisierungstheorien zeigt sich auch hier, dass

dieser Ansatz zumindest unvollständig ist. Vorwissen kann zu einer Veränderung der relevanten Merkmale und damit der für die Argumente relevanten Beziehungen führen.

((1)) 3b-7 AUSBLICK

Ausgangspunkt dieses Kapitels war die Überlegung, dass Kategorien es uns erlauben, verschiedene Objekte oder Ereignisse zu einer gemeinsamen Klasse zusammenzufassen. Dies versetzt uns in die Lage, bereits erworbenes Wissen auf neue Erfahrungen anzuwenden. Ein zentrales Forschungsthema betraf dabei die Frage, nach welchen Prinzipien wir unsere Erfahrungen zu Klassen zusammenfassen. Die klassische Sicht, die davon ausging, dass unsere Konzepte analog zu Definitionen repräsentiert werden, wurde bald aufgegeben und ersetzt durch die Annahme, dass unser kognitives System versucht, globale Ähnlichkeitsbeziehungen in der Welt durch seine Kategorien abzubilden. So erfolgreich die Ansätze, die ähnlichkeitsbasierte Kategorisierung annehmen, auch sind, erwiesen sie sich auch als zumindest unvollständig. Es wurde klar, dass die Frage, was Ähnlichkeit ist, keineswegs einfach zu klären ist. Herkömmliche Theorien gingen da von zu einfachen Annahmen aus, die erst langsam durch verbesserte Theorien ersetzt werden. Ein weiteres Manko früherer Ansätze ist, dass man sich zunächst fast ausschließlich auf die Beziehung zwischen der objektiven Welt und den sie abbildenden Konzepten als Basis von Kategorisierung konzentriert hat und dabei annahm, dass Kategorien kontextinvariante Repräsentationen von korrelativen Strukturen in der Umwelt darstellen. In den letzten Jahren wurde hingegen zunehmend klar, dass der Kontext der Nutzung von Kategorien, also ein Faktor, der von Menschen an die Kategorisierungsaufgabe herangetragen wird, mindestens ebenso bedeutend ist wie der Einfluss von Strukturen in der Welt. In gewissem Sinne ist natürlich immer schon klar gewesen, dass Ähnlichkeitsbeziehungen nicht ohne Bezug auf die perzeptiven Fähigkeiten unseres kognitiven Systems bestimmt werden können. Dennoch konnte darüber hinaus in den letzten Jahren gezeigt werden, dass bereichsspezifisches und abstraktes Vorwissen und die Zwecke der Kategoriennutzung einen großen Einfluss auf die Art der Kategorienbildung haben. Kategorien lassen sich häufig als Kompromisse zwischen verschiedenen Nutzungsanforderungen verstehen, gelegentlich findet man aber auch, dass die Art der Kategorien, die für eine bestimmte Klasse von Objekten gebildet werden, auch aufgabenabhängig variieren kann.

Kategorien lassen sich also weder als reines Abbild der Struktur der Umwelt, noch als umweltunabhängige Konstruktion der sie nutzenden Menschen betrachten, sondern sie sind ein variables Produkt der Interaktion dieser beiden Komponenten. Es ist zu erwarten, dass die Interaktion zwischen Kategorien und ihrer Nutzung ein zentraler Fokus der Forschung sein wird.

Das verstärkte Interesse an der Nutzung von Kategorien erklärt auch das neu erwachte Interesse an anthropologischen Studien. Nisbett, Peng, Choi und Norenzayan (2001) haben die provozierende Hypothese aufgestellt, dass viele Befunde der Kognitionspsychologie vielleicht nur für westliche Kulturen Geltung beanspruchen können. Eine interessante Forschungsfrage ist deshalb, ob es interkulturelle Unterschiede gibt in der Art und Weise, wie die Welt repräsentiert wird. Diese Forschung steht auch in engem Zusammenhang mit der in letzter Zeit wieder stark untersuchten Frage, wie Sprache und Denken zusammenwirken.

Ein weiteres interessantes aktuelles Forschungsfeld betrifft bisher vernachlässigte Typen von Kategorien. Objektkategorien wurden häufig untersucht. Wenig weiß man allerdings über Ereigniskategorien oder abstrakte Kategorien. Auch das Zusammenspiel zwischen kategorialen Repräsentationen und kausalem und funktionalem Wissen wird in letzter Zeit immer stärker untersucht.

Schließlich gibt es in der Kategorienforschung wie in anderen Bereichen der Kognitionspsychologie auch ein verstärktes Interesse an der Frage, wie Kategorien im Gehirn repräsentiert, gelernt und abgerufen werden. Viel versprechend scheinen hier Versuche, computationale, neurowissenschaftliche und experimentelle Methoden zusammenzuführen.

Wie sich alle in diesem Kapitel zusammengefassten Befunde einmal zu einer einheitlichen Theorie menschlichen Kategorisierens integrieren lassen, werden zukünftige Forschungen entscheiden müssen. Dabei ist absehbar, dass Methoden der Computermodellierung, der Neurowissenschaften und der Anthropologie die herkömmlichen verhaltensbezogenen psychologischen Untersuchungsstrategien

ergänzen werden. Nur wenn man konvergierende Evidenz mit unterschiedlichen Methoden erhalten kann, wird es möglich sein, dem Ziel einer integrierten Theorie menschlichen Kategorisierens näher zu kommen.

((2)) Kernsätze

Kategorien fassen Objekte aufgrund ihrer Gemeinsamkeiten zusammen. Auf diese Weise können wir Wissen, das wir über eine Kategorie erworben haben, auf neue Erfahrungen anwenden. Kategorien spielen eine zentrale Rolle in einer Reihe kognitiver Prozesse.

Ähnlichkeitsbasierte Theorien gehen davon aus, dass Ähnlichkeit zwischen Objekten oder Ereignissen der wichtigste Faktor der Kategorisierung ist. Der Einfluss von Vorwissen wird ausgeblendet. Derzeit konkurrieren eine Reihe, in der Regel mathematisch formalisierter, Modelle um die Erklärung experimenteller Befunde.

Ähnlichkeit ist ein hochkomplexes Konzept, das nur zum Teil geklärt ist. Vorwissen, Kontexteffekte und Relationen zwischen Merkmalen sind Faktoren, die in den bislang postulierten Theorien der Ähnlichkeit nur unzureichend berücksichtigt werden.

Gemäß der Theorien-Sicht lassen sich Kategorien nicht als Bündel von Merkmalen verstehen, sondern als strukturierte Gebilde, die ähnlich wie Theorien, funktionale oder kausale Relationen zwischen den Merkmalen spezifizieren. Vorwissen über eine Domäne ist ein wichtiger Einflussfaktor, der von ähnlichkeitsbasierten Theorien nur unzureichend berücksichtigt wird.

Vielen Kategorisierungstheorien liegt die Annahme zugrunde, dass die Repräsentation von Kategorien gleich ist für unterschiedliche Inhaltsbereiche. Im Rahmen der Theorien-Sicht wurden allerdings Studien durchgeführt, die zeigen, dass unterschiedliche Kategorieninhalte verschieden repräsentiert werden.

Kategorien können in einer Vielzahl von Relationen zueinander stehen. Sie können Teil einer hierarchischen taxonomischen Anordnung sein, sich aber auch in ereignisbezogenen Beziehungen befinden oder nicht-hierarchisch nebengeordnet sein.

Eine Reihe von Theorien konkurriert darum, wie man den Erwerb von Kategorien modellieren soll. Konnektionistische Netzwerkmodelle, die den Lernprozess als

Assoziationsvorgang ansehen, korrespondieren dabei mit ähnlichkeitsbasierten Kategorientheorien, während für die Theorien-Sicht bislang nur ansatzweise formale Modelle des Lernens vorliegen.

Kategorien dienen einer Reihe unterschiedlicher Funktionen. Dabei zeigt die Forschung, dass Kategorien nicht invariant sind, sondern sich in Abhängigkeit vom Kontext, in dem sie genutzt werden, aktuell oder dauerhaft verändern können.

((2)) Schlüsselbegriffe

Ad hoc-Kategorien (ad hoc categories) – Kategorien, die für einen bestimmten Zweck aktuell generiert werden (z.B. „Dinge, die man aus dem Haus trägt, wenn es brennt“).

Ähnlichkeit (similarity) – Ähnlichkeit von Objekten wurde lange Zeit als Basis der Kategorienbildung angesehen. In den letzten Jahren wird allerdings deutlich, dass Ähnlichkeit selbst ein hochkomplexes Produkt einer Reihe von Kontextfaktoren ist, die bisher nur ansatzweise erforscht sind.

Ähnlichkeitsbasierte Kategorisierungstheorien (similarity-based categorization theories) – Oberbegriff zu einer Reihe von Theorien (→ klassische Sicht, → Prototypentheorien, → Exemplartheorien, → Entscheidungsgrenzen), die davon ausgehen, dass Kategorisierung datengesteuert ist und von Ähnlichkeitsbeziehungen der Objekte oder Ereignisse determiniert wird.

Artefakte (artefacts) – Kategorien, die sich auf von Menschen geschaffene Objekte beziehen (z.B. Fahrzeuge, Instrumente, Möbel).

Basis-Kategorien (basic categories) – Psychologisch privilegierte Kategorienebene in Taxonomien (→ Taxonomien). Kategorien auf dieser Ebene (z.B. Hund) werden als Erstes gelernt, am schnellsten benannt und zeichnen sich häufig durch gleichartiges Aussehen aus. In der Regel handelt es sich um die mittlere Ebene in Taxonomien. Expertise und Kultur können aber mit beeinflussen, welche Ebene die Basis-Ebene ist.

Defizite, kategorien-spezifische (category-specific deficits) – Klinische Beobachtung, dass es bei einzelnen Patienten mit Gehirnläsionen zu selektiven Problemen mit spezifischen Inhaltsbereichen kommen kann (z.B. Lebewesen vs. Artefakte).

Determinismus, linguistischer (linguistic determinism) – Von Whorf (1956) vertretene Sicht, dass Sprache unser Denken determiniert.

Domänenspezifität (domain specificity) – Annahme, dass kognitive Prozesse und Kategorien nicht bereichsübergreifend gleichartig sind, sondern dass sie bereichsspezifisch operieren.

Entscheidungsgrenzen (decision boundaries) – Weiterentwicklung der Signalentdeckungstheorie. Die Theorie geht davon aus, dass sich Kategorienelemente in einem multidimensionalen Merkmalsraum repräsentieren lassen. Kategorisierung besteht darin, dass man diesen Raum mit Hilfe von Grenzen in Segmente unterteilt, die den einzelnen Kategorien entsprechen.

Ereignisbezogene Kategorien (event-related categories) – Kategorien, die Ereignisabläufe ausdrücken (z.B. Frühstück, Restaurantbesuch)(auch Skript-Kategorien genannt).

Essentialismus (psychological essentialism) – Annahme, dass natürliche Arten (→ natürliche Arten) sich durch eine gemeinsame, ihnen innewohnende Essenz auszeichnen, die entscheidend ist für die Kategorienzugehörigkeit.

Exemplar-Sicht (exemplar view) – Theorie der Kategorisierung, die davon ausgeht, dass beim Lernen die einzelnen Exemplare im Gedächtnis gespeichert werden. Kategorisierung wird auf der Basis von Ähnlichkeitsvergleichen mit den einzelnen Exemplaren vollzogen.

Explizites Kategorielernen (explicit learning of categories) – Bewusstes Suchen von Regeln, die die Kategorien unterscheiden (→ Klassische Sicht). Solche Prozesse setzen vergleichsweise einfach strukturierte Kategorien voraus.

Funktionen von Kategorien (functions of categories) – Klassifikation, Verstehen, Lernen, Inferenzen, Erklärungen, Vorhersage, Denken, Handlungsplanung, Sprache, Kommunikation

Implizites Kategorielernen (implicit learning of categories) – Unbewusster Erwerb von Kategorienwissen; wird bei komplexen Kategorienstrukturen vermutet.

Induktion, kategorienbasierte (category-based induction) – Schluss von bestimmten Kategorien (z.B. „Rotkehlchen haben X“) auf andere mehr oder weniger ähnliche Kategorien (z.B. „Straußenvögel haben X“).

Kausalmodelle (causal models) – Ereignisstrukturen, die Ursache-Wirkungszusammenhänge beinhalten. Kausalrelationen sind asymmetrisch, da Ursachen Wirkungen hervorrufen können, aber nicht umgekehrt. Kausale Kategorien beziehen sich häufig auf Kausalmodelle (z.B. Krankheiten, bei denen Symptome und Ursachen kausal verknüpft sind).

Kategorien (categories) – Klassen von Objekten oder Ereignissen, die aufgrund von Gemeinsamkeiten zusammengefasst werden.

Konnektionistische Lernmodelle (connectionist learning models) – Computermodelle des Lernens, die in Analogie zu neuronalen Netzwerken entwickelt wurden. Sie bestehen in der Regel aus Knoten, die Merkmale oder Kategorien kodieren und assoziativen Verbindungen, die Assoziationsstärken ausdrücken. Komplexere Modelle beinhalten auch vermittelnde weitere Knoten, die die Aufgabe haben, Merkmalskonfigurationen oder Lernexemplare zu repräsentieren. In den meisten Modellen wird eine Fehlerkorrektur-Lernregel angewendet, die die

Assoziationsgewichte in Abhängigkeit von der Lernrückmeldung so modifiziert, dass die Fehler, die das Netzwerk macht, minimiert werden. Es handelt sich um assoziationalistische Modelle, die dem Gedanken ähnlichkeitsbasierter Kategorisierungstheorien (→ ähnlichkeitsbasierte Kategorisierungstheorien) verhaftet sind.

Klassische Sicht (classical view) – Gemäß dieser Sicht werden Kategorien in Form von notwendigen und hinreichenden Merkmalen (d.h. Definitionen) repräsentiert. Diese Sicht wird durch eine Reihe empirischer Befunde (→ Prototypentheorie) in Frage gestellt, wie etwa dem Befund, dass nicht alle Exemplare als gleich typisch für eine Kategorie angesehen werden.

Kombination, konzeptuelle (conceptual combination) – Kombination von einfachen Konzepten zu komplexeren Konzepten (z.B. Haus+Tier→Haustier).

Kontextmodell, generalisiertes (generalized context model) – Formal präzisierte Fassung einer Exemplartheorie (→ Exemplartheorien) der Kategorisierung.

Kreuzklassifikationen (cross classifications) – Nicht-hierarchische alternative Kategorisierungen des gleichen Objekts oder Ereignisses (z.B. Professorin, Psychologin, Tennisspielerin).

Natürliche Arten (natural kinds) – Kategorien, die sich auf in der Natur vorfindliche Objekte beziehen (z.B. Tiere, Mineralien, Pflanzen).

Prototypen-Sicht (prototype theories, probabilistic view) – Theorie der Kategorisierung, die davon ausgeht, dass Kategorienrepräsentationen nicht nur notwendige und hinreichende, sondern auch charakteristische Merkmale enthalten. Kategorien sind Abstraktionen über Lernexemplare, die charakteristische Merkmale enthalten (Prototyp). Neue Exemplare werden auf der Basis von Ähnlichkeitsvergleichen zu den Prototypen klassifiziert.

Taxonomien (taxonomies) – Hierarchische Anordnung von Kategorien in über- und untergeordnete Kategorien, wobei zwischen den Ebenen Klasseninklusion besteht (z.B. Tier-Hund-Pudel).

Theorien-Sicht (theory-based categorization view; theory theory) – Diese Sicht geht davon aus, dass die meisten Kategorien nicht als Bündel definatorischer oder charakteristischer Merkmale gespeichert werden, sondern als strukturierte Gebilde, die auch funktionale und kausale Relationen zwischen Merkmalen enthalten. Kategorisierung wird in Analogie zur Wissenschaft als Vergleich zwischen Daten (Lernexemplaren) und Theorien (Kategorien) konzeptualisiert. Empirisch stehen Untersuchungen zur Rolle des Vorwissens bei der Kategorisierung im Vordergrund. Formale Modelle dieser Theorie fehlen weitgehend.

((2)) Weiterführende Literatur

- Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Science*, 5, 204-210. – (Kurze Zusammenfassung des aktuellen Forschungsstands zu den neurobiologischen Grundlagen der Kategorisierung)
- Bloom, P. (2000). *How children learn the meanings of words*. Cambridge: MIT Press. – (Überblick über entwicklungspsychologische Befunde zum Verhältnis von Kategorien und Sprache)
- Bruner, J. S., Goodnow, J., & Austin, G. (1956). *A study of thinking*. New York: Wiley. – (Klassischer Text zur klassischen Sicht der Kategorisierung)
- Medin, D. L., & Atran, S. (Eds.)(1999). *Folkbiology*. Cambridge, MA: MIT Press. – (Überblick über psychologische und anthropologische Forschung zur Repräsentation biologischer Kategorien)
- Eckes, T. (1991). *Psychologie der Begriffe*. Göttingen: Hogrefe Verlag für Psychologie. – (Überblick über ähnlichkeitsbasierte Theorien)
- Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press. – (Darstellung der Theorien-Sicht anhand vorwiegend entwicklungspsychologischer Befunde)

- Heit, E. (2000). Properties of inductive reasoning. *Psychological Bulletin & Review*, 7, 569-592. – (Übersichtsartikel über induktives Denken mit Kategorien)
- Heydemann, M. (1998). *Lernen von Kategorien*. Leverkusen: Deutscher Universitätsverlag. – (Einführung in konnektionistische Lernmodelle der Kategorisierung)
- Hoffmann, J. (1993). *Vorhersage und Erkenntnis*. Göttingen: Hogrefe Verlag für Psychologie. – (Enthält ein Kapitel über Kategorien)
- Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: “Category-specific” neuropsychological deficits. *Behavioral and Brain Sciences*, 24, 453-509. – (Umfassender Überblick über neuropsychologische Befunde und Theorien)
- Medin, D. L., & Heit, E. J. (1999). Categorization. In: D. Rumelhart & B. Martin (Eds.), *Handbook of cognition and perception* (pp. 99-143). San Diego: Academic Press. – (Einführung in Computermodelle der Kategorisierung)
- Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology*, 51, 121-147. – (Übersicht über aktuelle Forschung zu Typen von Kategorien)
- Waldmann, M. R. (1996a). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning* (pp. 47-88). San Diego: Academic Press. – (Übersicht über die Forschung zu kausalen Kategorien)

((2)) Einschlägige Zeitschriften

Annual Review of Psychology

Cognition

Cognitive Psychology

Journal of Experimental Psychology: General

Journal of Experimental Psychology: Learning, Memory, and Cognition

Journal of Memory and Language

Memory & Cognition

Psychological Bulletin

Psychological Bulletin & Review

Psychological Review

Psychological Science

The psychology of learning and motivation

Trends in Cognitive Science

((3)) Danksagungen

Ich danke M. Heydemann, J. Hoffmann, S. Macho und den Herausgebern für ihre zahlreichen hilfreichen Kommentare.

((2)) Literatur

- Ahn, W. (1998). Why are different features central for natural kinds and artifacts?: The role of causal status in determining feature centrality. *Cognition*, *69*, 135-178.
- Ahn, W., & Kim, N. (2001). The causal status effect in categorization: An overview. In D. L. Medin (Ed.), *The psychology of learning and motivation*, Vol. 40, 23-65.
- Ahn, W., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology*, *41*, 361-416.
- Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, *120*, 3-19.
- Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, *98*, 409-429.
- Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition*, *13*, 263-308.
- Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Science*, *5*, 204-210.
- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442-481.

- Ashby, F. G., Waldron, E. M., Lee, W. W., & Berkman, A. (2001). Suboptimality in human categorization and identification. *Journal of Experimental Psychology: General*, 130, 77-96.
- Atran, S. (1990). *Cognitive foundations of natural history: Towards an anthropology of science*. Cambridge: Cambridge University Press.
- Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Science*, 21, 547-611.
- Baillargeon, R. (1998). Infants' understanding of the physical world. In M. Sabourin & F. Craik (Eds.), *Advances in psychological science, Vol. 2: Biological and cognitive aspects* (pp. 503-529). Hove: Psychology Press/Erlbaum.
- Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition, Vol. 1: Basic processes* (pp. 1-40). Hillsdale: Erlbaum.
- Barr, R. A., & Caplan, L. J. (1987). Category representations and their implications for category structure. *Memory & Cognition*, 15, 397-418.
- Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, 11, 211-217.
- Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 629-654.
- Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101-140). New York: Cambridge University Press.
- Barsalou, L. W. (1989). Intraconcept similarity and its implications for interconcept similarity. In S. Vosniadon & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 76-121). New York: Cambridge University Press.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-609.

- Barsalou, L. W., & Sewell, D. R. (1985). Contrasting the representation of scripts and categories. *Journal of Memory and Language*, *24*, 646-665.
- Barsalou, L. W., Huttenlocher, J., & Lamberts, K. (1998). Basing categorization on individuals and events. *Cognitive Psychology*, *36*, 203-272.
- Barton, M. E., & Komatsu, L. K. (1989). Defining features of natural kinds and artifacts. *Journal of Psycholinguistic Research*, *18*, 433-447.
- Beggs, J. M., Brown, T. H., Byrne, J. H., Crow, T., LeDoux, J. E., LeBar, K. S., & Thompson, R. F. (1999). Learning and memory: Basic mechanisms. In M. J. Zigmond, F. E. Bloom, S. C. Landis, & L. R. Squire (Eds.). *Fundamental neuroscience* (pp. 1411-1486). New York: Academic Press.
- Berlin, B., Breedlove, D. & Raven, P. (1973). General principles of classification and naming in folkbiology. *American Anthropologist*, *74*, 214-242.
- Bloom, P. (2000). *How children learn the meanings of words*. Cambridge: MIT Press.
- Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conception of time. *Cognitive Psychology*, *43*, 1-22.
- Bourne, L. E. (1970). Knowing and using concepts. *Psychological Review*, *77*, 546-556.
- Bowerman, M. (1996). Learning how to structure space for language. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and space*. Cambridge, MA: MIT Press.
- Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hillsdale, NJ: Erlbaum.
- Brooks, L. R., Norman, G. R., & Allen, S. W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General*, *120*, 278-287.
- Bruner, J. S., Goodnow, J., & Austin, G. (1956). *A study of thinking*. New York: Wiley.
- Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: the animate-inanimate distinction. *Journal of Cognitive Neuroscience*, *10*, 1-34.

- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, Mass.: MIT Press.
- Choi, S., McDaniel, M. A., & Busemeyer, J. R. (1993). Incorporating prior biases in network models of conceptual rule learning. *Memory & Cognition*, *21*, 413-423.
- Coley, J. D., Medin, D. L., & Atran, S. (1997). Does rank have its privilege? Inductive inferences within folkbiological taxonomies. *Cognition*, *64*, 73-112.
- Coley, J. D., Medin, D. L., Proffitt, J. B., Lynch, E., & Atran, S. (1999). Inductive reasoning in folkbiological thought. In D. L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 205-232). Cambridge, MA: MIT Press.
- Corter, J. E., & Gluck, M. A. (1992). Explaining basic categories: Feature predictability and information. *Psychological Bulletin*, *111*, 291-303.
- Damasio, A. R. (1990). Category-related recognition deficits as a clue to the neural substrates of knowledge. *Trends in Neurosciences*, *13*, 95-98.
- Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5-18.
- Eckes, T. (1991). *Psychologie der Begriffe*. Göttingen: Hogrefe Verlag für Psychologie.
- Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*, 107-140.
- Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology*, *18*, 500-549.
- Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, *120*, 339-357.
- Fiske, S. T., Neuberg, S. L., Beattie, A. E., & Milberg, S. J. (1987). Category-based and attribute-based reactions to others: Some informational conditions of stereotyping and individuating processes. *Journal of Experimental Social Psychology*, *23* (5), 399-427.

- Flannagan, M. J., Fried, L. S., & Holyoak, K. J. (1986). Distributional expectations and the induction of category structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 241-256.
- Fodor, J. A. (1975). *The language of thought*. New York: Crowell.
- Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford: Clarendon Press.
- Fried, L. S., & Holyoak, K. J. (1984). Induction of category distributions: A framework for classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 234-257.
- Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*, 183-208.
- Gentner, D. (1981). Some interesting differences between nouns and verbs. *Cognition and Brain Theory*, *4*, 161-178.
- Gentner, D., & Markman, A. B. (1997). Structural alignment in analogy and similarity. *American Psychologist*, *52*, 45-56.
- Gluck, M. A., & Bower, G. H. (1988a). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*, 227-247.
- Gluck, M. A., & Bower, G. H. (1988b). Evaluating an adaptive network model of human learning. *Journal of Memory and Language*, *27*, 166-195.
- Goldstone, R. L. (1994a). The role of similarity in categorization: Providing a groundwork. *Cognition*, *52*, 125-157.
- Goldstone, R. L. (1994b). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, *123*, 178-200..
- Goldstone, R. L. (1996). Isolated and interrelated concepts. *Memory & Cognition*, *24*, 608-628.
- Goldstone, R. L., & Barsalou, L. (1998). Reuniting perception and conception. *Cognition*, *65*, 231-262.

- Goldstone, R. L., & Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General, 130*, 116-139.
- Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (Ed.), *Problems and projects*. New York: Bobbs-Merrill.
- Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.
- Hagmayer, Y., & Waldmann, M. R. (2000). Simulating causal models: The way to structural sensitivity. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society* (pp. 214-219). Mahwah, NJ: Erlbaum.
- Hahn, U., & Chater, N. (1997). Concepts and similarity. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 43-92). Hove: Psychology Press.
- Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18*, 441-461.
- Hampton, J. A. (1996). Conjunctions of visually based categories : Overextension and compensation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22*, 378-396.
- Heit, E. (1994). Models of the effects of prior knowledge on category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20*, 1264-1282.
- Heit, E. (1997). Knowledge and concept learning. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 7-41). Hove: Psychology Press.
- Heit, E. (2000). Properties of inductive reasoning. *Psychological Bulletin & Review, 7*, 569-592.

- Heydemann, M. (1998). *Lernen von Kategorien*. Leverkusen: Deutscher Universitätsverlag.
- Hirschfeld, L. A. (1994). The child's representation of human groups. In D. L. Medin (Ed.), *The psychology of learning and motivation, vol. 31* (pp. 131-185). New York: Academic Press.
- Hirschfeld, L. A. (1996). *Race in the making: Cognition, culture, and the child's construction of human kinds*. Cambridge, MA.: MIT Press.
- Hirschfeld, L. A., & Gelman, S. A. (Eds.)(1994). *Mapping the mind: Domain specificity in cognition and culture*. Cambridge: Cambridge University Press.
- Hoffmann, J. & Zießler, C. (1983). Objektidentifikation in künstlichen Begriffshierarchien. *Zeitschrift für Psychologie, 191*, 135-167.
- Hoffmann, J. & Zießler, M. (1982). Begriffe und ihre Merkmale. *Zeitschrift für Psychologie, 190*, 46-77.
- Hoffmann, J. (1986). *Die Welt der Begriffe. Psychologische Untersuchungen zur Organisation des menschlichen Wissens*. Weinheim: Beltz-Verlag.
- Hoffmann, J. (1993). *Vorhersage und Erkenntnis*. Göttingen: Hogrefe Verlag für Psychologie.
- Homa, D., & Vosburgh, R. (1976). Category breadth and the abstraction of prototypical information. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 322-330.
- Hull, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychological Monographs, 28*, (1, whole No. 123).
- Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: “Category-specific” neuropsychological deficits. *Behavioral and Brain Sciences, 24*, 453-509.
- Hunt, E., & Agnolli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review, 98*, 377-389.

- Jones, S. S., & Smith, L. B. (1993). The place of perceptions in children's concepts. *Cognitive Development, 8*, 113-140.
- Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review, 88*, 199-227.
- Keil, F. C. (1989). *Concepts, kinds, and conceptual development*. Cambridge, MA: The MIT Press.
- Kemler-Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior, 23*, 734-759.
- Kersten, A. W. (1998). A division of labor between nouns and verbs in the representation of motion. *Journal of Experimental Psychology: General, 127*, 34-54.
- Klix, F., van der Meer, E., Preuß, M. & Wolf, M. (1987). Über Prozeß- und Strukturkomponenten der Wissensrepräsentation des Menschen. *Zeitschrift für Psychologie, 195*, 39-61.
- Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: Parallel brain systems for item memory and category level knowledge. *Science, 262*, 1747-1749.
- Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*, 1399-1402.
- Knowlton, B. J., Squire, L. R., Paulsen, J. S., Swerdlow, N. R., Swenson, M., & Butters, N. (1996). Dissociations within nondeclarative memory in Huntington's disease. *Neuropsychology, 10*, 538-548.
- Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning and Memory, 1*, 106-120.
- Komatsu, L. K. (1992). Recent view of conceptual structure. *Psychological Bulletin, 112*, 500-526.
- Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*, 22-44.

- Kruschke, J. K. (1993). Three principles for models of category learning. In G. V. Nakamura, R. Taraban & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol.29 : Categorization by humans and machines.*(pp. 283-326). San Diego: Academic Press.
- Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45*, 812-863..
- Kuhl, J. & Waldmann, M. R. (1985). Handlungspsychologie: Vom Experimentieren mit Perspektiven zu Perspektiven fürs Experimentieren. *Zeitschrift für Sozialpsychologie, 16*, 153-181.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Landau, B. (1982). Will the real grandmother please stand up? The psychological reality of dual meaning representations. *Journal of Psycholinguistic Research, 11*, 47-62.
- Levinson, S. C. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and space*. Cambridge, MA: MIT Press.
- Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General, 130*, 3-28.
- Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology, 39*, 821-831.
- Lynch, E. B., Coley, J. D., & Medin, D. L. (2000). Tall is typical: Central tendency, ideal dimensions and graded category structure among tree experts. *Memory & Cognition, 28*,41-50.
- Macho, S. (1997). The effect of relevance shifts in category acquisition: A test of neural networks. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 23*, 30-53.
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51*, 93-120.

- Macrae, C. N., Milne, A. B. & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, *66*, 37-47.
- Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of classification. *Perception & Psychophysics*, *53*, 49-70.
- Malt, B. C. (1994). Water is not H₂O. *Cognitive Psychology*, *27*, 41-70.
- Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology*, *29*, 85-148.
- Malt, B. C., & Johnson, E. C. (1992). Do artifact concepts have cores. *Journal of Memory and Language*, *31*, 195-217.
- Malt, B. C., & Smith, E. E. (1983). Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior*, *23*, 250-269.
- Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, *40*, 230-262.
- Marcus, G. (1998). Rethinking eliminative connectionism. *Cognitive Psychology*, *37*, 243-282.
- Markman, A. B., & Makin, V. S. (1998). Referential communication and category acquisition. *Journal of Experimental Psychology: General*, *127*, 331-354.
- Massaro, D. W., & Friedman, D. (1990). Models of integration given multiple sources of information. *Psychological Review*, *97*, 225-252.
- McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 24, pp. 109-165). New York: Academic Press.
- McCloskey, M., & Glucksberg, S. (1978). Natural categories: Well-defined or fuzzy sets? *Memory & Cognition*, *6*, 462-472.

- McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology, 11*, 1-37.
- McKinley, S. C., & Nosofsky, R. M. (1995). Investigations of exemplar and decision-bound models in large-size, ill-defined category structures. *Journal of Experimental Psychology: Human Perception & Performance, 21*, 128-148.
- Medin, D. L., & Heit, E. J. (1999). Categorization. In: D. Rumelhart & B. Martin (Eds.), *Handbook of cognition and perception* (pp. 99-143). San Diego: Academic Press.
- Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-195). Cambridge: Cambridge University Press.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 5*, 207-238.
- Medin, D. L., & Schwanenflugel, P. J. (1981). Linear separability in classification learning. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 355-368.
- Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology, 20*, 158-190.
- Medin, D. L., Altom, M. W., & Murphy, G. L. (1984). Given versus induced category representations: Use of prototype and exemplar information in classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10*, 333-352.
- Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 8*, 37-50.
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254-278.

- Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology, 51*, 121-147.
- Medin, D. L., Lynch, E. B., Coley, J. D., & Atran, S. (1997). Categorization and reasoning among tree experts: Do all roads lead to Rome? *Cognitive Psychology, 32*, 49-96.
- Medin, D. L., Ross, B. H., & Markman, A. B. (2001). *Cognitive psychology* (3rd ed.). Fort Worth : Harcourt College Publishers.
- Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19*, 242-279.
- Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinate-, basic-, and superordinate-level categories. *Child Development, 53*, 258-266.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.
- Minsky, M. L., & Papert, S. A. (1969/1988). *Perceptrons*. Cambridge, MA.: MIT Press.
- Mooney, R. J. (1993). Integrating theory and data in category learning. In G. V. Nakamura, R. Taraban & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol.29 : Categorization by humans and machines.*(pp. 189-218). San Diego: Academic Press.
- Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science, 12*, 529-562.
- Murphy, G. L., & Allopenna, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20*, 904-919.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92*, 289-316.
- Murphy, G. L., & Wisniewski, E. J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15*, 572-586.

- Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*, 291-310.
- Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 104-114.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39-57.
- Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 700-708.
- Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception & Performance*, *17*, 3-27.
- Nosofsky, R. M. (1992). Exemplars, prototypes, and similarity rules. In A. F. Healy, S. M. Kosslyn & R. M. Shiffrin (Eds.), *From Learning theory to connectionist theory: Essays in honor of William K. Estes* (Vol. 1, pp. 149-167). Hillsdale, NJ: Erlbaum.
- Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of “multiple system” phenomena in perceptual categorization. *Psychonomic Bulletin & Review*, *7*, 375-402.
- Nosofsky, R. M., & Kruschke, J. K. (1992). Investigations of an exemplar-based connectionist model of category learning. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 28, pp. 207-250). New York: Academic Press.
- Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random-walk model of speeded classification. *Psychological Review*, *104*, 266-300.
- Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. T. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, *22*, 352-369.

- Nosofsky, R. M., Kruschke, J. K., & McKinley, S. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 211-233.
- Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*, 53-79.
- Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*, 185-200.
- Pauen, S. (2000). Early differentiation within the animate domain: Are humans something special? *Journal of Experimental Child Psychology*, *75*, 134-151.
- Pazzani, M. J. (1991). Influence of prior knowledge on concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 416-432.
- Pearl, J. (2000). *Causality. Models, reasoning, and inference*. Cambridge: Cambridge University Press.
- Pinker, S. (1994). *The language instinct*. New York: Harper Collins.
- Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology*, *13*, 564-574.
- Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353-363.
- Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, *83*, 304-308.
- Proffitt, J. B., Medin, D. L., & Coley, J. D. (2000). Expertise and category-based induction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 811-828.
- Pullum, G. K. (1991). *The great Eskimo vocabulary hoax, and other irreverent essays on the study of language*. Chicago: University of Chicago Press.
- Putnam, H. (1975). The meaning of "meaning". In H. Putnam (Ed.), *Mind, language, and reality. Philosophical papers* (Vol. 2, pp. 215-271). London: Cambridge University Press.

- Quine, W. V. O. (1960). *Word and object*. Cambridge, Mass.: MIT Press.
- Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, *130*, 323-360.
- Rescorla, R. A. (1973). Evidence for the "unique stimulus" account of configural conditioning. *Journal of Comparative and Physiological Psychology*, *85*, 331-338.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II. Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.
- Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21-59). Cambridge: Cambridge University Press.
- Rips, L. J. (2001). Necessity and natural categories. *Psychological Bulletin*, *127*, 827-852.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573-605.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382-439.
- Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 491-502.
- Ross, B. H. (1997). The use of categories affects classification. *Journal of Memory and Language*, *37*, 240-267.

- Ross, B. H., & Murphy, G. L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology*, 38, 495-553.
- Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology*, 15, 346-378.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, and the PDP Research Group (Ed.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1, pp. 318-362). Cambridge, MA: The MIT Press.
- Sartori, G., & Job, R. (1988). The oyster with four legs: a neuropsychological study on the interaction of visual and semantic information. *Cognitive Neuropsychology*, 5, 105-132.
- Schyns, P. G., & Rodet, L. (1997). Categorization creates functional features. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 23, 681-696.
- Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (1998). The development of features in object concepts. *Behavioral and Brain Sciences*, 21, 1-54.
- Shepard, R. N. (1958). Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. *Journal of Experimental Psychology*, 55, 509-523.
- Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237, 1317-1323.
- Slobin, D. (1996). From “thought and language” to “thinking for speaking.” In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). Cambridge, MA: Cambridge University Press.
- Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology*, 25, 231-280.
- Sloman, S. A. (1994). When explanations compete: The role of explanatory coherence on judgments of likelihood. *Cognition*, 52, 1-21.

- Smith, E. E. (1995). Concepts and categorization. In E. E. Smith & D. N. Osherson (Eds.), *Thinking. An invitation to Cognitive Science* (2nd ed.)(pp. 3-33). Cambridge, MA: MIT Press.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA.: Harvard University Press.
- Smith, E. E., & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, 8, 337-361.
- Smith, E. E., & Sloman, S. A. (1994). Similarity- versus rule-based categorization. *Memory & Cognition*, 22, 377-386.
- Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science*, 12, 485-527.
- Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition*, 65, 167-196.
- Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 24, 1411-1436.
- Smith, J. D., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 26, 3-27.
- Smith, J. D., Murray, M. J., & Minda, J. P. (1997). Straight talk about linear separability. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 23, 659-680.
- Smith, L. B., & Samuelson, L. K. (1997). Perceiving and remembering: Category stability, variability and development. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 161-195). Hove: Psychology Press.
- Sokal, R. R. (1974). Classification: Purposes, principles, progress, prospects. *Science*, 185, 1115-1123.
- Solomon, K. O., Medin, D. L., & Lynch, E. (1999). Concepts do more than categorize. *Trends in Cognitive Sciences*, 3, 99-105.

- Spelke, E. S., & Tsivkin, S. (2001). Language and number: A bilingual training study. *Cognition*, *78*, 45-88.
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*, 605-632.
- Spelke, E. S., Phillips, A., & Woodward, A. L. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack & A. J. Premack (Eds.), *Causal cognition: a multidisciplinary debate* (pp. 44-78). Oxford: Oxford University Press.
- Springer, K., & Keil, F. C. (1989). On the development of biologically specific beliefs: the case of inheritance. *Child Development*, *60*, 637-648.
- Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and applications. *Journal of Personality and Social Psychology*, *37*, 1660-1672.
- Sternberg, R. J., & Ben-Zeev, T. (2001). *Complex cognition. The psychology of human thought*. New York: Oxford University Press.
- Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgment. *British Journal of Psychology*, *54*, 101-114.
- Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*, 457-482.
- Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind* (pp. 19-136). Oxford: Oxford University Press.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327-352.
- Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, *113*, 169-193.
- Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Science*, *5*, 244-252.
- Van der Meer, E. (1986). What is invariant in event-related knowledge representation? In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities*:

- Mechanisms and performances* (Part A, pp. 339-352). Amsterdam: North Holland.
- Waldmann, M. R. (1990). *Schema und Gedächtnis*. Heidelberg: Asanger Verlag.
- Waldmann, M. R. (1996a). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning* (pp. 47-88). San Diego: Academic Press.
- Waldmann, M. R. (1996b). Wissensgeleitetes Lernen. In J. Hoffmann & W. Kintsch (Hrsg.), *Enzyklopädie - Kognition, Band II: Lernen* (S. 323-355). Göttingen: Hogrefe Verlag.
- Waldmann, M. R. (1997). Wissen und Lernen. *Psychologische Rundschau*, 48, 84-100.
- Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 53-76.
- Waldmann, M. R. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychological Bulletin & Review*, 8, 600-608.
- Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*, 121, 222-236.
- Waldmann, M. R., & Martignon, L. (1998). A Bayesian network model of causal learning. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 1102-1107). Mahwah, NJ: Erlbaum.
- Waldmann, M. R., Holyoak, K. J., & Fratianne, A. (1995). Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General*, 124, 181-206.
- Warrington, E. K., & Shallice, T. (1984). Category-specific semantic impairment. *Brain*, 107, 829-854.

- Wattenmaker, W. D. (1995). Knowledge structures and linear separability: Integrating information in object and social categorization. *Cognitive Psychology*, 28, 274-328.
- Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology*, 18, 158-194.
- Weinert, F. E. & Waldmann, M. R. (1988). Wissensentwicklung und Wissenserwerb. In H. Mandl & H. Spada (Hrsg.), *Wissenspsychologie. Ein Lehrbuch* (pp. 161-199). München: Urban & Schwarzenberg.
- Wellmann, H. M., & Gelman, S. A. (1992). Cognitive development: foundational theories of core domains. *Annual Review of Psychology*, 62, 1070-1090.
- Whorf, B. L. (1956). *Language, thought, and reality*. Cambridge, MA: MIT Press.
- Wisniewski, E. J., & Gentner, D. (1991). On the combinatorial semantics of noun pairs: minor and major adjustments to meaning. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 241-284). North Holland: Elsevier.
- Wisniewski, E. J., & Love, B. C. (1998). Properties versus relations in conceptual combination. *Journal of Memory and Language*, 38, 177-202.
- Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, 18, 221-281.
- Wittgenstein, L. (1953/1971). *Philosophische Untersuchungen*. Frankfurt/M.: Suhrkamp.
- Zaki, S. R., & Homa, D. (1999). Concepts and transformational knowledge. *Cognitive Psychology*, 39, 69-115.