

Funke, J. & Spering, M. (in Vorb.). Methoden der Denk- und Problemlöseforschung. In J. Funke (Ed.), *Denken und Problemlösen (=Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie II: Kognition, Band 8: Denken und Problemlösen)*. Göttingen: Hogrefe.

## **Methoden der Denk- und Problemlöseforschung**

Joachim Funke & Miriam Spering

Stand: 14.10.2004

### **1. Einleitung: Gegenstandsbereich und methodische Zugänge**

Denken und Problemlösen zählen seit der Antike zu den psychischen Funktionen, die über das bloße Erkennen hinausgehen und zu Erkenntnissen beitragen – von der Objektwahrnehmung über dessen Einordnung in begriffliche und funktionale Kategorien bis hin zum Verstehen, wofür dieses Objekt geschaffen ist, was damit getan werden kann und wie es in die aktuellen Pläne und Ziele des Organismus eingebaut werden kann. Natürlich können auch Prozesse und Abläufe zum Objekt derartiger Betrachtungen werden (Nachdenken über das eigene Denken, Metakognition).

Dass Menschen denken, erschließen wir zunächst aus der Selbstbeobachtung. Insbesondere dann, wenn wir uns in einer zunächst unlösbaren Situation befinden, in der eine intendierte Zielerreichung nicht gelingt (also ein „Problem“ vorliegt), machen uns die in solchen Situationen ablaufenden Prozesse der Lösungssuche darauf aufmerksam, dass wir mit dem Handeln innehalten müssen und ein paar gute Ideen brauchen, um dem Ziel näher zu kommen. Dass Selbstbeobachtung nicht der einzig denkbare Weg ist, der uns darauf bringt, dass wir denken, macht Gopnik (1993) deutlich: Nach ihrer Meinung erschließt sich uns unser eigenes Innenleben nicht aus der Introspektion, sondern primär aus der Beobachtung anderer Menschen und der Übertragung dieser Beobachtungen auf uns. Selbstbeobachtung ist allerdings eine Methode, die den Ansprüchen objektiver, intersubjektiv überprüfbarer Datenerhebungen nicht genügt. Hierfür muss die Fremdbeobachtung verwendet werden, die im Bereich des Denkens und Problemlösens ebenfalls zum Einsatz kommt.

Im vorliegenden Abschnitt werden wir zunächst den Gegenstand der Denk- und Problemlöseforschung reflektieren und uns Gedanken über die Wechselwirkung von Gegenstand und Methode machen. Es folgt als Hauptteil eine Darstellung von Methoden zur Präsentation, Diagnose und Auswertung kognitiver Prozesse, bevor abschließende Bemerkungen über zukünftige Entwicklungen im Bereich der Methoden der Denk- und Problemlösepsychologie gemacht werden.

### *1.1. Was ist Gegenstand der Denk- und Problemlöseforschung?*

Die Denk- und Problemlöseforschung befasst sich mit einem Gegenstand, der in Hinblick auf die Methoden aus verschiedenen Perspektiven betrachtet werden kann. Auf der phänomenologischen Seite sind die Bewusstwerdung von und Reflektion über bestimmte Objekte oder Prozesse markante Merkmale der Erste-Person-Erfahrung. Sich als Denkender zu erleben heißt, das Handeln erst einmal zurückzustellen und einen Plan für dasselbe zu machen. Die von Graumann (1964; 1965, S. 19f.) näher beschriebenen Aspekte des Denkens (Vergegenwärtigung des nicht sinnlich Präsenten, Ordnungsleistung durch Begriffsbildung, Orientierung der Innerlichkeit, Selektivität der hergestellten Assoziationen, Urteils- und Entscheidungsleistung, Reflexivität als Option zum Denken über das Denken) zeigen die phänomenologische Vielfalt dieses Zustands. Von phänomenologischer Seite aus ist der nahe liegende methodische Zugang die Introspektion.

Der Zugang über die behavioristische Seite verweist auf die von Denkenden gezeigten Verhaltensäußerungen (zu denen durchaus auch das Sprachverhalten zu rechnen ist). Aus dieser Perspektive geht es vor allem darum, die zwischen einem gegebenen Ausgangszustand und einem gewünschten Zielzustand liegenden Schritte möglichst lückenlos zu dokumentieren. Verständlicherweise werden aus methodischer Sicht Problemstellungen gesucht, deren Lösung nicht vollständig „im Kopf“ vollzogen werden können, sondern im Gegenteil die problemlösende Person zwingen, möglichst viele Zwischenschritte zu externalisieren. Der aus dieser Sicht nahe liegende methodische Zugang ist daher die Verhaltensbeobachtung, die dann möglich ist, wenn sich Teile des Denk- und Problemlöseprozesses in äußerlich sichtbaren Aktivitäten niederschlagen.

Aus physiologischer Sicht sind Denkprozesse vorrangig mit Aktivitäten im frontalen bzw. präfrontalen Kortex verknüpft. Dieser Teil des Gehirns, der etwa 30% seines Volumens in

Anspruch nimmt, ist – so Grafman (1994; 1995) – weitaus weniger gut erforscht als etwa der visuelle Kortex. Bildgebende Verfahren, aber auch klassische Verfahren wie das EEG (z.B. Jung-Beeman et al., 2004) versprechen hier neue Erkenntnisse, über die im weiteren Text noch zu sprechen sein wird (vgl. Kapitel 2.2.3). Der aus dieser Sicht nahe liegende methodische Zugang ist daher die möglichst detaillierte Erfassung neuronaler Aktivitäten und deren zeitlicher Verlauf während eines Denk- und Problemlöseprozesses. Als Herausforderung erweist es sich jedoch, aus dem Ort und der Art neuronaler Aktivität auf den Inhalt korrespondierender Denkprozesse zu schließen.

## *1.2. Wie beeinflussen die Methoden den Gegenstand?*

Methodische Zugänge haben immer auch Rückwirkungen auf den Gegenstandsbereich. Gigerenzer (1988; 1991) weist auf die enge Verbindung zwischen neu entdeckten Werkzeugen und deren metaphorischer Nutzung im Rahmen psychologischer Modellbildung hin, wie etwa bei der Entdeckung der Regressionsrechnung und der nachfolgenden Annahme, dass menschliche Kognition gut durch ein Regressionsmodell beschrieben werden kann.

Die Beeinflussung des Gegenstands durch die Methoden seiner Untersuchung kann für die Denk- und Problemlöseforschung am Beispiel der Untersuchungsparadigmen für einfaches Problemlösen demonstriert werden. Die von den Gestaltpsychologen verwendeten Einsichtsprobleme (siehe weiter unten) von der Art „Bilde aus sechs Streichhölzern vier gleichseitige Dreiecke“, wie sie etwa Duncker (1935/1974) verwendete, weisen die Charakteristik einer „Alles-oder-Nichts“-Lösung auf: Entweder findet man die richtige Antwort oder nicht. Wenn solche Probleme zum Einsatz kommen, die in einem einzigen Schritt zu lösen sind, hinterlässt der Lösungsprozess weniger deutliche Spuren als bei sequentiellen Problemen. Das Phänomen des „Aha-Effekts“, einer plötzlichen Einsicht, resultiert aus der Art des Problemraums: Wenn ich z.B. erkenne, dass zur Lösung eines bestimmten Streichholz-Problems die dritte Dimension erforderlich ist, „fällt der Groschen“. Erkenne ich dieses Prinzip nicht, tappe ich weiter im Dunkeln. Der Problemraum umfasst also im Wesentlichen die diskreten Zustände „falsche Lösungsvorschläge“ (als Sammelbegriff für alle Zustände, bei denen die Streichhölzer flächig arrangiert werden bzw. dort, wo sie räumlich werden, noch nicht die Lösung darstellen) oder „richtig“, aber keine differenzierten Zwischenzustände auf dem Weg vom Ausgangs- zum Zielzustand.

Die von den Informationsverarbeitungstheoretikern favorisierten sequentiellen Probleme wie der „Turm von Hanoi“ (siehe weiter unten) machen dagegen die Entdeckung eines „Aha-

Effekts“ eher unwahrscheinlich, da typischerweise eine schrittweise Annäherung an das Ziel zu beobachten ist und die Entdeckung eines „Tricks“ nicht stattfindet. Bei dieser Art von Problemen ist die Zielerreichung kontinuierlich gestuft, da man sich dem Ziel annähern oder sich weiter davon entfernen kann. Der Problemraum enthält differenzierbare Zwischenstufen. Mit der Entscheidung für einen der beiden eben genannten methodischen Zugänge ist somit bereits vorentschieden, was entdeckt werden kann: Im Streichholz-Paradigma werden Einsichtsphänomene regelrecht provoziert, die bei sequentiellen Problemstellungen kaum auftreten werden. Dafür ist dort die allmähliche, schrittweise Annäherung an den Zielzustand zu erkennen, die bei Einsichtsproblemen abrupt (oder gar nicht) erfolgt. Wähle ich statt eines einfachen Problems eine komplexe Anforderung (z.B. ein computersimuliertes Szenario, siehe weiter unten), tauchen andere, neue Phänomene auf, die in einer einfachen Problemsituation nicht zu entdecken sind.

Mit der Wahl eines Untersuchungsparadigmas ist somit bereits eine Vorentscheidung darüber getroffen, welches Phänomen überhaupt aufgedeckt und beschrieben werden kann. In gewissem Sinn ist diese Aussage eine Verallgemeinerung der These „Messung ist Modellbildung“ von Gigerenzer (1981), der nachweist, dass bereits mit der Wahl eines Skalenniveaus (nominale, ordinale, intervallskalierte Messung) eine Vorentscheidung über die Konzeptualisierung des Gegenstandsbereichs getroffen wird. Insofern ist von einer fundamentalen Beeinflussung des Gegenstands durch die Methoden zu sprechen. Die Entscheidung für eine bestimmte Untersuchungsmethode legt den Raum aufzudeckender Phänomene fest – eine triviale Aussage, die aber dennoch nicht vergessen werden sollte.

### *1.3. Wie beeinflusst der Gegenstand die Methoden?*

Nach dem eben Gesagten verwundert es kaum, wenn die Beziehung zwischen Gegenstand und Methode auch in der anderen Richtung einflussreich ist. Dies soll durch die Kontrastierung zweier unterschiedlicher Gegenstandsverständnisse illustriert werden.

Ausgehend von einem Gegenstandsverständnis „Denken ist ein intern ablaufender Prozess“ (z.B. Bourne, Ekstrand & Dominowski, 1971) wird der Fokus weg von der Umwelt hin zu den intern ablaufenden Prozessen verschoben. Die Beschränkung auf kognitive Prozesse führt damit notwendigerweise auch zur Beschränkung auf bestimmte Untersuchungsmethoden, die sich zur Erfassung der „im Kopf“ ablaufenden Prozesse beziehen. Die – nicht ganz alltägliche – Vorstellung, wonach Denkprozesse weniger im Kopf als vielmehr in der dinglichen Umwelt ablaufen, würde dagegen andere Untersuchungsmethoden favorisieren, die sich eher auf die

Erfassung von Umweltaspekten beziehen. So wie die ökologische Perspektive des Gedächtnisses (Graumann, 1997) nach anderen als den klassischen Prüfverfahren des Gedächtnisses ruft (z.B. Reproduktion und Rekognition von Wortlisten), würde eine ökologische Perspektive des Denkens ebenfalls ein spezielles Vorgehen verlangen: Neben einer Erfassung kognitiver Prozesse wären die Umweltbedingungen genauer zu dokumentieren und Methoden zur Abbildung vom Interaktionsprozess bereitzustellen, wie dies auch von Vertretern der „situated cognition“ (Clancey, 1997; Clark, 1997; Norman, 1993) oder Vertretern „ökologischer Rationalität“ (Gigerenzer & Selten, 2001) gefordert wird.

### *1.4. Methodologische Divergenzen*

Unterschiedliche methodologische Standpunkte gibt es auf allen Feldern der Psychologie. In einem gewissen Sinn sind diese Divergenzen treibende Kräfte der wissenschaftlichen Entwicklung, da in diesen Kontroversen häufig Zuspitzungen und Klärungen erfolgen, die als Fortschritt bei der Sammlung von Erkenntnissen über den jeweiligen Gegenstand gelten können. Auf einige dieser divergenten Standpunkte soll nachfolgend kurz eingegangen werden.

Intro- versus Extraspektion. Über die richtige Art des Zugangs zu Denk- und Problemlöseprozessen ist bereits in der Vergangenheit heftig gestritten worden. Dabei ging es vor allem um die Frage, inwiefern die Methode der Introspektion Daten produziert, die dem Anspruch der Wissenschaftlichkeit insofern Rechnung tragen, als sie objektiv, reliabel und valide sind. Introspektion nennt man den Prozess, durch den eine Person ihre Aufmerksamkeit auf die eigenen inneren Zustände richtet, in denen sie sich gerade befindet. Der Begründer der experimentellen Psychologie, Wilhelm Wundt, lehnte in seinem 1871 in Leipzig gegründeten ersten psychologischen Labor den Einsatz von Introspektion aus grundsätzlichen Erwägungen heraus ab, da wichtige Bestimmungsmerkmale eines Experiments wie willkürliche Bedingungsvariation oder Replizierbarkeit nicht gewährleistet seien. Auch die Erinnerungstäuschungen bei einer Variante der Introspektion, der Retrospektion, bei der im Nachhinein über einen abgelaufenen Denkvorgang zu berichten war, hielt er für so störend, dass er diese Verfahren rundheraus ablehnte (Wundt, 1907). Dagegen war es für Vertreter der „Würzburger Schule“ wie Külpe oder Bühler selbstverständlich, „Experimente ohne Instrumente“ durchzuführen (Wundt nannte sie despektierlich „Scheinexperimente“), indem sie dafür speziell geschulte Testpersonen baten, über ihre Einfälle und Erlebnisse bei der

Bearbeitung einfacher Aufgaben (Übersetzungen lateinischer Texte, Erklärung der Bedeutung bestimmter Aphorismen etc.) im Nachhinein zu berichten („Ausfragemethode“).

In heutiger Zeit sind verbale Selbstauskünfte unter definierten Bedingungen als Erkenntnisquelle akzeptiert. Dabei geht es allerdings weniger um Introspektion als vielmehr um lautes Denken. In Kapitel 2.2.1 werden wir im Detail darauf eingehen.

Experimentelle versus nicht-experimentelle Vorgehensweise. Bei der Untersuchung von Denk- und Problemlöseprozessen wird immer wieder die Frage aufgeworfen, inwiefern die dabei ablaufenden Prozesse einer experimentellen Manipulation zugänglich seien (vgl. Funke, 2003, S.139ff.). Die wohl heftigste Kritik an konventioneller Versuchsplanung und -auswertung im Bereich der Denkpsychologie stammt von Dörner (1989), der in seiner Karikatur eines varianzanalytischen Vorgehens im wesentlichen die Schwächen schlechter Experimente vorführt, die ohne theoretische Modellannahmen „blind“ Daten erheben und scheitern. Die von ihm später vorgetragene Argumentation (Dörner, 1992) richtete sich darauf, dass nicht die Methode der Dissektion (der analytischen Zerlegung), sondern nur diejenige der Kondensation (eine Technik der Reduktion, die eine vergrößerte Konstellation des „Urbilds“ liefert) zum Verständnis komplexer Systeme weiter helfe, weil ansonsten wichtige Interaktionen und Randbedingungen ignoriert würden. Diese methodologische Kontroverse soll hier nicht vertieft werden (siehe dazu Funke, 2003, Kap. 4.3). Die in diesem Beitrag berichteten experimentellen Befunde belegen u.E deutlich den Wert experimenteller Analysen im Bereich von Denken und Problemlösen.

Problemlösen versus Problemkonstruieren. Fast wie selbstverständlich beschränken Problemlöseforschende ihre Aktivitäten auf das Lösen von Problemen. Denken und Problemlösen werden als Herangehensweisen gesehen, aufgetretene Hindernisse im Dienste der Erreichung eines angestrebten Zieles zu überwinden. Dabei wird selten gefragt, wie Probleme entstehen. Stellt man sich dagegen auf den Standpunkt, dass Probleme gemacht und nicht gegeben sind (vgl. Funke, 2003, S. 19), rückt damit auch der Prozess der Problemstehung in den Fokus methodischer Erfassung. Ein möglicher Zugang hierfür ist von Nährer (1988) vorgeschlagen worden, der seinen Versuchspersonen die Aufgabe gestellt hat, unterschiedlich schwierige Probleme zu konstruieren. Das Interesse lag dabei auf den verwendeten Konstruktionskriterien, nicht auf der Lösung dieser Probleme.

Wissensgesteuerte versus heuristikgeleitete Verarbeitung. Mit dieser Dichotomie ist die Frage verbunden, ob man bei der Untersuchung von Denk- und Problemlöseprozessen seinen

Schwerpunkt auf das zur Problemlösung nötige Wissen richtet, wie dies etwa in der Expertiseforschung geschieht (vgl. Reimann, 1998), oder ob man seinen Schwerpunkt auf Strategien und Lösungsmethoden richtet, wie dies etwa in Arbeiten zur kognitiven Modellierung von Problemlöseprozessen zu finden ist (vgl. Gonzales, Lerch & Lebiere, 2003).

Tierische versus menschliche Intelligenz. Köhlers Schimpansenstudien (1921) haben den intelligenten Werkzeuggebrauch bei unseren Artverwandten eindrucksvoll demonstriert – aber kann man Denkprozesse tatsächlich am Tier untersuchen? Wie Hauser (2001) ausführt, lassen sich viele der beim Menschen eingesetzten Prüfverfahren auch im Tierversuch verwenden, insbesondere die Parallelen zur Säuglings- und Kleinkindforschung sind eindrucksvoll (Call & Tomasello, 1995; Tomasello, 2002). Bertrand Russell hat die Tierpsychologie des 20. Jahrhunderts dadurch charakterisiert, dass die Beobachtungsobjekte jeweils nationale Charakterzüge aufwiesen: „Die von den Amerikanern untersuchten Tiere laufen rastlos umher, machen sich zu schaffen und mobilisieren dabei eine unglaubliche Menge Energie, um schließlich das gewünschte Resultat rein zufällig zu erreichen. Die von den Deutschen beobachteten Tiere sitzen ruhig da und denken nach, um dann die Lösung der Probleme aus ihrer innersten Tiefe hervorzuziehen“ (zitiert nach Celli, 1986, S. 82).

Experiment versus Simulation. Eine letzte Dichotomie berührt den Streit, inwiefern maßgebliche Erkenntnisse nur aus sorgfältig kontrollierten Experimenten oder auch aus Simulationsstudien mit „künstlichen Seelen“ zu gewinnen sind. Die Vertreter Kognitiver Modellierung haben in den letzten Jahrzehnten eine eigenständige Methodologie entwickelt, die für einen kontrollierten Abgleich zwischen Simulation und Realität sorgt (vgl. Opwis & Spada, 1994) und damit diese Kontroverse entschärft.

## **2. Methoden zur Präsentation, Diagnose, und Auswertung kognitiver Prozesse**

In diesem Kapitel werden in einem ersten Teil zunächst statische und dynamische Aufgabentypen beschrieben, bei denen technische Hilfsmittel zur Präsentation des Stimulusmaterials üblich sind. Ein zweiter Teil befasst sich mit den diagnostischen Zugängen zu den während der Bearbeitung ablaufenden Denk- und Problemlöseprozessen. Hierzu zählen Selbstauskünfte, Verhaltensdaten und psychophysiologische Messungen. Schließlich geht ein dritter Teil auf spezielle Auswertungsverfahren ein. Dazu zählen kognitive

Modellierung, die Erzeugung synthetischer Versuchspersonen, Markov-Analysen sowie die latente semantische Analyse.

## *2.1. Aufgabentypen und Untersuchungsparadigmen*

Mit dem Titel wird auf die Tatsache Bezug genommen, dass die empirisch orientierte Denk- und Problemlöseforschung auf bestimmte Paradigmen zurückgreift, die heute zumeist unter kontrollierten Laborbedingungen präsentiert werden und technische Hilfsmittel zur Darbietung des Stimulusmaterials und/oder zur Registrierung des Verhaltens verwenden (mehr zur apparativen Psychologie allgemein bei Albert & Gundlach, 1997). Auf einer ersten Ebene unterscheidet man dabei einfache von komplexen Aufgabenstellungen. Innerhalb dieser beiden Klassen gibt es jeweils spezielle Aufgabentypen, von denen jeweils typische Vertreter beschrieben werden.

### 2.1.1. Einfache Aufgabentypen

Einfache Aufgabentypen unterscheiden sich von komplexen durch den geringen Instruktionsaufwand und die geringe Menge an Wissen, die zu ihrer Bearbeitung nötig ist. Man könnte in Bezug auf das benötigte Wissen auch von semantisch armen im Unterschied zu semantisch reichen Problemstellungen reden. Zudem benötigen einfache Aufgabentypen in aller Regel kurze Bearbeitungszeiten, die im Bereich bis zu 30 Minuten liegen, während komplexe Aufgabentypen eher Stunden oder Tage benötigen.

Zu den einfachen Aufgabentypen zählen klassische Denksportaufgaben (Einsichtsprobleme), kryptarithmetische Probleme sowie sequentielle Problemstellungen.

#### **Klassische Denksportaufgaben (Einsichtsprobleme)**

In der Frühzeit der Problemlöseforschung waren Denksportaufgaben und Einsichtsprobleme das Untersuchungsmaterial der Wahl (ausführlich dargestellt bei Knoblich & Öllinger, in diesem Band). Klassische Einsichtsprobleme sind von Duncker (1935/1974) im Rahmen seiner „Psychologie des produktiven Denkens“ vorgelegt worden. Den Problemlöseprozess hat er vor allem an zwei Problemstellungen näher untersucht:

1. Das praktische Bestrahlungsproblem: „Gesucht ein Verfahren, um einen Menschen von einer inoperablen Magengeschwulst zu befreien mit Hilfe von Strahlen, die bei genügender Intensität organisches Gewebe zerstören – unter Vermeidung einer

Mitzerstörung der umliegenden gesunden Körperpartien“ (S. 1) Dieses Problem hat er als „praktisch“ bezeichnet, weil die Leitfrage „Wie erreiche ich etwas?“ lautet.

2. Das theoretische Beweisproblem: „Gesucht eine Begründung dafür, dass alle sechsstelligen Zahlen vom Typus abcabc, z.B. 276276 durch 13 teilbar sind“ (S. 1). Dieses Problem hat er als „theoretisch“ bezeichnet, weil die Leitfrage „Wie, woraus sehe ich ein?“ lautet.

Beide Einsichtsprobleme suchen nach dem Grund einer vorgegebenen Folge, das praktische Problem nach einem Realgrund, das theoretische nach dem logischen Grund. Grund-Folge-Beziehungen (also die Suche nach Kausalität) spielen daher eine wichtige Rolle. Duncikers Erhebungsmethode war nicht die Selbstbeobachtung (Introspektion), wie sie z.B. von Vertretern der Würzburger Schule (Külpe, Marbe, Selz) praktiziert wurde, sondern das „Laut Denken“, ein Verfahren, bei dem „der laut Denkende unmittelbar auf die Sache gerichtet (bleibt), sie ... nur gleichsam ‚zu Worte kommen‘ lässt“ (S. 2). Seine Analyse der Lösungsvorschläge für das Bestrahlungsproblem zeigt, dass die verschiedenen Ideen nach ihrem „Funktionalwert“ geordnet werden können. Duncker nennt diese Liste „Lösungsstammbaum“.

Einsichtsprobleme unter Verwendung einer „Streichholz-Arithmetik“ untersucht Knoblich (z.B. 1999). Ein Einsichtsproblem tritt dann auf, wenn nach der ersten Exploration ein Hindernis auftaucht („impasse“, Sackgasse) und die Lösung subjektiv unmöglich erscheint (vgl. Metcalfe, 1986). Aus diesen mentalen Sackgassen kommt man nur durch eine Änderung der Problemrepräsentation wieder heraus. Zwei Beispiele aus der Arbeit von Knoblich, Ohlsson, Haider und Rhenius (1999) sollen dies verdeutlichen (vgl. Abbildung STREICH).

---

Hier bitte Abbildung STREICH einfügen

---

Problemstellungen im Bereich der Streichholzarithmetik bestehen aus falschen arithmetischen Ausdrücken, die aus römischen Zahlen (I, II, III usw.), arithmetischen Operationen (+, -) sowie dem Gleichheitszeichen zusammengesetzt sind. Durch Verlegung eines der Streichhölzer soll aus dem falschen ein richtiger Ausdruck gemacht werden. In Abbildung STREICH (a) ist etwa aus der IV eine VI zu machen. Dies ist die typische Repräsentation, bei der die Zahlwerte als variabel, die arithmetischen Operationen dagegen als konstant angesehen werden. Lockert man diese Randbedingung auf und lässt zu, dass auch die Operatoren variabel sein dürfen, kann man die Aufgabe in Abbildung STREICH (b) dadurch lösen, dass aus dem „+“ ein „=“ gemacht wird. Neben der Lockerung von Randbedingungen

lässt sich die Problemrepräsentation aber auch durch die Dekomposition von Chunks (= zu Gruppen zusammengefasste Einzelelemente) verändern. So werden „schwache“ Chunks wie „IV“ von „starken“ Chunks wie „X“ unterschieden, deren Zerlegung in „/“ und „\“ wegen fehlender Bedeutung der Einzelteile schwerer fällt.

Knoblich et al. können auf der Basis dieser beiden postulierten Mechanismen zur Veränderung der Problemrepräsentationen spezifische Vorhersagen über unterschiedliche Aufgabenschwierigkeiten und differentielle Transfereffekte machen, die sich in den berichteten Experimenten bestätigten. Auch begleitende Blickbewegungsanalysen (vgl. Knoblich, Ohlsson & Raney, 2001) konnten die theoretischen Annahmen in drei Punkten bestätigen: (a) an den Sackgassen-Stellen kommt es zu weniger Blickbewegungen und längeren Fixationszeiten, (b) aufgrund arithmetischen Vorwissens neigt man dazu, die Zahlenwerte und nicht die Operatoren als die veränderlichen Größen zu betrachten, und (c) führen Veränderungen der Problemrepräsentation aufgrund von Lockerungen der Randbedingungen und aufgrund von Dekompositionen perzeptueller Chunks zu Veränderungen in der Aufmerksamkeitsverteilung.

Die Streichholz-Arithmetik stellt eine interessante Problemstellung dar, anhand derer elementare Denkprozesse untersucht werden können. Gerade in Verbindung mit Blickbewegungsanalysen gewinnt dieses einfache Paradigma seinen Reiz, da sich Prozesstheorien überprüfen lassen, die sonst kaum der Empirie zugänglich wären. Festzuhalten bleibt aber auch, dass die geringe Menge an Wissen, die diese Probleme zu ihrer Bearbeitung erfordern, zwar forschungstechnisch einen Vorteil darstellt, in Hinblick auf die Aussagekraft für Problemlösen in Alltagssituationen aber wohl kaum aussagekräftig ist, da dort meist wesentlich mehr Weltwissen in die Problemlösung einfließt.

Ein anderer Zugang zu Einsichtsprozessen stammt aus der Analyse von Lösungsprozessen bei Anagramm-Aufgaben. Anagramme stellen Buchstabenverdrehungen dar, die korrigiert werden müssen (z.B. HEULSC -> SCHULE). Die Schwierigkeit kann in diesem Fall durch die Zahl verstellter Buchstaben sowie die Gesamtzahl der Buchstaben beeinflusst werden.

Dass es sich bei Anagramm-Lösungen um plötzliche Einsichtsprozesse und nicht um eine allmähliche, sequentielle Annäherung handelt, haben Metcalfe und Wiebe (1987) durch Erfassung von „Heiss-Kalt-Urteilen“ (eine alle 10 bis 15 Sekunden erhobene Angabe darüber, wie nah sich ein Problemlöser der Lösung fühlt) belegt: Während beim Lösen von

Gleichungen die Wärmeurteile allmählich anstieg, blieben sie bei Anagrammen konstant niedrig und stiegen erst kurz vor der Lösung steil an.

### **Kryptarithmetische Probleme**

Kryptarithmetische Probleme verlangen eine Dekodierung von Buchstaben in Zahlen unter Verwendung arithmetischer Prozeduren. Abbildung KRYPT veranschaulicht zwei Beispiele dieser Puzzles.<sup>1</sup>

---

Hier bitte Abbildung KRYPT einfügen

---

Kryptarithmetische Probleme können in ihrer Schwierigkeit manipuliert werden durch die Anzahl der Buchstaben, für die bereits eine richtige Ziffer vorgegeben wird, sowie (in geringem Maße) durch Anzahl und Verschiedenartigkeit der Buchstaben, die zu bearbeiten sind. Bei  $n$  verschiedenen Buchstaben ( $n \leq 10$ ) schrumpft durch die Restriktion, dass jede Ziffer höchstens einmal zugeordnet werden darf, von  $10^n$  auf  $10!/(10-n)!$  mögliche Zustände (nach Clearwater, Huberman & Hogg, 1991). Bei Newell und Simon (1972, S. 143f.) findet sich eine ausführliche Analyse dieses Aufgabentyps sowie eine detaillierte Beschreibung der Produktionsregeln, mit denen das Lösungsverhalten abgebildet werden kann. Ähnliche Überlegungen hat bereits Bartlett (1958, pp. 49-71) unter der Überschrift „Denken in geschlossenen Systemen“ angestellt.

### **Sequentielle Problemstellungen**

Sequentielle Problemstellungen sind solche, zu deren Lösung notwendigerweise eine Serie von Schritten vollzogen werden muss, die sich in äußerlich sichtbaren Veränderungen des Zustandsraums niederschlagen.

Turm von Hanoi: Das Problem besteht im Wesentlichen darin, eine gegebene Menge unterschiedlich großer, konzentrischer Scheiben, die der Größe nach geordnet auf einem Ausgangsstab gesteckt sind, unter Nutzung eines Hilfsstabs auf einen Zielstab zu bewegen. Dabei müssen nur zwei Regeln eingehalten werden: (1) Es darf jeweils nur eine Scheibe bewegt werden, (2) es darf nie eine größere Scheibe auf eine kleinere gelegt werden. Abbildung WEG illustriert das Problem durch eine Darstellung des gesamten Problemraums, d.h. aller möglichen Stellungen für den (einfachen) Fall von drei Scheiben.

---

<sup>1</sup> Das Trainer-Beispiel übernehmen wir von Claus Möbus (Universität Oldenburg), der es anlässlich aktueller Vorgänge im Jahre 2001 (Affäre Christof Daum) präsentiert hat.

---

Hier bitte Abbildung WEG einfügen

---

Der gezeigte Problemraum erklärt die Attraktivität des Problems für die Denkforschung: Jeder einzelne Zug des Problemlösers lässt sich als Bewegung in diesem Problemraum darstellen und zugleich daraufhin bewerten, wie weit entfernt er vom geforderten Zielzustand liegt. Außerdem lässt sich für jede beliebige Position zeigen, welcher Pfad am schnellsten zum Ziel führt. Der Problemlöseprozess wird als Trajektorie (eine zeitliche Abfolge von Zuständen) in diesem Problemraum beschreibbar.

Das Problem geht auf eine alte Legende zurück, die besagt, dass nach erfolgreicher Umlegung der 100 Scheiben des Turms von Hanoi das Ende der Welt kommen werde. Um diese Aussage verstehen zu können, muss der Problemraum noch etwas genauer angeschaut werden. Tatsächlich weist er eine charmante Symmetrie auf, wenn man die Zahl der Scheiben erhöht: Der jeweils nächst größere Problemraum besteht aus drei Teilen, die jeder für sich genommen den nächst niedrigeren Problemraum verkörpern. Dies wird in Abbildung TVH\_LÖS veranschaulicht.

---

Hier bitte Abbildung TVH\_LÖS einfügen

---

Die Einfachheit der geforderten Operationen ist dort – wie auch in „normalen“ Anwendungen – Vor- und Nachteil zugleich: Vorteil insofern, als tatsächlich eine einfache, leicht vermittelbare Problemstellung vorliegt, deren Umfang ohne großen Aufwand verändert werden kann (bereits wenige Scheiben mehr erhöhen die Zugzahl exponentiell); Nachteil insofern, als minimales Wissen verlangt wird und die Anforderung daher in ihrer Validität für Alltagssituationen eingeschränkt ist.

Sequentielle Vorhersageprobleme. Ein Aufgabentyp, den Hussy (1975) zur Untersuchung von informationsgenerierenden und –reduzierenden Prozessen eingesetzt hat, verlangt von der Testperson die schrittweise Vorhersage einer regelhaften Zeichenfolge. Typischerweise wird dazu auf einem Monitor eine Matrix mit einem entsprechenden Symbolrepertoire präsentiert, die in einer bestimmten Reihenfolge nacheinander aufleuchten. Während am Anfang nur geraten werden kann, offenbaren sich der Testperson während der laufenden Vorhersagen Stück für Stück die zugrunde liegenden Gesetzmäßigkeiten der Zeichenfolge und können immer besser antizipiert werden. Diesen Vorgang der Anpassung subjektiver Auftrittswahrscheinlichkeiten an die objektiven Wahrscheinlichkeiten kann man informationstheoretisch exakt quantifizieren und Übergänge zwischen Zeichen unterscheiden, die wenig oder viel Information enthalten; auch die Bildung von Superzeichen (Erkennen von

Abhängigkeiten höherer Ordnungen) kann auf diese Weise operationalisiert werden (siehe hierzu Hussy, 1982).

Superhirn. Ein sequentielles logisches Problem unter dem Namen „Superhirn“ (Mastermind) ist von Hussy (1989) erprobt worden. Dabei geht es für die Testperson darum, ein kombinatorisches Problem zu lösen: Aus einer gegebenen Ziffernmenge ist schrittweise eine verborgene Ziffernkombination herauszufinden, in dem die bei jedem Rateversuch gegebene Rückmeldung über richtige Ziffern und richtige Positionen genutzt wird. In einem Vergleich verschieden alter Stichproben (8, 12, 20, 40, 60 Jahre alte Vpn) konnte Hussy zeigen, dass bis zum Alter von 20 eine zunehmende Verfügbarkeit passender Reduktionsstrategien festzustellen ist, während die 60jährigen erhebliche Leistungsdefizite infolge mangelnder Nutzung solcher Strategien zeigten. Neuwirth (1982) hat eine Übersicht über erfolgreiche Strategien vorgelegt (mehr zu „Superhirn“ bei Best, 1990; Fritz & Funke, 1990; Hussy, 1991; Irving, 1978; Jeck, 1997; Laughlin, Lange & Adamopoulos, 1982).

### 2.1.2. Komplexe Aufgabentypen<sup>2</sup>

In die Rubrik der komplexen Aufgabentypen werden solche Paradigmen eingeordnet, zu deren Präsentation heutzutage Computer herangezogen werden müssen, damit die vielfältigen dynamischen Abhängigkeiten zwischen den beteiligten Variablen berechnet werden können. Um den Unterschied zu den im vorangegangenen Abschnitt dargestellten einfachen Aufgabentypen zu verdeutlichen: Natürlich kann auch der „Turm von Hanoi“ heute als Computerprogramm präsentiert werden, aber dies muss nicht sein. Die gleich vorgestellten Aufgabentypen wie z.B. „Lohhausen“ sind dagegen ohne Computer nicht realisierbar, sieht man einmal vom „Papiercomputern“ wie in Vesters' Spiel „Ökolopoly“ ab (Vester, 1983). Komplexe Probleme weisen Strukturen auf, die nicht in zwei oder drei Sätzen erschöpfend dargelegt werden können. Obwohl es sich wegen ihrer Implementation auf einem Computer um „geschlossene“ Probleme (im Unterschied zu offenen Problemen, bei denen prinzipiell immer neue Aspekte hinzukommen können) mit einem prinzipiell endlichen Problemraum handelt, ist dieser doch so umfangreich und vielfältig, dass er in der zur Verfügung stehenden Zeit nur ansatzweise exploriert werden kann.

---

<sup>2</sup> Die Darstellung dieses Teils ist in Auszügen angelehnt an die in Kapitel 4 und 5 gegebene Beschreibung komplexer Szenarien in Funke (2003, S. 146f.).

## **Inhaltlich orientierte Planspiele und Szenarien**

Die ersten Studien zum komplexen Problemlösen versuchten in eindrucksvoller Weise die komplexe Wirklichkeit ins Labor zu holen: Mit dem von Dörner programmierten Szenario „Schneiderwerkstatt“ (auch als „Tailorshop“ bezeichnet) wurde eine Management-Umgebung geschaffen, in der Probanden Rohmaterial einkaufen sowie mittels vorhandener Maschinen und Arbeiter Hemden produzieren und anschließend verkaufen mussten. Mit der ebenfalls in seiner Arbeitsgruppe erstellten Simulation „Lohhausen“ ging es von der Ebene einer kleinen Produktionsfirma noch eine Stufe weiter, zur Nachbildung einer kleinen Kommune gleichen Namens, der die Versuchspersonen als Bürgermeister vorstanden und um deren Wohlergehen sie sich zu bemühen hatten. Schließlich wurde als dritter Prototyp die Simulation eines afrikanischen Stammes („Tanaland“, „Dagu“ bzw. „Moro“) produziert, bei der die Versuchsperson die Rolle eines Entwicklungshelfers einzunehmen hatte. Alle drei Szenarien stellen Instantiierungen des Konzepts „Komplexes Problemlösen“ dar, wie es in den 1970er Jahren von Dörner begründet wurde. Die amerikanische Forschung mit semantisch reichhaltigen Problemen (z.B. Bhaskar & Simon, 1977) ist damit nur bedingt vergleichbar, da deren Szenarien wesentliche Merkmale komplexer Probleme wie Intransparenz und Polytelie (vgl. Funke, in diesem Band) nicht realisiert haben.

Lohhausen. Das Bürgermeister-Szenario „Lohhausen“ stellt vom Umfang her mit rund 2000 beteiligten Variablen eines der aufwändigsten Szenarien (gemessen an der Zahl beteiligter Variablen) dar. Die 48 studentischen Versuchspersonen konnten nicht direkt mit dem System interagieren, sondern die Versuchsleiter übernahmen eine zentrale Rolle: Sowohl die Wünsche der Versuchspersonen als auch das Feedback des Systems konnten nur über diese Zwischenstation vermittelt werden – und zwischen Eingabe und Ausgabe konnte (da das System auf einem entfernt stehenden Großrechner implementiert war) durchaus einiges an Zeit verstreichen.

Der Name „Lohhausen“ steht für die Simulation einer Kleinstadt gleichen Namens, in der Versuchspersonen die Rolle des Bürgermeisters für einen Simulationszeitraum von 10 Jahren übernehmen sollten und die Gemeinde so gut wie möglich leiten sollten (Dörner, Kreuzig, Reither & Stäudel, 1983). Nach der bei Dörner (1981, S. 165) gegebenen Beschreibung hat die Stadt etwa 3500 Einwohner und lebt hauptsächlich von der stadteigenen Uhrenfabrik. Neben der Stadtverwaltung gibt es Arztpraxen, Einzelhandelsgeschäfte, eine Bank, Schulen, Kindergärten usw. Im Computermodell wurden nicht nur die ökonomischen Beziehungen abgebildet, sondern auch soziale, demographische und psychologische Variablen (z.B.

Zufriedenheit der Einwohner). Versuchspersonen konnten in vielfältiger Weise in das System eingreifen: Sie konnten die Produktions- und Absatzpolitik der städtischen Fabrik beeinflussen, sie konnten Steuersätze ändern, Arbeitsplätze für Lehrer und Erzieherinnen schaffen, Arztpraxen einrichten und verpachten, Wohnungsbau betreiben, für Freizeiteinrichtungen sorgen usw. Die in Abbildung LOH dargestellte Grobstruktur mit gut 40 Variablen stellt eine erhebliche Vereinfachung des rund 2000 Variablen umfassenden Gesamtsystems dar.

---

Hier bitte Abbildung LOH einfügen

---

Nach Sternberg (1995) liegt eine Besonderheit europäischer im Vergleich zu amerikanischer Forschung im Umgang mit komplexen Problemen darin, dass hier (wie auch in anderen Studien europäischer Herkunft) Novizen als Versuchspersonen herangezogen wurden, die mit ihren ganz alltäglichen Routinen und ohne jede Schulung oder Vorbereitung Führungsaufgaben zu übernehmen hatten. Neuere amerikanische Überblicke zum computergestützten Messen der Problemlösekompetenz (z.B. Baker & O'Neil, 2002) ignorieren übrigens zumeist die europäischen Studien.

Schneiderwerkstatt. Das System „Schneiderwerkstatt“ stellt ein frühkapitalistisch organisiertes Unternehmen dar, in dem Rohmaterial von Arbeitern an Produktionsmaschinen zu Hemden verarbeitet wird, die auf dem Markt verkauft werden müssen. Das System besteht insgesamt aus 24 Variablen, von denen 11 durch Maßnahmen der Probanden direkt beeinflusst werden können. Kernvariable des Systems ist das Merkmal „Kapital (Bilanzwert)“, das mit 15 der 24 Variablen in Verbindung steht. Aufgabe des Problemlösers ist es, die „Schneiderwerkstatt“ über einen entsprechend ausgedehnten Simulationszeitraum hinweg so zu führen, dass nachhaltig Gewinn erwirtschaftet wird (Putz-Osterloh & Lüer, 1981, S. 313-315). Ohne Eingriffe in das System würde die „Schneiderwerkstatt“ in Kürze Konkurs anmelden müssen, da die laufenden Kosten (Lagerkosten, Lohnkosten, Miete, etc.) schnell zu negativen Zahlen führen. Dies lässt sich vermeiden, indem Rohmaterial eingekauft wird, die Maschinen gewartet werden und den Arbeitern ein vernünftiger Lohn gezahlt wird. Außerdem muss der Hemdenpreis marktfähig gemacht werden. Abbildung SCHNEI zeigt die Variablen der Schneiderwerkstatt und ihre Verknüpfungen.

---

Hier bitte Abbildung SCHNEI einfügen

---

In seiner ursprünglichen Form existierte dieses von Dietrich Dörner programmierte System als Programm auf einem Taschenrechner TI 59, das allerdings 1982 durch den Erstautor

dieses Beitrags in ein BASIC-Programm übersetzt und damit auf dem Bildschirm eines Kleinrechners bedienbar gemacht wurde. Dies ist deswegen bedeutsam, da die Ursprungsvariante, die in den Arbeiten von Putz-Osterloh (1981b; Putz-Osterloh & Lüer, 1981) zum Einsatz kam, keine direkten Eingriffe der Versuchspersonen in das Programm erlaubte, sondern zwischengeschaltete Versuchsleiter benötigte. Wie die bei Funke (1983) publizierten Details des Systems deutlich machen, sind nicht nur an mehreren wichtigen Stellen Zufallsfunktionen implementiert, sondern auch nicht-lineare Funktionen, die eine zumindest lineare Optimierung des Systems unmöglich machen. In der Zwischenzeit ist das ursprüngliche Programm „Schneiderwerkstatt“ in vielen Varianten erschienen (z.B. Funke, 1982; Süß & Faulhaber, 1990) – eine Reihe von Szenarien mit anderen Titeln gehen in ihrer Grundstruktur auf das Originalsystem zurück (z.B. „Heizölhandel“ und „Textilfabrik“).

Die ersten publizierten Arbeiten zur „Schneiderwerkstatt“ zeigten geringe bis keine Korrelationen zwischen Leistungsparametern des Szenarios und IQ-Werten (Putz-Osterloh, 1981b; Putz-Osterloh & Lüer, 1981). Dieser Befund trat später auch bei anderen Szenarien auf, wird jedoch heute anders erklärt. Die Befundlage hat sich inzwischen durch weitere Studien dahingehend geändert, dass man durchaus von substantiellen Zusammenhängen ausgeht (für einen Überblick siehe Süß, 1999).

Moro. Das überaus beliebte Programm „Moro“ (vgl. etwa Endres & Putz-Osterloh, 1994; Jansson, 1994; Kotkamp, 1999; Putz-Osterloh, Bott & Houben, 1988; Rigas, 2000; Schaub & Strohschneider, 1992; Schmuck & Strohschneider, 1995; Stäudel, 1987; Strohschneider, 1986, 1994; Strohschneider & Güss, 1999; Tisdale, 1990) simuliert in holzschnittartiger Weise die Situation eines Nomadenstammes der Sahel-Zone in der südlichen Sahara. In diesem Szenario haben Akteure praktisch diktatorische Möglichkeiten, in den Gang des Geschehens einzugreifen. Zugleich bekommen sie auf Nachfrage hin Rückmeldung über die Effekte ihres Handelns und können daher auch Einblicke in mögliche nicht bedachte Implikationen ihrer Entscheidungen nehmen.

Das Szenario „Moro“ verlangt von den Probanden, sich in die Rolle eines Entwicklungshelfers in Westafrika hinein zu versetzen. Dort leben am Südrand der Sahara die Moros, ein Halbnomadenstamm, der von Rinderzucht und Hirseanbau lebt. Die Stammesgruppe umfasst zu Beginn der Simulation etwa 650 Menschen, denen es nicht besonders gut geht. Die Rinder, von denen sie sich ernähren, sind klein und oft schlecht ernährt. Dem Nomadenbrauchtum entsprechend lassen die Moros ihre Rinderherden in der Nähe der einzelnen Wasserstellen weiden, wechseln aber die Stellen häufiger. Eine bestimmte

Anzahl von Rindern, die der eigenen Ernährung dienen, wird jährlich geschlachtet. All das ermöglicht ein Leben hart am Existenzminimum. Zudem wird ein großer Teil des Viehbestandes jährlich durch die Rinderschlafkrankheit vernichtet, die durch die Tse-Tse-Fliege verbreitet wird. Auch von Hungerkatastrophen bleibt der Stamm nicht verschont. Etwa 5 Quadratkilometer des insgesamt 1000 Quadratkilometer umfassenden Gebietes konnten bislang in Hirsefelder umgewandelt werden, die der Ernährung dienen. Medizinische Versorgung jenseits der Naturheilkunde fehlt den Moros.

Die „Moro“-Simulation besteht aus mehreren untereinander verbundenen Modellen (Wasserkreislauf, Vegetation, Rindervermehrung, Tse-Tse-Fliegen, Bevölkerung, Ernährung, Arbeit, Geld). Eine aktuelle Variante von „Moro“ ist als allgemein zugängliche Diagnose- und Trainingssoftware unter dem Namen „Moroland“ bzw. „Syrene“ erschienen (Lantermann, Döring-Seipel, Schmitz & Schima, 2000).

Reither (1981) verglich in seiner Studie Novizen und Entwicklungshilfe-Experten beim Bearbeiten eines derartigen Entwicklungshilfe-Szenarios und musste erstaunt feststellen, dass neben der Ignoranz von Novizen gegenüber wichtigen Problemen im Szenario auch eine Blindheit der Experten zu konstatieren war, die in der mangelnden Anpassung vertrauter Routinen an sich verändernde Bedingungen bestand.

Weitere inhaltlich orientierte Szenarien. Einen Überblick über die vielfältigen Szenarien findet man z.B. bei Funke (1992, S. 8-9) oder Funke (2003, Kap. 5). Die große Vielfalt rührt daher, dass recht bald schon Anwendungen solcher Szenarien im Rahmen von Personalauswahl und –entwicklung gut bezahlt wurden und sich recht schnell ein Markt dafür fand, der zumindest in der Anfangszeit kaum Qualitätsstandards kannte. Dies hat zu entsprechenden Warnungen geführt (z.B. Funke, 1998). Über den Einsatz im Rahmen von Personalauswahl und –entwicklung findet man mehr in den Beiträgen von Geilhardt und Mühlbradt (1995), Strauß und Kleinmann (1995), Wagener (2001) oder (kritisch) Kersting (1999).

### **Formal orientierte Szenarien**

Als Reaktion auf die anfängliche ad-hoc-Konstruktion komplexer Szenarien hat es Gegenbewegungen gegeben, die auf die Notwendigkeit einer präzisen Formalisierung komplexer Anforderungen hingewiesen haben (Funke, 1984; Hübner, 1987; Thalmaier, 1979). Die semantische Einkleidung wird dabei im Interesse einer präziseren Gestaltung der Aufgabenumgebung (und damit der Möglichkeit zu einer präziseren Gütebestimmung) der

formalen Struktur nachgeordnet. Der Schwerpunkt liegt auf einer Objektivierung des „task environments“.

Konkret bedeutet dies für die Systemkonstruktion: Zunächst wird eine formale Struktur festgelegt, für die dann im nächsten Schritt eine passende Semantik gesucht wird (oder auch nicht – den Versuchspersonen wird oftmals die Künstlichkeit des zu explorierenden Systems explizit mitgeteilt). Im Fall der bereits oben beschriebenen Forschung mit realitätsnahen Szenarien ist der Weg genau umgekehrt: Zunächst wird ein möglichst interessanter Realitätsbereich ausgewählt, dem dann später eine formale Modellierung unterlegt wird. Diese versucht die wichtigen Eigenschaften des Realitätsbereichs abzubilden, was aber keinesfalls immer gut gelingt – tatsächlich wurden artifizielle Systeme (Programme) geschaffen, die unter dem Deckmantel der Realitätsnähe kuriose Ergebnisse produzierten. So ist es etwa beim ursprünglichen „Schneiderwerkstatt“-Szenario möglich, durch einen starken Einsatz von Lieferwagen, an den der Programmierer wohl nicht gedacht hatte, die Nachfrage erheblich zu steigern.

Aus heutiger Sicht haben sich zwei Formalismen bewährt, die zwei der wichtigsten Anforderungsdimensionen eines komplexen Problems (Vernetztheit und Dynamik) realisieren helfen: zum einen der Ansatz linearer Strukturgleichungen für Systeme mit kontinuierlichen Variablen (vgl. Funke, 1985, 1993), zum anderen die Theorie finiter Automaten für Systeme mit diskreten Variablen (vgl. Buchner & Funke, 1993; Funke & Buchner, 1992). Beide Ansätze werden nachfolgend genauer vorgestellt.

### Lineare Strukturgleichungsmodelle

Typischerweise wird den Versuchspersonen als Einführung in die Problemstellung erklärt, dass sie in einem unbekanntem System einzelne Variablen (sog. exogene Variablen) verändern können, die dann andere Variablen (sog. endogene Variablen) beeinflussen. Diese Beziehungen sind zu eruieren, wobei neben den Wirkungen der exogenen auf die endogenen Variablen auch Wirkungen innerhalb der endogenen Variablen auftreten können und herausgefunden werden müssen.

Die Grundstruktur für ein derartiges einfaches lineares System ist in Abbildung LIN\_SYS exemplarisch an einem Vier-Variablen-System dargestellt (vgl. Vollmeyer & Funke, 1999).

---

Hier bitte Abbildung LIN\_SYS einfügen

---

Hier sind A und B die exogenen Variablen, die auf die endogenen Variablen Y und Z wirken. Die Zahlen an den Pfeilen geben Gewichte an, mit denen die jeweilige exogene auf die entsprechende endogene Variable wirkt. Formal ist das System durch zwei Gleichungen beschreibbar (für jede endogene Variable wird genau eine Gleichung benötigt), nämlich:

$$Y_{t+1} = 2 * A_t \quad (1)$$

$$Z_{t+1} = 3 * A_t - 2 * B_t + 0,5 * Y_t + 0,9 * Z_t \quad (2)$$

Dabei gibt der tiefer gestellte Index jeder Variablen den jeweiligen Zeitpunkt ( $t$  bzw.  $t+1$ ) des Systems an, der in diskreten Stufen getaktet wird. Aus Gleichung (1) ergibt sich, dass sich der Wert der Variable  $Y$  zum Zeitpunkt  $t+1$  errechnet aus dem Wert der Variable  $A$  zum Zeitpunkt  $t$ , multipliziert mit 2. Die Gleichung (2) verdeutlicht, dass die endogene Variable  $Z$  zum einen von den beiden exogenen Variablen  $A$  und  $B$  mit Gewicht 3 bzw. -2 beeinflusst wird, zum anderen zusätzlich von der anderen endogenen Variable  $Y$  (Gewicht 0,5) sowie von ihrem eigenen vorangegangenen Zustand abhängt (Gewicht 0,9).

Durch die Art der Darbietung des Systems am Computer ist erkennbar, welche Variablen manipuliert werden können. Nicht erkennbar ist, ob und welche exogene Variable welche endogene Variable beeinflusst. In manchen Systemen wirken endogene Variablen auf andere endogene Variablen (sog. Nebenwirkungen; im Beispielsystem als Wirkung von  $Y$  auf  $Z$  implementiert, vgl. Abbildung LIN\_SYS) oder eine endogene Variable verändert sich konstant, auch wenn das System nicht manipuliert wird (sog. Eigendynamik; im Beispielsystem als Wirkung von  $Z$  auf sich selbst implementiert, vgl. Abbildung LIN\_SYS). Durch gezielte Verwendung und Manipulation solcher Systemmerkmale können beliebig komplexe lineare Gleichungssysteme im Rahmen des sog. Dynamis-Ansatzes (siehe z.B. Funke, 1985, 1993; Funke & Steyer, 1985) konstruiert werden.

Die durch Zielvorgaben gesteuerten Anforderungen an die Versuchsperson lassen sich abstrakt beschreiben als (1) Wissenserwerb und (2) Wissensanwendung. Wissenserwerb bezieht sich dabei auf die Identifikation des Systems, mit dem man es zu tun hat, Wissensanwendung dagegen auf die Kontrolle dieses Systems im Hinblick auf ein durch den Versuchsleiter gesetztes Ziel. Für die diagnostische Situation ist es wichtig, die beiden Aspekte voneinander zu trennen. Wie dies geschehen kann, wird in den nächsten Abschnitten geschildert.

Identifikation eines Systems (Wissenserwerb). Da in vielen alltäglichen Situationen auf Vorwissen zurückgegriffen werden kann, ist bereits der Wissenserwerb von der Nutzung

dieses Wissens geleitet. Besonders in Fällen, wo nur ein Bearbeitungsdurchgang möglich ist (solche Situationen wurden in vielen Untersuchungen zum Komplexen Problemlösen realisiert, Bsp. „Lohhausen“), spielt die Anwendung von Vorwissen natürlich eine entscheidende Rolle bei den ersten Schritten des Wissenserwerbs (vgl. Süß, 1999).

Die Wissenserwerbssituation ist eine komplexe Lernsituation, in der die Vernetztheit und die Dynamik des jeweiligen Systems festzustellen ist. Die Erfassung der strukturellen Aspekte eines Systems (= Vernetztheit) kann allerdings nicht losgelöst von der Erfassung der prozessualen Aspekte dieses Systems (= Dynamik) erfolgen, da sich die Variablen des Systems nur im zeitlichen Verlauf analysieren lassen. Bezogen auf lineare Strukturgleichungen bedeutet Identifikation des Systems: finde heraus, wie die exogenen Variablen auf die endogenen wirken und wie die endogenen Variablen sich untereinander beeinflussen. Hierfür ist eine entsprechende Identifikationsstrategie zu entwerfen, da bei komplexeren Systemen eine einzelne Intervention und das Muster der daraus resultierenden Änderungen nicht aussagekräftig sind. Erst aufeinander abgestimmte Muster von Eingriffen (z.B. Strategie der isolierenden Bedingungsvariation) erlauben kausale Interpretationen der Datenvektoren.

Bei linearen Strukturgleichungssystemen kann das Identifikationsproblem in mehrere Teilziele zerlegt werden: Festzustellen sind (a) die Existenz einer Relation zwischen zwei Variablen, (b) deren Richtung (positives oder negatives Vorzeichen), sowie (c) deren genaue Spezifikation in Form des Gewichtungsfaktors, mit dem die erste Variable auf die zweite oder auch auf weitere Variablen einwirkt. Mit diesen Annahmen verbunden sind Vorstellungen darüber, wie die genannten Merkmale repräsentiert sein könnten (vgl. Funke, 1992, Kap. 3) und wie entsprechend diesen Repräsentationsannahmen eine theoriegeleitete Wissensdiagnostik aussehen könnte (z.B. nach Funke, 1985, als Kausaldiagramm-Diagnostik zur Erfassung expliziten Systemwissens). Die formulierten Repräsentationsannahmen sehen qualitativ verschiedene Stufen des erworbenen Wissens vor, zu denen Relations-, Vorzeichen- und Wirkstärken-Wissen gehören. Diese Wissensarten beziehen sich auf verschiedene Differenzierungsgrade des erworbenen Systemwissens, die im Rahmen der Wissensdiagnostik unterschieden werden können („Güte des Systemwissens“). Hinsichtlich der dafür verwendeten Kausaldiagramm-Diagnostik, die von Kluwe (1988, S. 371) wegen ihres hohen Aufforderungsgehalts kritisiert wurde, liegen sowohl überzeugende Reliabilitätsstudien vor (H. Müller, 1993) als auch Studien, in denen dieser Wissenserwerb erfolgreich multinomial modelliert wurde (Beckmann, 1994). Das Verfahren wird weiter unten genauer beschrieben.

Kontrolle eines Systems (Wissensanwendung). Wissensanwendung stellt das vorhandene bzw. erworbene Wissen in einen instrumentellen Zusammenhang: das Wissen soll genutzt werden, um einen bestimmten Systemzustand zunächst herzustellen und dann gegebenenfalls längerfristig aufrechtzuerhalten. Dieser zuletzt genannte Aspekt spielt vor allem bei solchen Systemen eine wichtige Rolle, bei denen durch die Eigendynamik des Systems ein stabiler Zustand nur durch ständiges Eingreifen sichergestellt werden kann.

Bezogen auf lineare Strukturgleichungssysteme bedeutet Kontrolle des Systems, den oder die Eingriffsvektoren zu spezifizieren, die einen gegebenen Ausgangszustand in den gewünschten Zielzustand überführen, und sodann den oder die Eingriffsvektoren zu bestimmen, die den Zielzustand aufrechterhalten. Die Messung erfolgt hier durch die Bestimmung des Abstands zwischen dem gewünschten und dem erreichten Zustand („Güte der Systemsteuerung“).

Zum Verhältnis von Wissenserwerb und Wissensanwendung. Natürlich stellt sich die Frage, in welchem Verhältnis Wissenserwerb und Wissensanwendung stehen. Die (naive) Annahme, vorangehender Wissenserwerb sei notwendige und hinreichende Bedingung für erfolgreiche Wissensanwendung, ist zumindest durch einige Studien in Frage gestellt worden. So fanden etwa Berry und Broadbent (1984; 1987b; 1988) wiederholte Dissoziationen zwischen der durch Steuerungsleistungen manifestierten Wissensanwendung und dem durch die Versuchspersonen verbalisierbaren Wissen. Diese Befunde werden zwischenzeitlich jedoch anders interpretiert (vgl. Buchner, Funke & Berry, 1995; Haider, 1992, 1993), nämlich als Effekte unterschiedlich tiefer Systemexplorationen. Weiterhin machen eine Reihe neuerer Arbeiten (Beckmann, 1994; Funke, 1992; Kersting, 1999; H. Müller, 1993; Preußler, 1996, 1998; Sanderson, 1989; Süß, 1999) klar, dass durchaus substantielle positive Korrelationen zwischen Wissenserwerb und Steuerleistung zu finden sind und vor allem dann auftreten, wenn die Lernenden zum Wissenserwerb angeregt wurden, entsprechende Erfahrungen sammeln konnten und wenn die Erfassung des erworbenen Wissens hinreichend spezifisch erfolgt.

Einordnung. Der Ansatz, menschlichen Umgang mit dynamischen Systemen aus der Perspektive linearer Strukturgleichungssysteme zu betrachten, ist aus mehreren Gründen nützlich: Erstens legt auch dieser Ansatz bestimmte Annahmen über Lernprozesse und die mentale Repräsentation des erworbenen Systemwissens nahe. Zweitens lassen sich daraus Methoden der Wissensdiagnostik ableiten. Drittens schließlich erlaubt der formale Ansatz die systematische Konstruktion und Beschreibung ganzer Klassen von Systemen und damit von

experimentell leicht variierbaren Aufgaben mit wählbarem Schwierigkeitsgrad. Zahlreiche mit diesem Ansatz realisierte Experimente sind bei Funke (2003, Kap. 5) beschrieben.

### Finite Automaten

Viele alltäglich genutzte technische Systeme wie Fahrkartenautomaten, Video- und Faxgeräte, Fotokameras, Getränke- oder Parkautomaten, aber auch Computerbetriebssysteme, Anwendungsprogramme und viele weitere, auf den ersten Blick sehr unterschiedlich erscheinende Systeme (z.B. bestimmte Formen sozialer Interaktionen) auf ganz unterschiedlichen Komplexitätsstufen sind auf einer abstrakteren Ebene durch eine Reihe gemeinsamer Eigenschaften charakterisiert. Dazu zählen (vgl. Buchner, 1999): (a) Sie nehmen nur eine begrenzte (finite) Anzahl von Zuständen an. (b) Aus einem gegebenen Zustand können sie in einen darauf folgenden Zustand entweder durch eine Benutzereingabe (Drücken einer Taste; Einreichen eines Formulars an eine Behörde) oder durch einen autonomen Prozess gelangen (bei einem Fahrkartenautomaten etwa das selbständige Zurücksetzen in den Ausgangszustand nach einer festgelegten Zeit ohne Münzeinwurf; Fristablauf eines Verfahrens). (c) In Abhängigkeit von Eingabe und erreichtem Zustand wird anschließend normalerweise ein Ausgabesignal produziert (z.B. standardisierte Antwort einer Behörde auf eine Bürgeranfrage). Solche Systeme lassen sich auf formaler Ebene als finite Automaten darstellen.

Das formale System. Ein deterministischer finiter Automat ist definiert durch eine endliche Menge  $X$  von Eingabesignalen, eine endliche Menge  $Z$  von Zuständen, eine endliche Menge  $Y$  von Ausgabesignalen und durch zwei Funktionen (z.B. Hopcroft & Ullman, 1988, S. 15f.). Die Übergangsfunktion stellt eine Abbildung von  $Z \times X$  auf  $Z$  dar; sie bestimmt, welchen Zustand der Automat als Konsequenz des in einem bestimmten Zustand eingegebenen Zeichens annimmt. Die Ergebnisfunktion stellt eine Abbildung von  $Z \times X$  auf  $Y$  dar; sie bestimmt, welches Zeichen der Automat als Konsequenz des in einem bestimmten Zustand eingegebenen Zeichens ausgibt. Ein Spezialfall liegt vor, wenn das Ausgabezeichen nur von dem erreichten Zustand, nicht aber von der vorangegangenen Eingabe abhängt. Hier wird die Ergebnisfunktion durch eine Markierfunktion ersetzt, die den Zustand mit einem Ausgabesignal verknüpft.

---

Hier bitte Abbildung FSA einfügen

---

Finite Automaten werden häufig als Zustandsübergangsmatrizen oder als gerichtete Graphen dargestellt. Je nach Darstellungsform werden bestimmte Eigenschaften des Systems

besonders gut veranschaulicht (vgl. Buchner, 1999; Buchner & Funke, 1992; Funke & Buchner, 1992).

Wissenserwerb bei finiten Automaten. Bei finiten Automaten ist das Identifikationsproblem nicht so einfach zu zerlegen wie bei den vorhin dargestellten Strukturgleichungssystemen: In diesem Fall ist – zumindest prinzipiell – die gesamte Zustandsübergangsmatrix zu identifizieren. Allerdings kann die Existenz wiederkehrender Muster in der Matrix (strukturelle Redundanz) deren Identifikation erheblich erleichtern, sofern diese erkannt werden. Auch hier ergibt sich aus diesen Vorstellungen nahtlos die Annahme, dass die Zustandsübergangsmatrix subjektiv repräsentiert sein muss. Daraus lassen sich bestimmte Verfahren zur Wissensdiagnostik herleiten (z.B. Erfassung von Ausschnitten der „subjektiven“ Zustandsübergangsmatrix, vgl. Funke & Buchner, 1992), mit denen die „Güte des Systemwissens“ bestimmt werden kann.

Wissensanwendung bei finiten Automaten. Bezogen auf finite Automaten bedeutet Kontrolle des Systems die Realisierung einer Sequenz von Eingabezeichen, mittels derer der gegebene Ausgangszustand in den gewünschten Zielzustand überführt werden kann. Anders als bei Strukturgleichungssystemen ist hier in der Regel das Verbleiben im Zielzustand keine besondere Anforderung, da das System im einmal erreichten Zielzustand verharrt. Die Messung erfolgt durch Bestimmung der Schrittzahl von Start- zum Zielzustand und dem Vergleich dieses Kennwerts mit der optimalen Schrittzahl bei vollständiger Kenntnis der Zustandsübergangsmatrix („Güte der Systemsteuerung“).

Experimentelle Arbeiten. Funke und Gerdes (1993) konnten durch die Rekonstruktion der in einem Manual erläuterten Zustandsübergänge eines Videorekorders als Übergänge eines finiten Automaten (Szenario „Videorekorder“) und durch deren Abgleich mit den tatsächlich implementierten Übergängen nicht nur Fehler in der Handanweisung aufdecken, sondern auch eine verbesserte Darstellung der Bedienungslogik darauf gründen. Buchner, Funke und Berry (1995) konnten das dynamische System „Sugar Factory“ von Broadbent und Aston (1978) als finiten Automaten rekonstruieren und gelangten dadurch zu einem neuen Verständnis häufig berichteter Dissoziationsphänomene bei diesem System. Berichtet wurde nämlich von Untersuchungen mit der „Sugar Factory“, bei denen ein überraschendes Auseinanderklaffen der Leistungen beim Steuern der simulierten Fabrik (= Können, implizit) und bei der nachträglichen Beantwortung von Fragen über deren Funktionsweise (= Wissen, explizit) konstatiert wurde (vgl. Berry & Broadbent, 1984, 1987a, 1988): Gute Steuerung ging paradoxerweise mit geringem Wissen über das System einher, hohes Wissen dagegen mit

schlechter Steuerung. Diese Dissoziation, die als Beleg für die Existenz zweier verschiedener Lernprozesse und zweier Gedächtnissysteme herangezogen wurde und dadurch theoretisch weitreichende Schlussfolgerungen bewirkte (zum impliziten Lernen siehe Frensch & Rüniger, 2003), erfuhr unter dem Blickwinkel der finiten Automaten eine neue, wesentlich einfachere Interpretation: Danach scheint es geradezu zwangsläufig, dass Personen, die das Zuckersystem gut steuern und sich daher oft im Zielzustand befinden, genau deswegen weniger Wissen über andere Zustandsübergänge dieses Automaten erwerben und daher in der anschließenden Befragung schlecht abschneiden; umgekehrt verhält es sich bei denjenigen, die nur selten den Zielzustand erreichen.

Einordnung. Es kann sehr nützlich sein, dynamische Systeme als finite Automaten aufzufassen, weil dadurch wichtige Eigenschaften des zu bearbeitenden Problems offenkundig und leicht zugänglich werden. Eine Reihe nützlicher Aspekte dieses Werkzeugs für die Problemlöseforschung (z.B. Annahmen über Lernprozesse und die mentale Repräsentation des Systems, Methoden der Wissenserfassung, systematische Konstruktion und Beschreibung ganzer Klassen von Systemen) sind bei Buchner (1999) näher beschrieben. Festzuhalten ist auch, dass die von Newell und Simon (1972) vorgeschlagene Konzeption eines Problemraums und seiner Zustände, die im Lösungsprozess durchlaufen werden, hervorragend zu diesem Formalismus passt, der ebenfalls von Zuständen und dem Wechsel zwischen diesen spricht.

## *2.2. Diagnostische Zugänge zu Denk- und Problemlöseprozessen*

Diagnostische Zugänge zu Denk- und Problemlöseprozessen lassen sich grob klassifizieren in (a) Selbstauskünfte, (b) Verhaltensdaten und (c) psychophysiologische Daten. Diese Einteilung entspricht weitgehend dem Konzept der Reaktionstrias, wie es von Lazarus, Averill und Opton (1970) für den Bereich der Emotionspsychologie vorgeschlagen wurde. Wir verallgemeinern hier diesen Rahmen und halten ihn auch zur Beschreibung von Daten aus dem Bereich der Denk- und Problemlöseforschung für angemessen.

### **2.2.1. Selbstauskünfte**

Auf Selbstauskünften basierende diagnostische Verfahren sind in der Denk- und Problemlöseforschung weit verbreitet und werden gelegentlich als Königsweg zur Erfassung kognitiver Prozesse und Fähigkeiten bezeichnet. Neuere Verfahren der

Verhaltensbeobachtung (Kapitel 2.2.2) und physiologische Messverfahren (Kapitel 2.2.3) haben bisher in der Denk- und Problemlöseforschung noch nicht viel Anwendung gefunden oder werden meist nur ergänzend eingesetzt. Die Gründe dafür liegen auf der Hand: Selbstauskünfte sind erstens im Vergleich zu Verhaltensdaten und physiologischen Maßen mit deutlich weniger materiellem und zeitlichem Aufwand bei der Erhebung und Auswertung verbunden. Zweitens stellen physiologische Korrelate nur ein indirektes Maß dar, wenngleich sie gegenüber den Selbstauskünften den Vorteil der Objektivität haben. Selbstauskünfte sind als Daten zwar zunächst objektiv erfassbar, der Inhalt dessen, was eine Vp sagt, ist jedoch nicht intersubjektiv fassbar.

Unter der Rubrik Selbstauskünfte werden im folgenden Teil alle diagnostischen Methoden zusammengefasst, mittels derer eine Person zur Auskunft über ihre kognitiven Prozesse und Leistungen vor, während und/oder nach einer Denk- oder Problemlöseaufgabe bewegt werden kann. Dazu zählen verbale Daten, insbesondere die Methode des Lauten Denkens, die Wissensdiagnostik und psychologische Testverfahren. Einen allgemeinen Überblick über verbale Daten und qualitative Methoden generell geben die Handbücher von König (2002) sowie Mayring (2002).

### **Verbale Daten**

Verbale Daten bezeichnen qualitative Methoden, die zum Ziel haben, die subjektive Konstruktion von Personen beim Denken und Problemlösen (ihre subjektive Weltansicht, ihre Erklärungsmuster und Handlungsprogramme) unter Berücksichtigung emotional-motivationaler Prozesse zu explorieren (G. L. Huber & Mandl, 1994). Verbale Daten entstehen unter dem Einsatz von Verbalisationsmethoden vor (präaktional), während (periaktional) oder nach (postaktional) zu untersuchenden Handlungsausführungen und lassen sich zudem nach dem Grad ihrer Strukturiertheit unterscheiden. Ein Tagebucheintrag hat beispielsweise eine niedrige Strukturiertheit, während ein Fragebogen üblicherweise als hoch strukturiert angesehen wird (G. L. Huber & Mandl, 1994). Zu den präaktionalen Methoden zählen z.B. Fragebogen, Ratingskala, strukturiertes Interview und die Gruppendiskussion; periaktionale Methoden umfassen die Methode des Lauten Denkens und Protokollnotizen. Beispiele für postaktionale Methoden sind Nachträgliches Lautes Denken (mit oder ohne Medienunterstützung), fokussiertes Interview, narratives Interview und Protokollnotizen.

In diesem Abschnitt sollen sowohl periaktionale Methoden, insbesondere die Methode des „Lauten Denkens“ (LD), als auch prä- und postaktionale Methoden besprochen werden. Dabei

soll ein Schwerpunkt auf der Methode des LD sowie den Weiterentwicklungen der Methode und Befunden aus Untersuchungen zum LD beim Denken und Problemlösen liegen. In Bezug auf Denk- und Problemlöseprozesse ist insbesondere die Anwendung periaktionaler Verfahren interessant, da sie einen Zugang zu den während einer Handlung ablaufenden kognitiven Prozessen ermöglichen.

Verbale Methoden, häufig auch als Methode der Introspektion bezeichnet, haben eine lange Tradition in der Denkpsychologie (Ericsson & Simon, 1993). Nachdem die Methode zunächst in die Kritik der Behavioristen geriet, werden Laut-Denk-Protokolle seit Anfang der 1970er Jahre in der Problemlöseforschung vermehrt eingesetzt (Dörner et al., 1983; Lüer, 1973), weil das Bedürfnis nach Datenquellen wuchs, in denen verstärkt prozedurale und dynamische Aspekte kognitiver Prozesse sichtbar werden.

Neben der Problemlöseforschung haben sich auch andere Forschungsbereiche verbaler Methoden bedient, z.B. die Spracherwerbs- und Leseforschung (Afflerbach, 2000; Pritchard, 1990), die Unterrichtsforschung (Goldman, Zech, Biswas & Noser, 1999; Weidle & Wagner, 1994), die Entscheidungsforschung (Backlund, Skånér, Montgomery, Bring & Strender, 2003), die Medienforschung (Eveland & Dunwoody, 2000) und die Forschung zur Mensch-Maschine-Interaktion (Ritter & Larkin, 1994).

Die Verwendbarkeit verbaler Informationen als Daten zur Identifizierung handlungsleitender Kognitionen ist an eine Reihe von Vorbedingungen geknüpft. Zunächst müssen Kognition und Handlung in einer engen zeitlichen Relation stehen, da sonst eine auf einer Kognition basierende Aussage nicht mit einer bestimmten Handlung in Verbindung gebracht werden kann. Zudem muss die (die handelnde Person muss prinzipiell z.B. Zugang zu ihren Kognitionen haben können, vgl. Nisbett & Wilson, 1977). Das Vorhandensein eines Kurzzeitgedächtnisses stellt eine weitere unabdingbare Voraussetzung für die Reliabilität der Daten dar. Ohne Erinnerung wäre eine Vp nicht in der Lage, ihre eigenen Handlungen zu kommentieren. Schließlich müssen die gewonnenen Informationen valide sein, eine Voraussetzung, die unter anderem durch den Einsatz anderer Verhaltensmaße (wie z.B. Blickbewegungen) überprüft werden kann.

Periaktionalen verbalen Methoden kommt eine besondere Rolle für die Denk- und Problemlöseforschung zu (Ericsson & Simon, 1980; Newell, 1990): Es gibt nur wenige Methoden, die Aufschluss über die während einer Handlung ablaufenden kognitiven Prozesse geben (z.B. Aufzeichnung von Blickbewegungen, Hand- und Körperbewegungen, Mimik, physiologische Korrelate). Betrachtet man die Methode des LD als periaktionaler Verfahren,

so unterscheidet man üblicherweise zwei Formen: LD mit Handlungsunterbrechung durch den Versuchsleiter (z.B. durch „metacognitive probes“, Dominowski, 1998) und kontinuierliches LD. Als postaktionale Methode betrachtet lassen sich das Nachträgliche Laute Denken und die Methode des „stimulated recall“, bzw. medienunterstützten LD (häufig auch als Videokommentiertechnik bezeichnet) unterscheiden. Diese Methoden der retrospektiven Verbalisation werden weiter unten ausführlich besprochen.

Verwendung des LD in der Denk- und Problemlöseforschung. Neben Parametern der Systemsteuerung und den Protokollen der Eingriffe der Versuchspersonen sind in Studien zum Denken und Problemlösen entsprechend Protokolle des LD erhoben worden, sowohl (a) beim einfachen Problemlösen, z.B. beim „Turm von Hanoi“ (bei Erwachsenen: Anzai & Simon, 1979; Hayes & Simon, 1977; bei Kindern: Klahr, 1981) als auch (b) beim komplexen Problemlösen (Dörner et al., 1983; Dörner & Reither, 1978; Heineken, Arnold, Kopp & Soltysiak, 1992; Putz-Osterloh, 1981b, 1985; Reichert & Dörner, 1988; Roth, Meyer & Lampe, 1991). Ziel ist es, Informationen über den kognitiven Problemlöseprozess, über Subroutinen und Strategien und begleitende emotional-motivationale Prozesse zu erhalten.

Die Befunde aus der Verwendung von LD-Daten beim Problemlösen betreffen häufig die Unterscheidung erfolgreicher und weniger erfolgreicher Problemlöser. Roth (1985; 1987) greift die These von Dörner et al. (1983) auf, wonach Denkstil und Sprachstil korreliert sind, und postuliert, dass sich erfolglose und erfolgreiche Problemlöser durch die Art ihrer Verbalisierung unterscheiden lassen. Schlechte Problemlöser verwenden häufiger Worte, die auf eine relative Geschlossenheit der kognitiven Organisation hinweisen (z.B. „alle“, „immer“, vs. „einige“, „manchmal“). Gute Problemlöser zeichnen sich dagegen z.B. durch geringere Verwendung des Konjunktivs und weniger Lexeme aus dem Dogmatismus-Kodierlexikon aus.

Im Kontext des LD ist häufig die Frage untersucht worden, ob Verbalisierung während des Problemlöseprozesses zu einer Veränderung kognitiver Leistungen führt (Berry & Broadbent, 1984; Dickson, McLennan & Omodei, 2000; Knoblich & Rhenius, 1995; Schooler, Ohlsson & Brooks, 1993). Hintergrund der Überlegungen zum Einfluss LD auf den Problemlöseprozess ist ein von Ericsson und Simon (1980) postuliertes Prozessmodell des LD, demzufolge Inhalte, die bereits im verbalen Code vorliegen, ohne zusätzlichen Aufwand in sprachliche Äußerungen umgesetzt werden können. Dagegen ist für Inhalte, die noch nicht im verbalen Code vorliegen, ein zusätzlicher Rekodierungsschritt notwendig, der eine verlangsamte Aufgabenbearbeitung und somit eine Interferenz der Versprachlichung mit dem

Problemlöseprozess zur Folge haben kann. Demnach müsste LD die Performanz beeinträchtigen. Andererseits ist vorstellbar, dass LD Metakognitionen (z.B. Nachdenken über eigene Strategien) hervorruft, was wiederum zu einer Performanzsteigerung führen kann. Studien zum Einfluss des Verbalisierens auf kognitive Leistung kommen jedoch zu widersprüchlichen Ergebnissen. Es wurden entweder (a) keine Reaktivität / keine Performanzunterschiede (z.B. Biggs, Rosman & Sergenian, 1993; Deffner, 1984), (b) ein positiver Effekt, vor allem die Förderung eines analytischen Urteils (z.B. Franzen & Merz, 1988; Merz, 1969; Short, Evans, Friebert & Schatschneider, 1991) oder (c) ein negativer Effekt, d.h. schlechtere Leistung, z.B. beim Lösen von Einsichtsproblemen (Schooler et al., 1993) gefunden. Knoblich und Rhenius (1995) zeigen in einer Untersuchung zum LD beim Problemlösen mit gleichzeitiger Aufzeichnung von Blickbewegungen, dass sich Unterschiede zwischen Laut- und Stilldenkern im Problemlöseverhalten zeigen; Lautdenker halten bspw. länger an einer einmal gewählten Strategie fest. Dickson, McLennan und Omodei (2000) fassen zusammen, dass LD dann eher förderlich für einen Problemlöseprozess ist, wenn die Aufgabe (a) gut strukturiert, (b) über verbalisierbare Regeln beschrieben werden kann und (c) nicht zeitkritisch ist. Anhand des Szenarios „Fire Chief“ (Omodei & Wearing, 1995) wird gezeigt, dass LD in komplexen, dynamischen Problemsituationen einen negativen Einfluss auf die Problemlösequalität (hier: prozentualer Anteil des geretteten Waldes) und die Anzahl der eingeleiteten Maßnahmen hat. Dieser Befund wird auf die zusätzliche Bindung von Aufmerksamkeitsressourcen durch die Verbalisierungsaufgabe zurückgeführt, wodurch es zu Verzögerungen im Problemlöseprozess kommt.

Neben dem LD spielen auch andere periaktionale Methoden in der Denk- und Problemlöseforschung eine Rolle. Die Analyse von Problemlösesitzungen in Gruppen beispielsweise bietet eine Möglichkeit der Erfassung von Problemlöseprozessen, insbesondere der Problemstrukturierung, Lösungssuche und Entscheidungsfindung. In Studien zu Interaktionsprozessen beim Problemlösen (Boos, Morguet, Meier & Fisch, 1990; Boos, Scharpf & Fisch, 1991) werden die Anzahl der angesprochenen Problemelemente, ihre Neuheit, Differenziertheit, Kreativität und Verknüpfung kodiert.

Insbesondere in der Unterrichtsforschung wurde das Verfahren des LD weiterentwickelt mit dem Ziel, mittels der Verbalisierung die Notwendigkeit von Verhaltensmodifikationen zu erkennen und diese durchzuführen (Aktionsforschungsansatz, Groeben, 1986). In diesem Kontext entstand die Methode des Nachträglichen Lauten Denkens (NLD, Wagner, Uttendorfer-Marek & Weidle, 1977; Weidle & Wagner, 1994) und des medienunterstützten

Verbalisierens (sog. „stimulated recall“, Kagan & Kagan, 1991; McLennan & Omodei, 1996), auch Videokommentiertechnik genannt. Bei der Methode des nicht-medienunterstützten NLD wird als problematisch angesehen, dass bestimmte Aspekte eines kognitiven Prozesses unzugänglich bleiben (Omodei & McLennan, 1994). Beim „stimulated recall“ hingegen, bei dem der Proband z.B. eine Videoaufnahme von sich selbst beim Problemlösen oder von einem seinem Gesichtsfeld entsprechenden Monitorauschnitt sieht, können auch vermeintlich unzugängliche Aspekte des kognitiven Prozesses erinnert und verbalisiert werden.

Auswertung von LD-Daten. LD-Daten werden meistens mit der Methode der Inhaltsanalyse ausgewertet. Inhaltsanalytische Auswertungsmethoden dienen der Klassifikation von verbalen und non-verbalen Kommunikationsinhalten und ordnen diese nach festgelegten Regeln a priori aufgestellten Kategorien zu (G. L. Huber & Mandl, 1994; Mayring, 2002).

Bei der quantitativen Inhaltsanalyse werden für jede Dimension Analysekatoren entwickelt, denen das vorliegende Material unter dem Gesichtspunkt der semantischen Ähnlichkeit (z.B. mit Hilfe von Prototypen) zugeordnet wird. Das Auftreten einzelner sprachlicher Äußerungen, z.B. bestimmter Wörter, wird nach ihrer Häufigkeit ausgezählt (sog. Frequenzanalyse) und ggf. nach dem Intensitätsgrad eingeschätzt (sog. Intensitätsanalyse, welche das Ausmaß vorgenommener Wertungen des Materials erfasst). Neben der Inhaltsanalyse einzelner sprachlicher Äußerungen können auch andere Textmerkmale ausgewertet werden, etwa grammatikalische oder rhetorische Konstruktionen, Diskurs- und Wissensstrukturen. Hier geht es vor allem um Beziehungen zwischen den kodierten Äußerungen (vgl. Schnotz, 1994). Das Ziel einer quantitativen Vorgehensweise besteht folglich darin, die erfassten Sinngehalte in Form von Häufigkeiten bzw. Assoziationsmustern auszuwerten, um so zu statistisch analysierbaren Vergleichen, Trendmustern, Gruppen- oder Zeitvergleichen zu kommen.

Die qualitative Inhaltsanalyse bezieht auch Kommunikationsinhalte, die nicht explizit ausgesprochen werden, in die Analyse ein. Durch eine systematische Interpretation wird die inhaltliche Bedeutung von Aussagen ermittelt, ohne das Material auf quantifizierbare Aussagen zu reduzieren. Dabei wird zunächst nach Sichtung des Materials ein System von Kategorien festgelegt, anhand dessen durch die interpretativen Techniken der Zusammenfassung, Explikation und/oder Strukturierung Aussagen aus dem Text herausgefiltert werden. Damit wird das Ziel verfolgt, die subjektive Perspektive des

Handelnden und seine Weltsicht (Wirklichkeitskonstruktion) herauszuarbeiten, um auf dieser Grundlage Verhalten verstehen und erklären zu können.

Unter den praktischen Problemen der Inhaltsanalyse ist vorrangig das Problem der Reliabilität zu nennen, d.h. der Zuverlässigkeit der Einordnungen der Textbestandteile in die vorgegebenen Kategorien. Es ist sowohl damit zu rechnen, dass ein und dasselbe zu verkodende Element von verschiedenen Personen unterschiedlich beurteilt wird, als auch damit, dass eine Person zu verschiedenen Zeitpunkten unterschiedliche Kategorienzuordnungen vornimmt. Mit der Analyse der Inter- und Intra-Rater-Übereinstimmung ist zumindest eine Überprüfung der Reliabilität möglich.

Computergestützte Methoden, die für die Inhaltsanalyse vermehrt zum Einsatz kommen, können möglicherweise für eine höhere Reliabilität des Verfahrens sorgen, indem die Kodierung einzelner Elemente computergestützt vorgenommen wird (Mayring, 2002). Für die Bildung der Kategorien sind solche Verfahren jedoch nicht anwendbar, somit bleibt hier das Problem der Reliabilität bestehen. Die Ergebnisse der Inhaltsanalyse werden maßgeblich von der Qualität der Auswertungskategorien beeinflusst. Hier ist kritisch zu bemerken, dass eine a priori Bildung von Analysekategorien (also z.B. eine deduktive, theoriebasierte Kategorienbildung) häufig nicht möglich ist, sondern Kategorien und damit Kodierschemata erst auf Basis des vorliegenden Materials induktiv gebildet werden können. Dieses Problem kann durch Vortests zur Bildung und Testung von Analysekategorien und Kodierschemata umgangen werden. Vortests können gleichzeitig dem Training von Beurteilern dienen und so die Inter- und Intra-Rater-Reliabilität erhöhen. Häufig werden LD-Protokolle jedoch bewusst explorativ eingesetzt; dann kann eine induktive Kategorienbildung sinnvoll sein.

Fazit. Wie ist nun der Einsatz verbaler Daten insgesamt zu bewerten? Im Vergleich zur Befragung ist durch die Methode des LD ein „Online“-Zugang zu den während der Aufgabe ablaufenden mentalen Operationen und dadurch zu den handlungsleitenden Kognitionen möglich. Mit der Methode sind jedoch auch zahlreiche Nachteile verbunden (vgl. Weidle & Wagner, 1994), so z.B. (a) das Problem der Kapazität/Auswahl, (b) das Problem der Bewusstheit, (c) das Problem der sozialen Situation (z.B. Problem der sozialen Erwünschtheit und Problem der Reaktivität), (d) das Problem der sprachlichen Umsetzung. Darüber hinaus vertreten Nisbett und Wilson (1977) in Bezug auf retrospektiv erhobene Daten das Argument, dass Personen nur Zugang zu den Produkten ihrer kognitiven Prozesse hätten und nicht zu den kognitiven Prozessen selbst. Hinzu kommen Probleme bei der inhaltsanalytischen Auswertung, z.B. betreffend der Reliabilität oder der Bildung von Analysekategorien.

Festzuhalten bleibt, dass LD-Protokolle ein nützliches Verfahren für explorative Untersuchungen darstellen, obgleich in solchen Fällen die Auswertung der Daten schwierig ist (vgl. auch Kluwe, 1995). LD-Daten können zudem in Kombination mit anderen prozessorientierten Methoden eingesetzt werden. Auf Studien, in denen neben LD-Daten Blickbewegungen zur konkurrenten Validierung eingesetzt wurden (vgl. Deffner, 1987; Knoblich & Rhenius, 1995), wird im Kapitel 2.2.2.2 eingegangen.

### **Befragungen**

Fragebogenverfahren werden in der Denk- und Problemlöseforschung in erster Linie zur Erfassung relevanter Persönlichkeitsmerkmale, wie z.B. der Kompetenzüberzeugung und der Entscheidungsfähigkeit (Friedman, 1996; Köller & Strauß, 1994; Stäudel, 1988) eingesetzt. Andere Verfahren zielen darauf ab, Copingstrategien von Versuchspersonen in belastenden Problemlösesituationen, den Umgang mit Stress, die Bewältigung negativer Emotionen oder ihre Motivation für die gegebene Aufgabe zu untersuchen (Otto, Döring-Seipel, Grebe & Lantermann, 2001; Stäudel, 1987). Verfahren aus dem neuropsychologischen Bereich schließlich dienen der Diagnose von Problemlöse-, Planungs- und Handlungsstörungen (Gauggel, Deckersbach & Rolko, 1998, siehe nächsten Abschnitt).

### **Psychometrische Tests**

Natürlich ist als diagnostisches Mittel zur Untersuchung von Denk- und Problemlöseprozessen auch ein standardisierter Test vorstellbar, der neben objektiver Messung auch Daten zur Standardisierung und Normierung liefert.

Ein klassisches Beispiel für diesen Zugang bietet der „Denksport-Test“ (DST) von Lienert (1964). Dieser Test besteht aus 15 Denksportaufgaben mit steigender Schwierigkeit, die als eingekleidete Textaufgaben mit je 5 Antwortmöglichkeiten vorgegeben werden. Für die Bearbeitung stehen 45 Minuten Zeit zur Verfügung. Ein Beispielitem enthält Abbildung DST\_ITEM.

---

Hier bitte Abbildung DST\_ITEM einfügen

---

Die Spearman-Brown-Reliabilität liegt bei einer Stichprobe von N=323 Abiturientinnen und Abiturienten bei  $r_{tt}=0.83$ , die Validität wurde über ein Lehrerurteil sowie eine Kontrastgruppenanalyse (Abiturienten gegen Berufsschüler) belegt. Die Standardisierung erfolgte anhand einer Stichprobe von weiteren N=836 Abiturientinnen und Abiturienten.

Lienert schränkt daher den Anwendungsbereich seines Tests auf Oberschüler, Abiturienten, Studenten und Akademiker ein.

Ein anderes Beispiel ist das Verfahren „Plan A Day“ (PAD) von Funke und Krüger (1995), das unter standardisierten Bedingungen eine Planungsaufgabe derart stellt, dass von der Testperson eine bestimmte Menge an Aufträgen an verschiedenen Orten und zu verschiedenen Zeiten zu erledigen sind und hierfür ein optimaler Plan zu finden ist.

Abbildung PAD zeigt das Display mit den verschiedenen Orten und dem Wegenetz, innerhalb dessen die raumzeitliche Planung vorzunehmen ist.

---

Hier bitte Abbildung PAD einfügen

---

Durch die Anzahl der Aufträge sowie durch die zeitlichen Einschränkungen bei deren Erledigung kann die Schwierigkeit systematisch gesteigert werden. Die zur Planung benötigte Zeit wie auch die Anzahl korrekt erledigter Aufträge stellen mögliche Messgrößen dar.

Anwendungen dieses Messverfahrens finden sich z.B. in der Studie von Gouzoulis-Mayfrank et al. (2002) über die Auswirkungen des Ecstasy-Konsums auf kognitive Funktionen oder in der Arbeit von Kliegel et al. (2002) über altersbedingte Veränderungen des prospektiven Gedächtnisses.

Ein Vorläufer des PAD ist der „Bogenhausener Planungstest“, der von Stoltze (1991) als Testverfahren für den Einsatz im neuropsychologischen Kontext entwickelt wurde (siehe von Cramon, Matthes-von Cramon & Mai, 1991) und bis heute trotz gewisser Schwächen im Einsatz ist (S. V. Müller, von der Fecht, Hildebrandt & Münte, 2000). Eine Übersicht über planungsdiagnostische Instrumente aus den Bereichen der Arbeits- und Betriebspsychologie, der Entwicklungspsychologie und der Neuropsychologie findet sich bei Fritz und Funke (1995).

Schließlich ist noch der „Fragebogen zum Problemlösen“ (PLF) von König, Liepmann, Holling und Otto (1985) zu erwähnen, der aus 50 problembezogenen Items besteht, die den fünf Skalen (1) Problemerleben (hoch ladendes Beispiel-Item: „Ich werde mit Problemen gut fertig“), (2) Problemverleugnung („Ich kann nur staunen, welche Probleme Leute haben“), (3) Neigung zu unkonventionellen Problemlösungen („Unter vielen albernen Einfällen habe ich auch gute“), (4) Problembearbeitung („Wenn ich Schwierigkeiten habe, überlege ich, wodurch ich sie ändern kann“) sowie (5) Neigung zu konservativen Problemlösungen („Ich vertraue erprobten Lösungen“) zugeordnet sind. Die an 441 Personen ermittelte interne Konsistenz und die aufgrund einer Teilstichprobe von 56 Personen errechnete Retest-

Reliabilität können lediglich für die zuerst genannte Skala als gut bezeichnet werden. Die Validitätsbelege konvergenter und diskriminanter Art sind jedoch recht heterogen und lassen weitere Validierungsstudien als wünschenswert erscheinen.

### **Kontrollierte Wissensdiagnostik**

Versteht man wie Kluwe (1988, S. 359) unter Wissen die mitteilbaren Kenntnisse über Sachverhalte und Vorgänge der Realität, so sind in gewissem Sinn alle diejenigen diagnostischen Verfahren als Wissensdiagnostik zu bezeichnen, die etwas über die bei Denk- und Problemlöseakten benutzten Wissens Elemente in Erfahrung bringen, also auch verbale Daten und Befragungen. Um hier eine sinnvolle Abgrenzung vornehmen zu können, wird der Begriff „kontrollierte Wissensdiagnostik“ für diejenigen Erhebungsmethoden reserviert, die modellgeleitet und systematisch vorgehen. Modellgeleitet bedeutet, dass bestimmte theoretische Annahmen über die Struktur des zur Aufgabenbearbeitung benötigten Wissens gemacht werden. Systematisch bedeutet, dass gezielt nach theoretisch postulierten Komponenten gesucht wird. Die erhobenen Wissensbestände können – wiederum Kluwe (1988) folgend – hinsichtlich ihrer Qualität (inhaltliche Beschaffenheit, Richtigkeit) sowie ihrer Struktur (Anzahl, Umfang) bewertet werden.

Kausaldiagramm-Diagnostik (KDD). Als Beispiel für eine derartige kontrollierte Wissensdiagnostik soll die Technik der Kausaldiagramm-Diagnostik (KDD) beschrieben werden, die im Kontext der Untersuchung des Umgangs mit dynamischen Strukturgleichungssystemen Anwendung findet (z.B. bei Funke, 1986, 1992). Die Vpn müssen aufgrund ihrer Interaktion mit einem unbekanntem dynamischen Kleinsystem herausfinden (Anforderung der Identifikation eines Systems), wie bestimmte beeinflussbare Größen (die exogenen Variablen) auf andere Größen (die endogenen Variablen) wirken. Die Struktur solcher Systeme wird formal durch lineare Strukturgleichungssysteme beschrieben (siehe oben).

Die KDD erhebt die vermuteten Kausalbeziehungen einer Vp durch ein Diagramm (entweder auf Papier oder am Bildschirm vorgegeben), in dem die exogenen den endogenen Variablen als Boxen gegenübergestellt werden und die Vp die vermuteten Kausalbeziehungen als Kanten in dieses Netzwerk eintragen soll. Entsprechend den Modellannahmen von Funke (1992, Kap. 3.3) können diese Kanten von der Vp jeweils typisiert werden als Relations-, Vorzeichen- oder Gewichts-Beziehung. Das von der Vp generierte Diagramm kann mit dem tatsächlich implementierten dann auf Übereinstimmung verglichen werden, wobei für jeden

Kantentyp gesonderte Bewertungen vorgenommen werden können („Güte der Kausaldiagramme“). Bei wiederholter Messung kann aus den Veränderungen im Diagramm auf Veränderungen der zugrunde liegenden Wissensstruktur geschlossen werden (Lernverlauf). Abbildung KDD zeigt ein subjektives (links) im Vergleich mit dem tatsächlich implementierten Modell (rechts).

---

Hier bitte Abbildung KDD einfügen

---

Die Unterscheidung von Relations-, Vorzeichen- oder Gewichts-Beziehungen in kausal verknüpften Systemen folgt einer Theorie des Wissenserwerbs bzw. der Wissensrepräsentation, die bei der Identifikation von Systembeziehungen ein schrittweises Vorgehen postuliert. Zunächst wird das Vorliegen einer wie auch immer gearteten Relation zwischen zwei Variablen erkannt (Stufe des Relationswissens), dann wird die Wirkrichtung genauer erkannt (Stufe des Vorzeichenwissens), schließlich kann sogar die genaue Stärke dieser Relation bestimmt werden (Stufe des Wirkstärkewissens). Vor diesem Hintergrund können auch subjektive mit den objektiv implementierten Kausaldiagrammen verglichen werden. Somit kann jede Angabe eindeutig klassifiziert werden – ist bspw. die Vermutung der Vp eine Aussage über eine objektiv existierende Relation („Treffer“), kann sie hinsichtlich Relation, Vorzeichen und Wirkstärke beurteilt werden. Höhere Stufen schließen dabei untere ein (d.h. eine Wirkstärkenangabe kann falsch sein, aber Vorzeichen und damit auch Relation können richtig sein). Auch die falschen Angaben („Fehler“) lassen sich entsprechend einteilen (nur falsche Relation, falsche Relation mit Vorzeichen, falsche Relation mit Wirkstärke). Das daraus bestimmbare Maß „Güte der Kausaldiagramme“ (GdK) von Funke (1992, S. 82) bestimmt sich für jede der drei Komponenten Relation, Vorzeichen und Wirkstärke als Differenz zwischen Treffer- und Fehleranteilen, die auf das im jeweiligen Kausaldiagramm mögliche Maximum an Treffern oder Fehlern normiert werden:

$$GdK = \frac{Treffer}{Max(Treffer)} - \frac{Fehler}{Max(Fehler)} \quad (1)$$

Das Maß GdK hat sich in zahlreichen Arbeiten als Indikator des Wissenserwerbs bewährt (z.B. Beckmann, 1994; Funke, 1992; H. Müller, 1993). Weiter unten wird beschrieben, wie Beckmann (1994) den hier postulierten Prozess des Wissenserwerbs in verschiedenen Stufen mittels multinomialer Modellierung rekonstruiert hat.

Die zugrunde gelegte Entwicklungslogik des Wissenserwerbs findet sich in ähnlicher Weise bei Plötzner et al. (1990), wo der Wissenserwerb über elastische Stoßvorgänge auf vier Stufen

beschrieben wird (vgl. auch Plötzner & Kneser, 1998; Plötzner & Spada, 1993): (1) qualitative Stufe, (2) semi-quantitative Stufe (Repräsentation monoton steigender oder fallender Relationen), (3) quantitative relationale Stufe (Repräsentation proportionaler Relationen) und (4) quantitative Stufe (Repräsentation in Form quantitativer Gleichungen). Die Ähnlichkeit zu dem zuvor beschriebenen Stufenmodell ist unübersehbar.

Erfassung aktiver Informationssuche. Eine Variante kontrollierter Wissensdiagnostik stellt das Verfahren der „Aktiven Informationssuche“ dar, das von Oswald Huber im Kontext der Entscheidungsforschung entwickelt wurde. Huber (im Druck) berichtet in seinem Übersichtsartikel zum Entscheiden unter Risiko über Schwächen konventioneller Entscheidungsforschung. Dabei hebt er hervor, dass die in klassischen Untersuchungen standardmäßig praktizierte, versuchsbedingte Vorgabe von Handlungsalternativen und deren Wahrscheinlichkeiten ein Phänomen völlig zum Verschwinden bringt, das eigentlich von zentraler Bedeutung ist: die aktive Suche nach Mitteln, die ein Risiko vermindern.

In den von Huber gestalteten Versuchssituationen werden quasi-realistische Risikosituationen vorgegeben und mit der Methode der „Aktiven Informationssuche“ (AIS) untersucht, die die von den Entscheidern aktiv gesuchten Informationen dokumentiert (vgl. O. Huber, Wider & Huber, 1997). Die Vpn können dabei entweder vorgegebene Zusatzinformationen abrufen oder vom Versuchsleiter (VL) im kontrollierten Dialog erfragen. Es zeigt sich, dass die Vpn nicht an Wahrscheinlichkeiten interessiert sind, sondern vor allem nach Risikoentschärfung suchen.

Williamson, Ranyard und Cuthbert (2000) berichten über Erfahrungen mit einer Variante der AIS von Huber et al. (1997), die sie „conversational AIS technique“ nennen und die den sozialen Kontext zu berücksichtigen versucht, insofern als Konversationsmaximen wie etwa die Grice'schen (1975) Kriterien der Qualität, Quantität, Relevanz und Stil auch in denkpsychologischen Untersuchungssituationen zur Anwendung kommen. Die typischen Komponenten konversationsbasierter Verfahren zur Erfassung von Denkprozessen lassen sich vom Ablauf her in (1) Orientierung, (2) Szenario-Darbietung und (3) Debriefing unterteilen.

Im Hinblick auf die als zweiten Punkt benannte Szenario-Darbietung sind verschiedene Varianten zu unterscheiden:

- (1) Vorgehende Merkmalerfassung („preliminary attribute elicitation“, PAE): Nach der Kenntnisnahme des Szenarios wird verlangt, die zur Entscheidung benötigten Informationen zu benennen (schriftlich wie mündlich). Ein Beispiel hierfür findet sich bei Hershey et al. (1990).

(2) Konversationsbasierte AIS: Der VI gibt in der Rolle eines Beraters (standardisierte) Antworten auf faktenbezogene Fragen der Versuchsperson; sobald die Vp 60 Sekunden schweigt, gibt der VI einen Hinweis, der zu weiteren Fragen führt, bis die Vp ihre Entscheidung trifft. Ein Beispiel hierfür ist die Studie von Huber et al. (1997).

(3) Konversationsbasiertes Lautes Denken: Hierbei spielt der VI die Rolle eines nicht-urteilenden Zuhörenden, damit die laut denkende Person nicht in den leeren Raum hinein reden muss. Im Unterschied zur Standard-Prozedur des LD spricht die Vp hier also zu einem Zuhörer und nicht nur ins Mikrophon; zudem findet eine Interaktion dadurch statt, dass der VI zuhört und Interesse bekundet, sowohl durch verbale (z.B. „Verstehe!“, „Ah so!“) wie non-verbale (z.B. Nicken) Signale. Ein Beispiel dafür findet sich bei Williamson et al. (2000).

(4) Nicht-direktives Nachentscheidungs-Interview bzw. gezieltes Nachentscheidungs-Interview. Beide Varianten fordern einen retrospektiven Bericht über die subjektiv wahrgenommene Entscheidungsfindung, in einem Fall in freier Form und eher zusammenfassend, im anderen Fall mit gezielten Fragen zum Vorgehen. Ein Beispiel für diese Methode findet sich bei Klein, Calderwood und MacGregor (1989).

Die hier berichteten Varianten der Erfassung aktiver Informationssuche sind insofern modellgeleitet und systematisch, als sie gezielt zur Aufdeckung des Prozesses der Risikoentschärfung eingesetzt werden. Der Prozess der Entscheidungsfindung bei quasi-realistischen Situationsbeschreibungen, der nach Huber in der klassischen Entscheidungsforschung nur selten thematisiert wird, setzt mit dem Aufbau einer mentalen Repräsentation ein und wird erst durch die AIS-Technik sichtbar. Sobald man entdeckt, dass eine ansonsten attraktive Alternative negative Konsequenzen haben kann, kommt es zur Suchentscheidung (Soll ein Operator zur Risikoentschärfung gesucht werden? Lohnt sich der Aufwand?), die bei positivem Urteil zu einer Akzeptanzentscheidung über etwaige Risikoentschärfungsoperatoren führt.

## 2.2.2. Verhaltensdaten

Neben den subjektiven Daten interessieren sich Denk- und Problemlöseforschende auch für Verhaltensdaten. Verfahren der Zeitmessung, die Erfassung und Auswertung von Blickbewegungen sowie die kontrollierte Verhaltensbeobachtung sind hier zu erläutern.

Die Verwendung von Verhaltensdaten hat eine lange Tradition in der Denkpsychologie. Reaktionszeitmessungen wurden bereits Ende des 19. Jahrhunderts eingesetzt; Blickbewegungen und psychophysiologische Verfahren dagegen werden erst in jüngerer Zeit verstärkt auch in der kognitiven Psychologie verwendet. Gegenüber Selbstauskünften haben Verhaltensbeobachtungen generell den Vorteil, dass sie weitgehend objektiv sind. Allerdings ist ihre Verwendung meist mit einem vergleichsweise höheren Aufwand verbunden und an eine Reihe von theoretischen Vorannahmen geknüpft, auf die weiter unten näher eingegangen wird. Neben den genannten Verfahren sind in der Denkforschung – analog zu Laut-Denk-Protokollen – auch andere Verhaltensprotokolle eingesetzt worden, mittels derer z.B. die Bewegungen der Maus und die Bedienung der Tastatur (Card, Moran & Newell, 1983; Ritter & Larkin, 1994) analysiert wurden (für einen Überblick siehe Salvucci & Anderson, 2001).

Beim Versuch einer Klassifikation der verschiedenen verhaltensbeobachtenden Zugänge zum Denken und Problemlösen fällt auf, dass sich ein großer Teil der Methoden zur Verhaltensbeobachtung in der kognitiven Psychologie auf Effekte im Sekunden- und Millisekundenbereich konzentriert. Insbesondere bei der Untersuchung komplexerer kognitiver Prozesse, wie dem Planen oder Problemlösen, stellt die Einengung der Betrachtung auf diesen Zeitbereich ein Problem dar, wenn man an der Entwicklung und Veränderung kognitiver Prozesse interessiert ist. Mit seiner Zeitskala menschlicher Handlungen hat Alan Newell (1990) versucht, eine Verbindung zwischen den einzelnen Bereichen menschlicher Aktivität – vom biologischen „Band“, über das kognitive, rationale, bis hin zum sozialen Band – herzustellen (vgl. Tabelle NEWELL).

---

Hier bitte Tabelle NEWELL einfügen

---

Anderson (2002) argumentiert für eine Dekomposition von Lernprozessen auf die Ebene der Aufgabe (1-10 min), Aufgabeneinheit (10 sek) oder Körperbewegung (200-300 msec). Nur eine „feinkörnige“ Untergliederung kognitiver Prozesse in ihre Komponenten macht diese überhaupt zu messbaren Ereignissen. Die Herausforderung besteht dabei in zweierlei Hinsicht: Zum einen muss eine feinkörnige Analyse mit allen daran geknüpften technischen und theoretischen Voraussetzungen möglich gemacht werden. Zum anderen müssen die Befunde anschließend auf eine höhere zeitliche Ebene überführt werden, um diagnostische Konsequenzen ziehen zu können (siehe dafür bspw. Kapitel 2.3.1 zur kognitiven Modellierung). Die postulierte Notwendigkeit der Dekomposition kognitiver Prozesse in kleinste zeitliche Einheiten rechtfertigt die Schwerpunktsetzung bei den hier vorgestellten Methoden auf Blickbewegungsmessungen und psychophysiologische Methoden.

## **Zeitmessung**

Die Zeitmessung oder mentale Chronometrie (Posner, 1978) versucht mit Hilfe von Reaktionszeitaufgaben, die Struktur und den zeitlichen Verlauf elementarer mentaler Vorgänge aufzuklären. Donders (1969/1868) hat den ersten Versuch unternommen, die Prozesse einer simplen kognitiven Aufgabe durch Zeitmessung zu analysieren. Seinen frühen Experimenten liegt die Annahme zugrunde, dass einer einfachen Reaktion auf einen Reiz (Detektion) grundsätzlich andere mentale Prozesse zugrunde liegen als der Diskrimination oder der Auswahl eines Reizes und dass diese verschiedenen Prozesse unterschiedlich viel Zeit benötigen. Die Zeitdauer für die Ausführung einer Reaktion wird somit als Indikator für den Ablauf mentaler Prozesse angesehen. Die Reaktionszeit ist nicht nur ein Korrelat der mentalen Aktivität, sondern wird als minimal erforderliche Zeit für die Durchführung einer Reaktion gesehen. Die Interpretation von Reaktionszeitergebnisse erfolgt daher unter der Voraussetzung, dass möglichst wenig Zeit für die Durchführung der Aufgabe benötigt wird und die Fehlerrate gleichzeitig gegen Null geht (sog. speed-accuracy trade-off, vgl. Nährer, 1986).

Beispiele für die Anwendung der Reaktionszeitmethode stellen die klassischen Untersuchungen zum Buchstaben-Matching (letter matching, Posner, 1978), zu Gedächtnissuchaufgaben (memory-scanning tasks, S. Sternberg, 1966, 1969), zur mentalen Rotation (Cooper & Shepard, 1973a, 1973b; Just & Carpenter, 1980; Shepard & Metzler, 1971) und zur Satz-Bild-Verifikationsaufgabe (sentence-picture verification task, Clarke & Chase, 1972) dar. Cooper und Shepard (1973a; 1973b) zeigten beispielsweise in ihren klassischen Untersuchungen zum mentalen Umgang mit visuellen Vorstellungen, dass es umso länger dauerte, bis eine Entscheidung darüber gefällt werden konnte, ob der gezeigte Buchstabe R normal oder gespiegelt dargeboten wurde, je weiter ein Buchstabe von seiner normalen Orientierung rotiert wurde. Offensichtlich wird das mentale Bild der Figur soweit rotiert, bis es seine gewohnte vertikale Lage erreicht hat, bevor ein Urteil gefällt wird. Ähnliche Reaktionszeitbefunde zeigen sich beim visuellen Abtasten (image tracing task, Kosslyn, 1980).

Eine häufig genannte Kritik an diesem klassischen Ansatz der Zeitmessung bezieht sich auf den Mangel an Information über den zugrunde liegenden Bearbeitungsprozess. Um die kognitive Struktur bei Reaktionsaufgaben befriedigend erklären zu können, ist der Informationsgehalt von Reaktionszeiten zu gering. Neuere Untersuchungen verwenden daher zunehmend häufig eine Variation der Reaktionszeitmethode, nach welcher der

Aufgabenlösungs- bzw. Reaktionsprozess segmentiert wird und aktuelle Stadien der Aufgabenbearbeitung registriert werden (S. Sternberg, 1998). Neuerdings werden in der mentalen Chronometrie aber auch ergänzend psychophysiologische Verfahren eingesetzt. Eine Reihe von Untersuchungen widmet sich dabei der Frage, welche Verarbeitungsabschnitte zwischen Reizaufnahme und Reaktionsausführung (z.B. frühe sensorische oder späte motorische Verarbeitung) durch die Vorbereitung einer Handlung verkürzt werden können (Leuthold, Sommer & Ulrich, 1996; Sanders, 1998). Da sich die Reaktionszeit aus der Dauer vieler Teilprozesse (z.B. Wahrnehmung, Handlungsauswahl, motorische Ausführung) zusammensetzt, ist eine Kombination traditioneller Methoden der Reaktionszeitforschung mit modernen Methoden der Chronopsychophysiologie, wie z.B. mit dem Lateralisierten Bereitschaftspotential (lateralized readiness potential, LRP; vgl. Abschnitt 2.2.3.1); siehe hierzu besonders Müller-Gethmann, Rinke, Stahl und Ulrich (2000).

Reaktions- oder besser Lösungszeiten und Fehlerraten stellen auch in Untersuchungen zum Problemlösen häufig verwendete Erfolgsmaße mit hoher Reliabilität dar. Dabei wird die Lösungszeit jedoch meistens nicht als diagnostisches Maß für elementare kognitive Prozesse (z.B. die Geschwindigkeit der Informationsverarbeitung) gesehen, sondern die für die Lösung zur Verfügung gestellte maximale Zeitdauer ist als Randbedingung bei der Bearbeitung des Problems zu berücksichtigen und kann auch als Ausschlusskriterium verwendet werden. Die Bestimmung der Problemlösequalität kann dann der Bearbeitungszeit, der Verarbeitungstiefe bzw. dem Ausmaß des erworbenen Wissens, der Fehlerrate o.ä. gegenübergestellt werden.

Reaktionszeitmessungen sind aus mehreren Gründen kritisch zu sehen. Neben der Frage, inwieweit Zeitmessungen überhaupt valide sind, also Rückschlüsse auf kognitive Prozesse zulassen, ist der mit Reaktionszeitmessungen verbundene hohe experimentelle Aufwand zu nennen. Aufgrund der intraindividuellen Variabilität und den auftretenden Lern- und Ermüdungseffekten sind meist viele Versuchsdurchgänge notwendig. Hinzu kommt eine hohe interindividuelle Varianz der Reaktionsgeschwindigkeit, was vor allem Probleme für Between-subjects-Designs schafft. Aussagen über die Reaktionsgeschwindigkeit sind zudem immer mit Aussagen über die Reaktionsqualität konfundiert. Auf den wesentlichsten Kritikpunkt im Kontext kognitionspsychologischer Untersuchungen deutet jedoch die Erweiterung von Reaktionszeitmessungen um psychophysiologische Verfahren hin: Durch bloße Reaktionszeitmessung bleibt ungeklärt, welcher kognitive Prozess die mentalen Ressourcen beansprucht. Zwar können Aussagen über bestimmte Eigenschaften, nicht jedoch über seinen genauen Ablauf gemacht werden. Um Reaktionszeitdaten sinnvoll nutzen zu

können, ist daher ein streng deduktivistischer Ansatz (Erdfelder & Funke, 2004) und die Verwendung zusätzlicher, z.B. verbaler, bildgebender oder verhaltensbeobachtender Verfahren notwendig.

### **Blickbewegungen**

Nahezu alle unsere Handlungen und die kognitive Vorbereitung derselben sind von Bewegungen unserer Augen begleitet (Deubel, 1994; Land & Hayhoe, 2001; Yarbus, 1967). Augenbewegungen (a) dienen der Ausrichtung des Auges auf ein Sehobjekt, (b) ermöglichen scharfes Sehen und (c) verhindern ein Verwischen des Umgebungsbildes bei eigener Bewegung ("saccadic suppression", Carpenter, 1988; Land, 1999). Als Blickbewegungen werden solche Bewegungen des Auges bezeichnet, die in Verbindung mit den vom Auge aufgenommenen Informationen interpretiert werden. Bei der Erfassung von Blickbewegungen muss neben der Augenbewegung definitionsgemäß also auch immer der Zielort der Augen mit erfasst werden (Rötting, 2001). Zwischen sensorischer Wahrnehmung, kognitiver Verarbeitung und Handlungsmotorik besteht also ein enger Zusammenhang.

Blickbewegungsuntersuchungen im Bereich der kognitiven Psychologie liegt meistens die dogmatische Annahme zugrunde, dass Blickbewegungen Aufschluss über höhere psychologische Prozesse erlauben (sog. „eye-mind hypothesis“, Just & Carpenter, 1980). Dabei wird argumentiert, dass exploratorische Blickbewegungen ja keine sensomotorischen Reflexe seien und daher zumindest zu einem gewissen Ausmaß mentale Prozesse abbilden müssen (Viviani, 1990). Es wird postuliert, dass ein Großteil der visuellen Information während einer Fixation, bei der sich das Auge in relativer Ruhe zu einem Sehobjekt befindet, aufgenommen und verarbeitet wird. Die Dauer der Fixation, so die Annahme, entspreche der Dauer der zentralen Verarbeitung (Just & Carpenter, 1980). Zentrale Vorgänge der Informationsverarbeitung seien somit bei visuell dargebotenem Aufgabenmaterial an extern verfügbare Informationen gebunden. Somit lassen sich aus Blickbewegungen Aussagen über nicht direkt beobachtbare serielle Vorgänge der Informationsverarbeitung treffen (Just & Carpenter, 1976; Schroiff, 1987). Dass dieses zugrunde gelegte Dogma aus verschiedenen Gründen problematisch ist (z.B. Anderson, Bothell & Douglas, 2004), wird am Ende dieses Abschnittes gezeigt.

Blickbewegungen werden in der kognitiven Psychologie seit den 1960er Jahren eingesetzt. Mit der Weiterentwicklung und Vereinfachung der Messmethoden werden Blickbewegungen als Methode für kognitionspsychologische Untersuchungen zunehmend interessanter. Neben

Studien aus dem Bereich der Denk- und Problemlöseforschung sollen hier auch Studien aus anderen Bereichen der kognitions- und sprachpsychologischen Forschung dargestellt werden, sofern sie von Relevanz für problemlösendes Denken sind. Zunächst sollen aber relevante Augenbewegungsparameter (insbesondere Sakkaden und Fixationen) kurz dargestellt werden sowie verschiedene Messmethoden erläutert und verglichen und beispielhaft Anwendungsbereiche für das Verfahren der Blickbewegungsmessung vorgestellt werden.

### Blickbewegungsparameter: Sakkaden und Fixationen

Die für die Analyse von Blickbewegungen im Kontext der Denk- und Problemlöseforschung relevanten Augenbewegungsparameter sind vor allem Sakkaden und Fixationen und das sich aus diesen Parametern ergebende Blickmuster (scanpath).

Sakkaden sind schnelle, sprunghafte, ballistische Bewegungen des Auges, die unwillkürlich (reflexhaft) oder gezielt ausgelöst werden können (Carpenter, 1988) und dazu dienen, relevante Information auf die Fovea, den Punkt des schärfsten Sehens, zu bringen. Sakkaden lassen sich durch ihre räumlichen und zeitlichen Eigenschaften, ihre Schnelligkeit, Amplitude, Dauer, Latenz und ihren zeitlichen Verlauf beschreiben. Welche Aspekte analysiert werden, hängt von der Zielsetzung der jeweiligen Studie ab. Eine Übersicht über Sakkadeneigenschaften sowie Kriterien zur Bestimmung von Sakkaden und Fixationen geben Deubel (1994) und Rötting (2001).

Sakkadenlatenzen können in der kognitiven Psychologie analog zu Reaktionszeiten verwendet werden. Die Dauer zwischen dem Auftauchen (oder Verschwinden) eines Reizes bis zum Onset der Sakkade ca. 200 msek später kann als Zeit betrachtet werden, die das kognitive System bis zum Fällen eines Urteils (z.B. für die Detektion, Diskrimination oder Auswahl) benötigt. Darüber hinaus kann die Landungsposition einer Sakkade bzw. ihre Genauigkeit bzgl. eines bestimmten Zielobjekts ein relevantes Maß sein. In diesem Kontext werden üblicherweise sog. „regions of interest“, also virtuelle Grenzen zwischen den einzelnen interessierenden Objekten, Wörtern, Satzteilen, Bildaspekten, o.ä., definiert. Sakkaden innerhalb dieser Regionen werden meistens als Korrektursakkaden verstanden.

Für kognitionspsychologische Untersuchungen sind die Fixationsdauer und die Fixationshäufigkeit interessante Parameter. Die Fixationsdauer – so die Annahme – lässt Rückschluss auf die Intensität und Tiefe der Informationsverarbeitung zu, während die Fixationshäufigkeit und insbesondere die Abfolge mehrerer Sakkaden und Fixationen (Blickmuster oder scanpath) Aufschluss über verwendete Strategien liefern können.

Gelegentlich werden mehrere Fixationen gemeinsam betrachtet, insbesondere wenn sie nur von Korrektursakkaden unterbrochen werden (sog. Blickdauer). Die gesamte Blickdauer auf ein Element („total viewing time“) wird ebenfalls analysiert.

Während ein Großteil der visuellen Informationsverarbeitung während der Fixation stattfindet, ist die Informationsaufnahme während einer Sakkade eher beschränkt (vgl. hierzu die Diskussion zum Phänomen der sakkadischen Suppression, Burr, Morrone & Ross, 2001). Dennoch werden auch während einer Sakkade Informationen verarbeitet (Irwin, 1992; Irwin & Gordon, 1998). Einen Überblick über neurologische Mechanismen zur Generierung von Sakkaden und Modelle der Sakkadensteuerung geben Carpenter (1988), Deubel (1994) sowie Leigh und Zee (1999).

### Messung von Blickbewegungen

Zur Registrierung von Blickbewegungen gibt es eine Reihe von Messverfahren, von denen im Folgenden die gängigsten Methoden erläutert und verglichen werden sollen. Für eine umfassendere Darstellung aller Verfahren sei auf Carpenter (1988; Collewyn, 1999; Rötting, 2001) verwiesen. In Tabelle MESS sind die Verfahren und die zugrunde liegenden anatomisch-physiologischen Prinzipien zusammengefasst. Sowohl Eigenschaften der Pupille, das corneo-retinale Potenzial zwischen dem negativen Pol der Retina (Netzhaut) und dem positiven Pol der Cornea (Hornhaut), die Krümmung der Cornea (Hornhaut) und Reflexionen an Grenzflächen des dioptrischen Apparats können zur Bestimmung der Augenbewegungen dienen.

---

Hier bitte Tabelle MESS einfügen

---

Die Blickbewegungsregistrierung mit dem Elektrookulogramm (EOG) basiert auf der zwischen Retina und Cornea bestehenden elektrischen Potenzialdifferenz. Um die geringen Spannungsdifferenzen erfassen zu können, werden rechts und links, bzw. oberhalb und unterhalb der Augen Oberflächenelektroden platziert, mit denen die horizontale bzw. vertikale Ausrichtung gemessen werden kann. Ein Blick nach rechts oben bspw. hat zur Folge, dass der rechten und der oberen Elektrode mehr positiv geladene Hornhaut zugewandt wird, was mit Hilfe der jeweiligen Verstärker als die entsprechende Blickbewegung aufgezeichnet wird. Aufgrund der erheblichen Einschränkung im Messbereich und der geringen örtlichen Genauigkeit, insbesondere bezüglich der vertikalen Komponente, sind EOG-Messungen für Fragestellungen aus dem kognitionspsychologischen Bereich nur bedingt geeignet und werden zumeist eher im klinischen Kontext eingesetzt.

Bei der Cornea-Reflex-Methode (CRM) wird die Auslenkung des Auges mit Hilfe von Reflexionen des Lichts auf der Cornea erfasst. Das auf der Cornea auftreffende Licht spiegelt sich in einer punktförmigen Quelle, dem Cornealreflex oder 1. Purkinje-Bild, das etwa 3,5 mm hinter der Augenoberfläche zu liegen scheint. Bei einer Augenbewegung kommt es zu einer Verschiebung des Cornealreflexes, der mittels eines videobasierten Systems erfassbar ist. Für die CRM existieren kopfgestützte und berührungslose Systeme. Bei kopfgestützten Systemen sind die Videokameras zur Beobachtung der Augen am Kopf befestigt. Auf diese Weise wird lediglich die Augenbewegung relativ zum Kopf gemessen, so dass zusätzliche Verfahren zur Registrierung von Kopfbewegungen notwendig werden. Da kopfgestützte Eyetracker den berührungslosen Systemen an Genauigkeit der Blickregistrierung überlegen sind, werden sie in der physiologischen und psychologischen Forschung meist vorgezogen. Auch Kosten-Nutzen-Betrachtungen machen diese Systeme zu beliebten Instrumenten kognitionspsychologischer Forschungslabore.

Die Kontaktlinsemethode, bei welcher dem Proband eine Kontaktlinse mit magnetischen Spulen eingepasst wird, gehört dagegen zu den genauesten, aber auch aufwändigsten Verfahren zur Blickbewegungsregistrierung. Der Kopf des Probanden befindet sich in einem magnetischen Wechselfeld. Die sich bei Bewegung proportional mit der Lage der Spulen im Raum verändernde, über die Kontaktlinse induzierte Spannung, wird gemessen. Da die Tragedauer der Kontaktlinsen auf etwa 30 Minuten pro Messung begrenzt ist, ist die Methode in ihrer Anwendung auf Aufgaben mit kurzer zeitlicher Dauer begrenzt.

Die Doppelte Purkinje-Bild-Technik (DPI) macht sich – wie die Cornea-Reflex-Methode – die Tatsache zunutze, dass Lichtstrahlen an verschiedenen Oberflächen des Auges gebrochen werden. Neben dem 1. Purkinje-Bild gibt es weitere Reflexionen an der Grenzfläche zwischen Cornea und Kammerwasser (2. Purkinje-Bild), Kammerwasser und Linse (3. Purkinje-Bild) und Linse und Glaskörper (4. Purkinje-Bild). Das 1. und das 4. Purkinje-Bild sind besonders geeignet zur Bestimmung der Blickrichtung, da sich ihre Lage zueinander nur durch rotatorische und nicht durch translatorische Bewegungen des Auges ändert. Für eine ausführliche Beschreibung des Messprinzips sei auf Crane und Steele (1985) verwiesen. Die DPI-Technik kommt aufgrund ihrer hohen zeitlichen Auflösung und räumlichen Genauigkeit vornehmlich in physiologischen und psychophysikalischen Untersuchungen zum Einsatz. Da für kognitionspsychologische Fragestellungen, d.h. für die Untersuchung insbesondere von Fixationszeiten und Sakkaden, die hier erreichte hohe Auflösung nicht notwendig ist, kommt

die DPI-Methode aufgrund der erheblichen Anschaffungs- und Wartungskosten des Messgeräts sowie des für die Probanden nötigen Trainings eher selten zum Einsatz.

Neben der Hardware (dem Eyetracker) steht für bestimmte Messgeräte zur Erhebung von Blickbewegungen und zur Präsentation der Stimuli gut dokumentierte Software zur Verfügung, so dass die Steuerung des Eyetrackers und die Stimuluspräsentation relativ komfortabel erfolgen können. Ein Beispiel ist die als Freeware erhältliche Eyelink-Toolbox (Cornelissen, Peters & Palmer, 2002), die ein Interface zwischen MATLAB, einem Programm für mathematische Berechnungen, und dem Eyelink-Blickbewegungsmessgerät (SR Research Ltd., Ontario, Canada) darstellt. Mit der Toolbox lassen sich Blickbewegungen messen, während über die ebenfalls MATLAB-basierte Psychtoolbox (Brainard, 1997) die Stimuluspräsentation erfolgt.

#### Auswertung von Blickbewegungsdaten

Blickbewegungsdaten können mit hoher zeitlicher und räumlicher Auflösung gemessen werden und bieten sich als mögliches motorisches Korrelat zur Untersuchung raum-zeitlicher Charakteristika der zugrunde liegenden kognitiven Prozesse an. Mit der fortgeschrittenen Erhebungstechnik ist vergleichsweise wenig Übung bzw. Aufwand für die Probanden und Versuchsleiter verbunden. Ein erhebliches Problem kann aber nach wie vor die Datenauswertung darstellen, und zwar aus drei Gründen: (1) Aufgrund der hohen zeitlichen Auflösung wird eine große Datenmenge und damit ein großer Suchraum produziert (dies erfordert zudem eine gewisse Rechnerkapazität), (2) abhängig vom System und der Art der Kalibrierung können die Daten verrauscht sein und müssen u.U. aufwändig nachbearbeitet werden und (3) auch bei exakter Messung spiegeln die Daten ein erhebliches Maß an individueller Variabilität wider. Neben gängigen Auswertungsprogrammen (z.B. MATLAB, The MathWorks, 2003) gibt es auch andere „Tracing“-Verfahren, wie die sog. Protokollanalyse (Salvucci & Anderson, 2001) zur Umwandlung von Rohdaten in Sequenzen von Sakkaden und Fixationen. Die Hersteller einiger Blickbewegungsmessgeräte stellen auch Software zur Auswertung zur Verfügung (wie z.B. der Data Viewer von Eyelink, SR Research Ltd.). So lassen sich bspw. einfach „regions of interest“ auswählen und Blickdauer und –häufigkeit für den gewählten Bildausschnitt oder das relevante Objekt berechnen sowie graphisch darstellen.

### Empirische Befunde: Blickbewegungsdaten in kognitionspsychologischen Studien

Studien zu kognitiven Prozessen, in denen Blickbewegungsanalysen verwendet werden, lassen sich einerseits nach dem Ausmaß der Kontrolle durch den Beobachter bzw. andererseits des Versuchsleiters über die Wirkung des Stimulus kategorisieren (Findlay & Gilchrist, 2003). Blickbewegungsstudien mit einfachen Paradigmen stammen bspw. aus der Forschung zur visuellen Suche und zum Betrachten und Wiedererkennen von Bildern und natürlichen Szenen. In Untersuchungen zu höheren kognitiven Prozessen wie z.B. einfachen Denk- und Problemlöseaufgaben und Entscheidungssituationen hat der Beobachter mehr Kontrolle über den Verlauf der Handlung. Manipulative Aufgaben aus dem Bereich alltäglicher Verrichtungen (z.B. Tee kochen, Hände waschen) sowie dynamische Situationen (Sport, Auto fahren) stellen ein Höchstmaß an Anforderungen an den Beobachter. Gleichzeitig hat der Versuchsleiter hier weniger Kontrolle über die Wirkungsweise des Reizes und kausale Rückschlüsse auf kognitive Prozesse sind schwieriger.

In den folgenden Abschnitten sollen aus den Bereichen (a) Denken und Problemlösen sowie (b) Handlungsplanung und –verrichtung Studien exemplarisch vorgestellt werden, wobei das Hauptaugenmerk auf Studien zum simplen und komplexen Problemlösen liegt.

#### *Blickbewegungsdaten in Studien zum Denken und Problemlösen*

Blickbewegungsstudien zu Denk- und Problemlöseaufgaben lassen sich danach einteilen, welche kognitiven Prozesse im Mittelpunkt des Interesses stehen. Blickbewegungen sollten unter anderem Aufschluss über zugrunde liegende Gedächtnisprozesse, den Vorgang der Informationsverarbeitung und –reduktion, den Aufbau mentaler Modelle, den Unterschied zwischen Experten und Novizen und die Frage der Strategiewahl beim Problemlösen liefern. Darüber hinaus wird mit Hilfe von Blickbewegungsstudien der Versuch unternommen, den Problemlöseprozess in verschiedene Phasen aufzuteilen und zu analysieren.

Was zugrunde liegende Gedächtnisprozesse betrifft, so zeigen Blickbewegungsanalysen bei der Betrachtung einer visuellen Szene oder bei der visuellen Suche, dass das Arbeitsgedächtnis für Inhalte einer visuellen Szene relativ beschränkt ist. Mit Hilfe von Sakkaden werden bestimmte physikalische Bestandteile einer Szene oder Aufgabe immer wieder betrachtet (Ballard, Hayhoe & Pelz, 1995; Epelboim & Suppes, 2001; Zelinsky & Loschky, 1998). Irwin (1992) postuliert, dass aufgrund der begrenzten Kapazität des Arbeitsgedächtnisses über sakkadische Augenbewegungen hinweg eine nur unzureichende Repräsentation der Identität und Lokalisation von Objekten in einer visuellen Szene erreicht

wird. Studien zur Veränderungsblindheit (Simons & Levin, 1997) scheinen diese These zu belegen. Entsprechend behaupten Groner und Fraisse (1982), O'Regan (1992) sowie O'Regan et al. (2000), dass der Organismus die Umwelt als externalen Speicher nutze, um die Arbeitsgedächtnisbelastung zu reduzieren, und dass er auf diese Informationen genau dann zurückgreife, wenn es für die Aufgabe günstig sei. Das Abtasten der visuellen Umgebung durch Blickbewegungen erfolge analog zum Durchsuchen mentaler Modelle im Arbeitsgedächtnis. Somit könne aus dem Blickmuster darauf geschlossen werden, welche Gedächtnisinformationen zur Lösung einer Aufgabe als relevant erachtet werden.

Gleiches gilt jedoch nicht für alltägliche Planungsaufgaben (Hayhoe, Shrivastava, Mruczek & Pelz, 2003, siehe unten; Land & Furneaux, 1997) und für Problemlöseaufgaben: Lüer, Hübner und Lass (1985) fanden, dass Probanden mit zunehmender Erfahrung konstante Teile der Umgebung in einer Problemlöseaufgabe (im komplexen Szenario „Schneiderwerkstatt“) weniger häufig fixierten. Die Autoren vermuten, dass diese Information im Gedächtnis gespeichert wurde und daher keine visuelle Aufmerksamkeit mehr benötigt. Die unterschiedlichen Blickmuster bei Betrachtungsaufgaben versus bei Denk- bzw. Planungsaufgaben sind durchaus vereinbar, wenn man berücksichtigt, dass beide Aufgabentypen ganz unterschiedliche Anforderungen an die Versuchsperson stellen. Eine Versuchsperson, die in einer Situation planen, entscheiden und handeln muss, ist ungleich stärker gefordert, als in einer Aufgabe zur visuellen Suche. Eine erhöhte Aufmerksamkeit kann dann bspw. zur Folge haben, dass die Sensitivität für redundante Objekte erhöht wird und diese nicht mehr fixiert werden.

Untersuchungen zu Informationsverarbeitungs- und Informationsreduktionsprozessen beim Problemlösen (z.B. bei kryptarithmetischen Problemen, bei komplexen Problemlösesimulationen, in Entscheidungssituationen und bei verbalen Aufgaben) stammen u.a. von Haider und Frensch (1999), Lüer, Hübner und Lass (1985), Newell und Simon (1972), Nodine, Kundel, Polikoff und Toto (1988) sowie Russo und Rosen (1975). In diesen Studien werden erfolgreiche und weniger erfolgreiche Problemlöser hinsichtlich ihrer Informationsverarbeitungsstrategien verglichen.

Im Kontext der Unterschiede zwischen Experten und Novizen beim Problemlösen sind zunächst die Arbeiten zu Expertiseunterschieden beim Schach (Chase & Simon, 1973; Simon & Barenfeld, 1969) zu nennen. Experten führen im Gegensatz zu Laien ständige Paarvergleiche zwischen Figuren aus, die in einer strategischen Beziehung (z.B. einem Angriffs-Verteidigungs-Verhältnis) zueinander stehen. Zudem lassen sich aus dem

Blickbewegungsmuster von Experten die alternativen, hypothetisch ausgeführten Züge ablesen, die allerdings nur mit Hilfe von online erhobenen Laut-Denk-Protokollen klar einem bestimmten Blickmuster zugeordnet werden konnten (de Groot, 1978).

Blickbewegungen erlauben offensichtlich Rückschluss auf das Ausmaß der Erfahrung bzw. Expertise, das ein Problemlöser mit einer Aufgabe erworben hat. Eine Vielzahl von Studien (Dillon, 1985; Haider & Frensch, 1999; Lüer et al., 1985; Lüer, Lass & Shallo-Hoffmann, 1988; Suppes, Cohen, Laddaga, Anliker & Floyd, 1982), in denen unterschiedliche Aufgaben (arithmetische Aufgaben, induktives Schließen, verbale Aufgaben, komplexes Problemlösen im „Tailorshop“, Konzeptlernen, Lösen von Algorithmen) bei Erwachsenen und Kindern verwendet wurden, kommt zu ähnlichen Ergebnissen: Experten identifizieren schneller die für die Lösung relevanten Aspekte einer Aufgabe (mehr Fixationsorte), nehmen schneller die relevanten Informationen auf (kürzere Fixationszeiten) und weisen insgesamt ökonomischere Blickmuster auf. Novizen hingegen verharren häufiger und länger auf für die Lösung irrelevanten Aspekten der Aufgabe. Gleichzeitig werden relevante Teile der dargebotenen Aufgaben wenig oder gar nicht beachtet. Anhand dieser plastischen Unterschiede könnte eine Vorhersage über den Erfolg eines Probanden schon anhand seines Blickmusters möglich sein (siehe Grant & Spivey, 2003). Groner und Groner (1989) unterscheiden zwei Arten von für das Problemlösen relevanten Fixationen: Aufnahme-fixationen (ca. 100-200 msec) und Verarbeitungsfixationen (ca. 300 msec). Möglicherweise lassen sich Experten und Novizen auch hinsichtlich dieser Fixationsarten und ihres Auftretens in verschiedenen Phasen des Problemlöseprozesses unterscheiden.

Die Aufzeichnung von Blickbewegungen bei Denk- und Problemlöseaufgaben (geometrische Puzzle-Aufgabe, Deffner, 1984; z.B. Raven-Matrizen, Putz-Osterloh, 1981a; Würfelaufgabe, Putz-Osterloh & Lüer, 1979; Satzanagrammaufgaben, Rhenius & Locher, 1992) soll weiterhin Aufschluss über den Ablauf eines Lösungsprozesses und die verwendeten Lösungsstrategien geben. Just und Carpenter (1985) sowie Putz-Osterloh und Lüer (1979) untersuchten die Augenbewegungen von Probanden bei verschiedenen räumlich-visuellen Aufgaben, unter anderem den Würfelaufgaben aus dem IST-70. Bei diesen Aufgaben geht es darum, den richtigen von mehreren möglichen gedrehten Würfeln einem Modellwürfel zuzuordnen. Diese Aufgabe kann man durch einfache Zuordnungsstrategien oder durch mentale Rotation und räumliche Vorstellung lösen. Putz-Osterloh und Lüer (1979) konnten anhand von Augenbewegungsdaten entscheiden, welche Strategie die jeweiligen Probanden benutzen: Steigen die Fixationszeiten erheblich an und nimmt die Häufigkeit von Sakkaden

stark ab, so deutet dies auf eine Verwendung visueller mentaler Modelle hin.

Performanzunterschiede können dagegen nicht als Indikator für die Verwendung einer bestimmten Strategie genutzt werden, da je nach Art der Aufgabe unterschiedliche Strategien zum Erfolg führen können. Umgekehrt können prototypische Strategien definiert werden, die in spezifischen Blickbewegungsmustern resultieren, etwa Flächen- versus Raumstrategien beim Bearbeiten von Raumvorstellungsaufgaben (vgl. Plata, 2000).

Zur Beziehung zwischen der internalen Repräsentation von Problemen und der daraus resultierenden Problemlösestrategie untersuchte Krause (1981; 1988) Blickbewegungen beim Lösen von linearen Ordnungsproblemen der Form: „F ist größer als B und K ist kleiner als B“. Diesen Satz verstehen Probanden schneller als den Satz „B ist größer als K und B ist kleiner als F“, was sich in der Anzahl der Regressionssakkaden niederschlägt. Regressionen dienen, so der Autor, dem Aufbau eines mentalen Modells und sind in ihrer Anzahl abhängig von der Zahl der nötigen Operationen und dem Vorwissen des Probanden. In einer ähnlichen Untersuchung zu Relationsproblemen identifizierte Deffner (1987) zwei Lösungsstrategien: die Serienbildung, bei welcher ein räumliches Bild der Relationsbeziehung aufgebaut wird, und die Eliminationsstrategie, die ohne räumliche Vorstellung auskommt und auf Vergleichsoperationen beruht. Aus Simulationsversuchen wurden ideale Blickbewegungsmuster für die jeweiligen Strategien abgeleitet und die empirisch gefundenen Blickmuster dann mit diesen Idealnormen verglichen. Durch Validierung mittels Laut-Denk-Daten konnte gezeigt werden, dass Blickbewegungsmaße ein Indikator für die benutzte Strategie beim Problemlösen sein können.

Nodine et al. (1988) untersuchten Entscheidungsverhalten und Blickbewegungen von Radiologen bei der Interpretation von Lungenradiogrammen. Interessanterweise wiesen Radiologen bei der Betrachtung von Gewebe, welches sie später als Falsch-Negativ (d.h. übersehenes Krebsgewebe) klassifizierten, die gleichen Augenbewegungsmuster (z.B. längere Fixationszeiten) auf wie bei der Betrachtung von Gewebe, welches sie richtig als krankhaft identifizierten. Blickbewegungen bei der Klassifizierung von Falsch-Positiven Gewebeteilen (d.h. als Krebsgewebe erkanntes gesundes Gewebe) waren unauffällig. Die Augenbewegungsmuster der Experten können also etwas über deren implizite Informationsverarbeitung aussagen und erlauben so unter Umständen eine bessere Zuordnung von krankem und gesundem Gewebe als das explizite Urteil der Experten.

Dass sich Strategieunterschiede in Blickbewegungen widerspiegeln und sich Problemlöseerfolg anhand von Fixationsorten und –zeiten vorhersagen lässt, wird auch in

Untersuchungen zur Bearbeitung des „Turm von London“ (TVL; eine Abwandlung des „Turm von Hanoi“) durch gesunde Probanden und Morbus Parkinson Patienten (Timothy L. Hodgson, Adnan Bajwa, Adrian M. Owen & Christopher Kennard, 2000a; Hodgson, Tiesman, Owen & Kennard, 2002) und zum Lösen von Einsichtsproblemen (Streichholzarithmetik, Knoblich et al., 2001) deutlich. Diesen Studien ist erstens das Ergebnis gemeinsam, dass sich erfolgreiche und weniger erfolgreiche Problemlöser anhand ihrer Blickbewegungsmuster unterscheiden lassen. Problemlöser, die in der für die Lösung vorgesehenen Zeit keine Lösung bei den Streichholzarithmetik-Aufgaben finden, weisen durchschnittlich niedrigere Fixationszeiten auf als erfolgreiche Problemlöser (Knoblich et al., 2001). Erfolgreiche Problemlöser fixieren zudem in der letzten Phase ihrer Lösungsfindung das lösungsrelevante Element in der Aufgabe (bspw. den Operator, der verändert werden muss). Zweitens lassen sich aufgabenspezifische und phasenspezifische Blickmuster erkennen. Im TVL (Neuanordnung von Bällen im „Problemraum“ nach einem Vorbild im „Lösungsraum“) wenden Probanden ihren Blick vor allem in der ersten und letzten Phase der Problemlösung dem Lösungsraum zu, während sie ihren Blick in der mittleren Phase der Problemlösung schwerpunktmäßig auf den Problemraum richten (vgl. Abbildung TVL).

---

Hier bitte Abbildung TVL einfügen

---

Daraus lässt sich ableiten, dass die Problemlösung bei Aufgaben, die mehr als zwei Schritte erfordern, in mehreren Phasen (einer Analyse-, einer Elaborations- und einer Verifizierungsphase) abläuft (siehe dazu auch die Befunde von Putz-Osterloh (1981a) zu Problemlösephasen bei Würfelaufgaben). Hodgson et al. (2000a) ziehen aus diesen Befunden den Schluss, dass Blickbewegungsstrategien gezielt eingesetzt werden, um die Arbeitsgedächtnisbelastung zu jedem Zeitpunkt des Problemlösevorgangs möglichst gering zu halten, m.a.W.: Probleme werden „mit den Augen“ und nicht „mit dem Kopf“ gelöst (siehe auch Ballard, Hayhoe, Pook & Rao, 1997).

#### Handlungsplanung und -ausführung bei Alltagsverrichtungen

Untersuchungen zu Augenbewegungen bei dynamischen, automatisierten Handlungen, die höhere kognitive Prozesse erfordern (z.B. Tee kochen, Auto fahren; siehe Abbildung TEA), haben ergeben, dass aufgabenrelevante Objekte und Lokationen sofort fixiert werden, ohne dass durch sakkadische Augenbewegungen erst relevante Informationen einer visuellen Szene ausgewählt werden müssen (Land & Furneaux, 1997; Land & Hayhoe, 2001).

---

Hier bitte Abbildung TEA einfügen

---

Aufgabenirrelevante Objekte werden zu einem erstaunlich geringen Ausmaß (Anteil von weniger als 5%, gemessen an allen Blickbewegungen) fixiert, auch wenn sie eine gewisse Salienz besitzen. Die Studien ergeben, dass nicht nur motorische Handlungen visuell gesteuert sind, sondern dass auch kognitive Planungstätigkeiten von Blickbewegungen begleitet werden und Blickbewegungen somit objektbezogene Handlungen vorbereiten und steuern (Land, 1999; Land & Lee, 1994). Insgesamt wurden die Hände relativ selten fixiert, dafür aber das Objekt, das gegriffen wurde. Sobald sich das Objekt jedoch in der Hand befand, wanderte der Blick vom Objekt zur Zielposition. Konkret haben Blickbewegungen die Funktionen (a) der Lokalisation handlungsrelevanter Objekte im Raum, (b) der Bestimmung der Richtung der Bewegung, (c) der visuellen Kontrolle von Handlungen und (d) des Überprüfens, ob durch eine Handlung das angestrebte Ziel erreicht wurde, z.B. durch ständiges Monitoring des Wasserstandes beim Füllen des Wasserkessels (Land & Hayhoe, 2001).

### Fazit

Ein Vergleich von Blickbewegungsdaten und LD-Protokollen zur konkurrenten Validierung beider Verfahren deutet auf eine relativ hohe Übereinstimmung zwischen den aus verbalen Auskünften gewonnenen Informationen über den Aufmerksamkeitsfokus des Probanden und Fixationsorte und -dauer hin (Knoblich & Rhenius, 1995; Newell & Simon, 1972; Russo & Rosen, 1975). Diese Übereinstimmung spricht nochmals dafür, dass der Ort der Blickbewegung unter optimalen Bedingungen Rückschluss auf den Gegenstand der Informationsverarbeitung zulässt, über den der Beobachter verbal Auskunft geben kann.

Kritisch anzumerken bleibt: Die dargestellten Studien verwenden ausschließlich statische Bilder und Szenen. Interessant wäre jedoch – insbesondere vor dem Hintergrund der komplexen Problemlöseforschung mit computersimulierten Szenarien – die Verwendung dynamischer, bewegter Oberflächen. Von diesem Ziel ist die kognitionspsychologische Blickbewegungsforschung jedoch noch weit entfernt, da schon die Abstrahierung mentaler Prozesse aus der Fixation statischer Information genug Probleme aufwirft.

Aus der genannten Kritik wird deutlich, dass eine Verwendung von Blickbewegungsdaten in der kognitiven Psychologie nur unter der Annahme einer streng deduktivistischen Vorgehensweise möglich ist: „Eye movements will only become useful in the context of a deductive approach: not until we conjure up a reasonably articulated theory of these processes will data on overt behavior truly contribute to our understanding [...]” (Viviani, 1990, S. 354). In der Literatur sind daher Anforderungen an eine Theorie zu Blickbewegungen als

Indikatoren kognitiver Prozesse formuliert worden (vgl. z.B. Ballard et al., 1997; Kowler, 1990; Viviani, 1990), die im Folgenden als Desiderate erörtert werden.

- (1) Aus dem Zitat von Eileen Kowler „Many different eye movement patterns can be observed with the same visual stimulus. Similarly, large changes in the visual stimulus often have no systematic effect on eye movements” (Kowler, 1990, S. 1) folgt die Erkenntnis, die den Blickbewegungen zugrunde liegenden perzeptuellen und okulomotorischen Prozesse und ihre Berechnungsmechanismen zu verstehen. So kann es zu einer hilfreichen Einschränkung des Spielraums bei der Interpretation der Daten kommen und es können sich wesentliche Implikationen für das experimentelle Design sowie die Wahl der Stimuli bzw. die Gestaltung der Präsentationsoberfläche ergeben.
- (2) Blickbewegungen stellen ein seriell ablaufendes Verhalten dar und können daher nur entsprechende kognitive Prozesse abbilden. Es folgt, dass bei parallel ablaufenden Prozessen aus dem seriellen Blickmuster entweder nur ein Prozess oder alle Prozesse überlagert abgeleitet werden können. Der angenommene zugrunde liegende kognitive Prozess muss also im Grunde seriell sein oder in serielle Komponenten unterteilbar sein; nur so können Blickbewegungen Aufschluss über eine Sequenz unbekannter kognitiver Abläufe geben. Ein Beispiel für die Verwendung von Blickbewegungen bei seriell ablaufenden kognitiven Prozessen bieten Ballard et al. (1997) mit der sog. „block copying“-Aufgabe.
- (3) „It appears then that the heuristic validity [der Annahme, dass sich aus Blickbewegungsdaten Rückschluss auf kognitive Prozesse ziehen lässt] hinges upon the crucial issue: what can we say about the structure of the cognitive process? The debate on this issue is quite open” (Viviani, 1990, S. 356). Mit diesem Zitat leitet Viviani (1990) sein Plädoyer für eine Neuformulierung des informationstheoretischen Ansatzes in ein plausibles Subset kognitiver Abläufe, z.B. bei der Lösung eines Problems, ein. Als Basis für die Ableitung mentaler Prozesse aus Blickbewegungen ist erforderlich (a) die Kenntnis von a priori Annahmen des Probanden über mögliche Zustände und Ereignisse, aus denen plausible Hypothesen zur Problemlösung abgeleitet und abgewogen werden, (b) Vereinfachung holistischer und komplexer Hypothesen in simplere Alternativen mit einer visuell-figuralen Korrespondenz und Übersetzung dieser figuralen Hypothesen in einer Sequenz im visuellen Raum (also in ein idealtypisches Blickmuster).

- (4) Der visuelle Informationsgehalt einer Szene kann nicht reliabel gemessen werden: „The trouble is that the information is not out there in the visual field. [...] From the fact that a given point of the scene has been fixated we can only infer that some information might have been gleaned.” (Viviani, 1990, S. 358). Notwendigerweise muss daher etwas über die Vorannahmen und das Vorwissen des Problemlösers bekannt sein. Da sich Vorannahmen und Wissen jedoch im Laufe eines Problemlöseprozesses verändern, besteht hier eine grundsätzliche Schwierigkeit (die möglicherweise durch den zusätzlichen Einsatz von Verbaldaten behoben werden kann).
- (5) Es muss sichergestellt werden, dass der Informationsgehalt einer visuellen Oberfläche (z.B. des dargebotenen Screenshots eines Problemlöseszenarios) nur durch inhaltliche Merkmale und nicht etwa durch physikalische Objekteigenschaften (wie Kontraste, Farben, Konturen) bestimmt wird (vgl. Mackworth & Morandi, 1967; Marr, 1982).
- (6) Bei der Interpretation der Daten aus Blickbewegungsuntersuchungen ist zu berücksichtigen, dass es eine empirisch belegte teilweise Unabhängigkeit von Blickbewegungen und kognitiven Prozessen gibt (Anderson et al., 2004). Zudem sind Aufmerksamkeit und Fixationsort nicht immer koinzident. Mittlerweile gilt es als gesichert, dass der Aufmerksamkeitsfokus ohne jegliche Augenbewegungen verlagert werden kann (z.B. Reeves & Sperling, 1986). Wenn jedoch eine Augenbewegung erfolgt, so geht ihr immer eine Verlagerung der fokalen Aufmerksamkeit voraus (Hoffman & Subramaniam, 1995; McPeck, Maljkovic & Nakayama, 1999). Ob die Augen auch auf ein visuelles Ziel gerichtet werden können, ohne dass eine Aufmerksamkeitsverlagerung notwendigerweise dazu parallel liefere, wird von einigen Autoren als bislang unbeantwortete Frage gesehen (Deubel, 1994).

Zusammenfassend kann festgehalten werden, dass verhaltensbeobachtende Messmethoden nur dann sinnvoll Anwendung finden können, wenn die zugrunde liegenden internalen kognitiven Prozesse ihrer Art und ihrem Ablauf nach im wesentlichen bekannt sind und Reaktionslatenzen, Fixationszeiten und Blickmustern zugeordnet werden können. Die von Just und Carpenter (1980) geäußerten Vorannahmen, die für die Verwendbarkeit von Blickbewegungsdaten in der Kognitionspsychologie erfüllt sein müssen, treffen also auch heute noch voll zu: (1) Es wird das Objekt fixiert, das tatsächlich gerade Gegenstand der zentralen Verarbeitung ist („eye-mind assumption“), (2) die Fixationsdauer entspricht der Dauer der zentralen Verarbeitung („immediacy assumption“) und (3) aus der Sequenz der

Fixationen lässt sich die Abfolge der zentralen Verarbeitungsschritte rekonstruieren (siehe auch Schroiff, 1987). Solange die hier kritisch aufgeworfenen Fragen nach der Validität der Methode nicht geklärt sind, sollten Blickbewegungsdaten in der kognitiven Psychologie nur für gut formulierte und operationalisierte Fragestellungen angewendet und mit Vorsicht interpretiert werden.

### **Ausdrucksbeobachtung (Mimik, Gesten und Bewegungen)**

Neben ihren Blickbewegungen kann auch der Ausdruck einer Person, also ihre Gestik und Mimik, beim Problemlösen kognitive Relevanz besitzen. (Garber & Goldin-Meadow, 2002) überprüften die Übereinstimmung von gestischen und verbalen Äußerungen von Erwachsenen und Kindern beim Bearbeiten des „Turm von Hanoi“. Ausgehend von der Beobachtung, dass die Verbalisierung von Problemlösestrategien häufig durch Gestik begleitet wird, wird angenommen, dass sich aus der Gestik eines Problemlösers der Fokus der geplanten Handlung bzw. die Problemlösestrategie ablesen lasse. Gestik und sprachliche Äußerung können hierbei übereinstimmend eine Strategie zum Ausdruck bringen, oder bei Nichtübereinstimmung (im Fall eines sog. „mismatch“) eine zweite Strategie anzeigen. Aus solchen Nichtübereinstimmungen zwischen Zeige- und Handbewegungen in Richtung der Aufgabe und verbalen Äußerungen leiteten Garber und Goldin-Meadow (2002) die Stellen im zeitlichen Verlauf der Problemlösung ab, an denen Versuchspersonen zwischen zwei möglichen Lösungswegen entscheiden mussten. Nichtübereinstimmungen stehen zudem – so die Autoren – für die gleichzeitige Aktivierung verschiedener Problemlösestrategien bzw. aus der Analyse der Gestik lässt sich eine zukünftige Strategie ableiten, die sich in der Sprache noch nicht zeigt. Gestik stellt eine sinnvolle Ergänzung zur Analyse verbaler Äußerungen dar, da sich in der Gestik somit implizite kognitive Strategien widerspiegeln, die den Problemlösern teils nicht bewusst sind (Garber, Alibali & Goldin-Meadow, 1998; Goldin-Meadow, 1999): „Gesture, it turns out, has access to speakers’ implicit thoughts—thoughts that speakers cannot articulate and may not even know they have“ (Garber & Goldin-Meadow, 2002, S. 819).

Der Erhebungs- und Auswerteprozess von Ausdrucksdaten ist vergleichsweise unkompliziert. Häufig werden Problemlösevorgänge dazu durch Videotechnik dokumentiert und verbal und gestisch beschriebene Lösungswege anschließend kodiert und verglichen. In der Untersuchung von Garber und Goldin-Meadow (2002) werden zunächst ein optimaler und ein nicht-optimaler Lösungsweg definiert. Anschließend wird überprüft, welchen Weg die

Versuchsperson laut ihrer verbalen Beschreibung wählt und inwiefern dieser mit der gestischen Beschreibung übereinstimmt. Dabei werden verschiedene „mismatches“ unterschieden: (1) Die Gestik zeigt einen anderen Weg an als die sprachliche Äußerung (z.B. Beschreibung, dass die letzte Scheibe vom ersten auf den dritten Stab gelegt wurde mit gleichzeitiger Zeigebewegung vom zweiten auf den dritten Stab), (2) die Gestik geht über den verbal beschriebenen Weg hinaus, oder (3) sprachlich wird ein allgemeiner Lösungsweg beschrieben, während durch Gestik eine detailliertere Beschreibung geliefert wird.

Neben dieser videobasierten Kodiermethode sind auch genauere Analyseverfahren denkbar. So können zur Messung von Körperbewegungen z.B. infrarotbasierte 3D-Erfassungssysteme zum Einsatz kommen, wie das System Optotrak (Northern Digital Inc., Ontario, Canada). Bei diesem System werden bspw. an der Hand oder an den Fingern Infrarotleuchtdiodenmarker (LED-Marker) befestigt, deren reflektierter Infrarotlichtstrahl von zwei oder drei Videokameras mit einer sehr hohen räumlichen Auflösung (etwa 0.1 mm) aufgenommen wird. Über die Disparität der aufgenommenen Bilder kann die Position des interessierenden Körperteils im Raum errechnet werden.

### 2.2.3. Psychophysiologische Messmethoden

Auch wenn es auf den ersten Blick merkwürdig erscheinen mag, psychophysiologische Messmethoden zur Erfassung von Denk- und Problemlöseprozessen anführen zu wollen, wird bei genauerer Betrachtung deutlich, dass sowohl bildgebende Verfahren (strukturelle wie funktionelle) als auch pupillometrische Daten eine wichtige zusätzliche Erkenntnisquelle darstellen.

#### **Bildgebende Verfahren**

Nach Craik und Lockhart (1972) erfolgt die Informationsverarbeitung auf hierarchischen Ebenen, den sogenannten levels of processing (LOP), die sich von der perzeptuellen bis zur höheren kognitiven Verarbeitung erstrecken. Diesem Modell entsprechend erfolgt zunächst eine Dekodierung der Oberflächen- oder perzeptuellen Merkmale eines Objekts, bevor eine weitergehende semantische Kategorisierung stattfindet. In neueren Untersuchungen wird noch eine weitere, metakognitive (kreative bzw. selbstreferentielle) Ebene unterschieden. Diese Informationsverarbeitungsprozesse bauen, entsprechend der zunehmenden Verarbeitungstiefe, hierarchisch aufeinander auf. Kognitionspsychologischen Untersuchungen, die sich bildgebender Verfahren bedienen, liegt die Annahme zugrunde, dass sich bestimmte Ebenen

der Informationsverarbeitung mit unterschiedlichen Hirnarealen in Verbindung bringen lassen. Mit modernen bildgebenden Verfahren (Kohärenzverfahren im EEG, PET, funktionelle MRT) konnte inzwischen gezeigt werden, dass unterschiedliche Ebenen der Informationsverarbeitung neurophysiologische Korrelate aufweisen, die sich verschiedenen kognitiven Prozessen zuordnen lassen.

Die Psychophysiologie befasst sich unter anderem mit zentralen und peripheren physiologischen Veränderungen, die auf eine Änderung des psychischen Zustands hin auftreten und stellt somit die Verbindung zwischen Anatomie, Physiologie und dem Verhalten des Organismus in einer physikalischen und sozialen Umwelt dar. Von besonderer Relevanz für den Bereich des Denkens ist die kognitive Psychophysiologie, die sich mit der Beziehung zwischen Elementen der menschlichen Informationsverarbeitung und physiologischen Ereignissen beschäftigt (Cacioppo, Tassinary & Berntson, 2000). Ob es aber tatsächlich möglich ist, psychologische Bedeutung auf physiologische Ereignisse bzw. ihr neuronales Korrelat zurückzuführen, hängt (a) von der Qualität des experimentellen Designs, (b) den psychometrischen Eigenschaften der Messmethoden und (c) der Angemessenheit der Datenanalyse und Interpretation ab (Problem der psychophysiologischen Inferenz). Diese Fragen werden am Ende dieses Abschnitts kritisch diskutiert. Bevor auf einzelne Befunde eingegangen wird, sollen zunächst die wichtigsten psychophysiologischen Verfahren kurz beschrieben werden. Dabei wird insbesondere auf bildgebende Verfahren und solche Methoden eingegangen, die in der kognitiven Psychophysiologie neu sind. Für detaillierte Informationen zu den Verfahren sei auf Cacioppo et al. (2000), Moonen und Bandettini (1999) sowie Toga und Maziotta (2002) verwiesen.

Psychophysiologische Messmethoden generell sind noninvasive Registriertechniken, mit denen biophysikalische Signale erfasst und mit psychologischen Konzepten in Beziehung gesetzt werden können (Rösler, 1996). Biosignale, die den Aktivitätsverlauf einzelner Organsysteme über die Zeit erfassen, können z.B. als Spannungsdifferenzen zwischen zwei Elektroden abgegriffen und verstärkt werden, wie beim Elektroenzephalogramm (EEG), Elektromyogramm (EMG), Elektrokardiogramm (EKG), oder beim Elektrookulogramm (EOG, siehe Abschnitt 2.2.2.2). In Bezug auf die Untersuchung kognitiver Prozesse sind insbesondere ereigniskorrelierte Aktivitätsänderungen, die einem Ereignis folgen oder ihm vorausgehen, von Bedeutung. Das Lateralisierte Bereitschaftspotential (LRP), das neuerdings häufig in Verbindung mit Reaktionszeitmessungen eingesetzt wird (siehe Abschnitt 2.2.2.1), wird aus dem Bereitschaftspotential gewonnen, einer langsamen Negativierung vor

selbstinitiierten oder durch spezifische Reize ausgelösten Handlungen. Während die genannten Methoden zur Erfassung von Biosignalen eine sehr hohe zeitliche Auflösung haben, ist die räumliche Auflösung eher niedrig. Abbildung RESOLUT zeigt einen Überblick über die Auflösung der relevanten Verfahren.

---

Hier bitte Abbildung *RESOLUT* einfügen

---

Bildgebende Verfahren, die zwar eine schlechtere zeitliche, dafür aber eine höhere räumliche Auflösung haben, versuchen hauptsächlich die neuroanatomischen Strukturen zu beschreiben, die durch bestimmte mentale Operationen aktiviert werden. Bei diesen Methoden steht also mehr die neuroanatomische Lokalisation von mentalen Leistungen im Vordergrund, während es bei elektrophysiologischen Methoden eher um die Abbildung des zeitlichen Verlaufs eines Prozesses geht. Um das Verständnis für Handlungsvorbereitung zu verbessern, erscheint eine Integration beider Herangehensweisen sinnvoll (siehe unten). Bildgebende Verfahren können aber nicht nur Hinweise auf die Hirnareale geben, die durch eine kognitive oder behaviorale Aufgabe besonders aktiviert werden, sondern lassen auch Rückschluss auf kognitive Prozesse und die Funktion der spezifischen Hirnstruktur zu: „Combined or integrated effectively, cognitive psychology, functional neuroimaging, and behavioral neurology can map elementary information processing operations onto specific brain areas and map sets of those operations/areas onto system-level models of specific psychological tasks“ (Parsons, 2001, S. 156).

Bildgebende Verfahren oder Brain-Imaging-Techniken lassen sich in strukturelle und funktionelle Methoden unterteilen. Für die Untersuchung von Fragestellungen aus dem Bereich der kognitiven Psychologie sind vor allem funktionelle Verfahren von Bedeutung. Zu den funktionellen Methoden gehören die Einzelphotonen-Emissionscomputertomographie (SPECT), die Positronen-Emissions-Tomographie (PET) und die funktionelle Magnetresonanztomographie (fMRT).

#### Positronen-Emissions-Tomographie (PET )und funktionelle Magnetresonanztomographie (fMRT)

Die im Kontext der kognitiven Psychologie am häufigsten verwendeten Verfahren sind die PET-Technik und die fMRT. Ziel beider Verfahren ist es, gemessene Änderungen der funktionellen (und damit neuronalen) Aktivierung des Gehirns (PET: z.B. Glukose-Metabolismusrate, fMRT: z.B. Sauerstoffgehalt des Blutes, sog. „blood oxygenetic level“) bestimmten mentalen Vorgängen zuordnen zu können. Bei der fMRT wird angenommen, dass

die Zunahme der neuronalen Aktivierung zu einer Zunahme des Sauerstoffbedarfs führt. Dieser Sauerstoffbedarf wird durch die Erhöhung der Zufuhr sauerstoffreichen Bluts gedeckt, bzw. überkompensiert (sog. neurovaskuläre Kopplung). Die kompensatorische Anflutung einer Region mit oxygenisiertem Hämoglobin führt zu einer Homogenisierung des Magnetfeldes, das durch die Gabe eines elektromagnetischen Impulses (derzeit üblich ca. 1,5-7 Teslar) im Kopf erzeugt wird. Wenn das elektromagnetische Feld abgeschaltet wird, wird Energie in Form elektromagnetischer Wellen abgegeben. Die Stärke (Amplitude) des abgegebenen Signals hängt unmittelbar mit der Protonendichte an diesem Ort (und somit mit der Menge an oxygenisiertem Hämoglobin) zusammen (für eine detaillierte Darstellung siehe Raichle, 2001; Reiman, Lane, Van Petten & Bandettini, 2000; Springer, Patlak, Pályka & Huang, 2000). Die fMRT stellt ein räumlich besonders hoch auflösendes Verfahren dar. Ein Problem ist jedoch die zeitliche Auflösung: Da das Signal verzögert ist, können schnelle Veränderungen der neuronalen Aktivität nur bedingt sichtbar gemacht werden.

Mit der fMRT sind gegenüber der PET-Technik zahlreiche Vorteile, aber auch Nachteile verbunden. Während der Proband bei der PET durch die Gabe einer radioaktiven Markersubstanz einem gesundheitlichen Risiko ausgesetzt ist, wird die fMRT bisher als völlig ungefährlich erachtet. Im Gegensatz zur PET erlaubt die fMRT außerdem ereignisbezogene und wiederholte Messungen an einer Versuchsperson. Ein Nachteil ist jedoch die relativ große Anzahl an Artefakten (z.B. Bewegungsartefakte durch eigene oder physiologische Bewegung oder Suszeptibilitätsartefakte durch Störung der Gradientenfelder, hervorgerufen u.a. durch leitfähige Materialien am oder im Körper des Pb), die bei der fMRT zu Datenverlusten führen kann. Schließlich ist fMRT ein lautes Verfahren, PET dagegen nicht. Bei Sprachforschung ist PET meist die zu bevorzugende Methode.

Die Analyse und Interpretation der im Scanner erhobenen Daten ist aufwändig und erfolgt in mehreren Schritten: Zunächst müssen die Bilder um Bewegungsartefakte korrigiert werden (sog. Realignment). Wenn Daten über mehrere Personen aggregiert werden sollen, muss zudem eine Normalisierung erfolgen, bei der die Bilder an ein Referenz-Template derselben Modalität angepasst werden, um interindividuelle Unterschiede der Hirnmorphologie auszugleichen. Schließlich werden die Bilder in der Regel zur Verbesserung des Signal/Rauschen-Verhältnisses geglättet (sog. Smoothing). Einen Überblick über diese Vorverarbeitung (Preprocessing) der Bilder liefern z.B. Frackowiak et al. (1997). Bei der statistischen Analyse, die als „region of interest“-Analyse erfolgen kann, wenn eine a priori Hypothese über den Ort der Aktivierung besteht, stellt die Hypothesenprüfung ein Problem

dar, wenn voxelweise auf Unterschiede in der Aktivierung zwischen einer Experimental- und einer Kontrollbedingung getestet wird. Um eine Inflation des Alpha-Fehlers zu vermeiden, also der Wahrscheinlichkeit, auch nur einen falsch positiven Befund zu erhalten, müssen die Daten etwa mittels Bonferroni-Korrektur, Random-Field-Theorie oder False-Discovery-Rate korrigiert werden (vgl. Friston, Holmes, Poline, Price & Frith, 1996; Genovese, Lazar & Nichols, 2002).

Auch die Auswertung von elektrophysiologischen Daten ist mit viel Aufwand verbunden. Die Extraktion ereignisrelevanter Signale ist aufgrund ihrer kleinen Amplitude und der sie begleitenden Artefakte methodisch schwierig (vgl. Rösler, 1996). Die extrahierten und artefakt-bereinigten Signale müssen auf nicht-redundante Kennwerte (z.B. Latenzen und Amplituden) reduziert und auf eine Baseline (pre-stimulus baseline) oder einen Extremwert (Peak-to-Peak Messung) relativiert werden.

#### Transkranielle Magnetstimulation (TMS)

Die transkranielle Magnetstimulation (TMS), die 1985 erstmals zur Untersuchung des motorischen Kortex eingesetzt wurde (Barker, Jalinous & Freeston, 1985), ist eigentlich keine Meßmethode, sondern eine Methode zur Stimulation des Gehirns durch die Schädeldecke mittels einer magnetischen Spule, die über dem ausgewählten Hirnareal angelegt wird. Ein ca. 1 msek dauernder TMS-Impuls unterbricht durch die Auslösung eines inhibitorischen Signals die normale Aktivität in der Zielregion und u.U. in der entsprechenden kontralateralen Region für etwa 30-250 msek und erlaubt die Messung der Beeinträchtigung bei motorischen, perzeptuellen oder kognitiven Aufgaben (für einen Überblick siehe Jahanshahi & Rothwell, 2000; Walsh, 1998). Da der Effekt der Stimulation für den Probanden kaum spürbar und zudem vollständig reversibel ist, können auf diese Weise „virtuelle Patienten“ erzeugt und die Effekte von Hirnläsionen simuliert werden. Nachdem TMS zunächst zur Untersuchung motorischer Areale und des primären visuellen Kortex verwendet wurde, ist die Anwendung inzwischen auf andere Areale ausgedehnt worden. Da die örtliche Genauigkeit (ca. 0.5 cm) bei tieferen Zellschichten – ähnlich wie beim EEG und MEG – schnell kleiner wird, ist die Anwendung jedoch eingeschränkt. Ein wesentliches Problem bei der Anwendung dieser Methode besteht aber darin, dass bei TMS immer nur in einer Zielregion stimuliert werden kann, während bei anderen Verfahren (z.B. PET, fMRT) die Aktivität des gesamten Kortex beobachtet werden kann. Für exploratorische Studien wird TMS aufgrund des erheblich größeren Aufwandes praktisch nicht eingesetzt. Im Gegensatz zu anderen Verfahren ist die mit TMS gewonnene Aussage dagegen nicht nur eine korrelative, sondern bei festgestellter

Beeinträchtigung der Leistung des Probanden kann davon ausgegangen werden, dass die Aktivität in dem entsprechenden Areal tatsächlich kritisch (i.S. von notwendig) für die Ausführung einer Aufgabe ist. Für die Vorgehensweise ist daher zu beachten: Zunächst sollten die für eine Aufgabe in Frage kommenden Hirnareale durch andere Methoden bestimmt werden. TMS kann dann klären, ob das Gebiet tatsächlich kritisch ist und gegebenenfalls für welchen Aspekt der Aufgabe und zu welchem Zeitpunkt der Aufgabenlösung. Aufgrund der hohen zeitlichen Auflösung (ca. 50 msek) kann so der Verlauf der Aufgabenlösung relativ genau nachgezeichnet werden. Unabdingbar sind dafür die genaue Konfiguration und Platzierung der Magnetspule, über welche die Stimulation erfolgt. Noch nicht gänzlich geklärt sind die mit TMS verbundenen möglichen beeinträchtigenden Auswirkungen auf den Probanden (Pascual-Leone et al., 1993), zumindest aber kann die Anwendung sensorische Nebeneffekte (wie z.B. den Lidschlussreflex, Gesichtszucken) produzieren, die nicht nur unangenehm sind, sondern auch mit der Messung interferieren können (Cowey & Walsh, 2001; Jahanshahi & Rothwell, 2000).

### Empirische Befunde

Studien mit bildgebenden Verfahren haben sich bisher im Wesentlichen mit relativ einfachen Denkaufgaben beschäftigt. Parsons (2001) fasst in einem Übersichtsartikel bspw. Befunde zur mentalen Rotationsaufgabe zusammen, die mittels fMRT gewonnen wurden. Neben der Identifizierung der während der Aufgabenbearbeitung aktiven Hirnareale (v.a. Cerebellum) ist ein häufiger Befund die deutliche Übereinstimmung der Länge der Reaktionszeit mit der Dauer der sog. BOLD-Antwort (BOLD = „blood oxygenetic leveling response“ als Maß für die neuronale Aktivierung).

Das wohl am häufigsten verwendete Paradigma ist der „Turm von Hanoi“ (TVH) bzw. der „Turm von London“ (TVL, siehe Abschnitt 2.1.1.3). Beide Verfahren sind bei gesunden Probanden (Fincham, Carter, van Veen, Stenger & Anderson, 2002; Rowe, Owen, Johnsrude & Passingham, 2001) und bei Patienten, meist mit Läsionen im Frontalhirn (Stuss & Alexander, 2000; van den Heuvel et al., 2003), eingesetzt worden. Im Ergebnis geben diese Studien Aufschluss über die mit der Planungs- bzw. Problemlöseaufgabe verbundenen Hirnareale. Im Fall des TVH/TVL sind das der dorsolaterale präfrontale Kortex, das prämotorische Areal, Teile des parietalen Kortex und das Cerebellum. Der dorsolaterale präfrontale Kortex scheint insbesondere mit der Generierung, Auswahl und Erinnerung an mentale Bewegungen der Scheiben im TVH befasst zu sein. Aktivität in ähnlichen Arealen, insbesondere den superioren Frontallappen und dem parietalen Kortex, wurde in einer fMRI-

Studie zum Lösen von Schachproblemen gefunden (Atherton, Zhuang, Bart, Hu & He, 2003). Newman et al. (2003) finden bei der Bearbeitung des TVL Unterschiede im Aktivitätsmuster im parietalen Kortex zwischen der linken und rechten Hemisphäre und ziehen die Schlussfolgerung, dass Areale im rechten präfrontalen Kortex eher an der Generierung eines Plans, Areale im linken präfrontalen Kortex dagegen bei Ausführung eines Plans aktiv sind. Die im präfrontalen und superior parietalen Kortex gemessene Aktivität wurde mit zunehmender Aufgabenschwierigkeit stärker. Entsprechend ist der rechte superiore Parietallappen mit Aufmerksamkeitsprozessen verbunden, während das entsprechende Areal links eher an visuell-räumlichen Prozessen beteiligt zu sein scheint.

Zu den untersuchten komplexeren Aufgaben gehören z.B. Entscheidungsaufgaben und das Lösen von Alltagsproblemen (Channon & Crawford, 1999; Dickson et al., 2000). Volz et al. (2003) verwenden Aufgaben, in denen Probanden häufigkeitsbasierte Vorhersagen unter Unsicherheit machen müssen. Die bei diesen Aufgaben aktiven Hirnareale liegen im posterioren frontomedianen Bereich und sind bekannt als Korrelate des Hypothesentestens.

Mittels ereigniskorrelierter Messungen wurden ebenfalls einfache Aufgaben untersucht, z.B. zur mentalen Arithmetik (Roland & Friberg, 1985). In einer Studie zu komplexen Additionsaufgaben, in denen Probanden angeben mussten, ob ein Ergebnis über oder unter 100 liegt, fanden El Yagoubi et al. (2003) mittels Ereigniskorrelierter Potenziale (EKP), dass je nach Abweichung des Ergebnisses von 100 unterschiedliche Strategien eingesetzt werden: Weicht das Ergebnis stark von 100 ab, wird lediglich geschätzt, ist das Ergebnis nahe an 100, wird genau nachgerechnet. Diese unterschiedlichen Strategien weisen möglicherweise verschiedene neurophysiologische Korrelate auf, zumindest unterscheiden sich die EKP-Signale je nach gewählter Strategie im Zeitraum zwischen 250 und 600 msek nach Aufgabenonset. Mittels des EKP-Signals kann somit eine Hypothese über den möglichen zeitlichen Ablauf der zugrunde liegenden kognitiven Prozesse – visuelle Enkodierung der Aufgabe und Strategiewahl nach 250 msek, Strategiewahl bis 600 msek, Entscheidung und motorische Reaktion (zwischen 687 msek bei Rechnung und 872 msek bei Schätzung) – formuliert werden.

Die transkranielle Magnetstimulation (TMS) ist bereits vermehrt im Bereich kognitiver Aufgaben eingesetzt worden (für Übersichtsarbeiten siehe Jahanshahi & Rothwell, 2000; Robertson, Theoret & Pascual-Leone, 2003). In den meisten Untersuchungen wurden allerdings – ebenso wie bei den bildgebenden Verfahren – relativ simple Paradigmen verwendet. Zu den am häufigsten untersuchten Bereichen gehören Aufmerksamkeitsprozesse

(z.B. bei der visuellen Suche) und Sprachverarbeitungsprozesse (für eine Übersicht siehe Jahanshahi & Rothwell, 2000; Stewart, Ellison, Walsh & Cowey, 2001; Volz et al., 2003).

### Fazit

Die exemplarisch dargestellten Studien insbesondere zu PET und fMRT zeigen, dass es sich hier um ein schnell expandierendes Forschungsfeld handelt. Gleichwohl sind die Befunde zu kognitionspsychologischen Fragestellungen noch dürftig (Ausnahme: Untersuchungen zur Einsicht von Luo & Niki, 2003). Die Gründe hierfür liegen auf der Hand: Studien mit bildgebenden Verfahren fordern ein Höchstmaß an Kontrolle über die Stimulussituation und lassen daher relativ wenig Spielraum, was die Auswahl komplexerer und dynamischer Stimuli anbelangt. Um zu validen und reliablen Ergebnissen zu kommen, müssen Studien mit bildgebenden Verfahren einer Reihe von Anforderungen genügen, die hier nur kurz skizziert werden sollen:

- (1) Es muss sichergestellt werden, dass die beobachtete neurophysiologische Veränderung tatsächlich auf den Stimulus zurückgeführt werden kann. Dazu ist der Vergleich der Daten mit einer Baseline (Messung vor Darbietung des Stimulus) nötig. Bei kognitiven Aufgaben stellt die Aufgabenschwierigkeit in diesem Kontext ein besonderes Problem dar. Ist die Aufgabe zu leicht, kommt es möglicherweise zu keiner Veränderung. Ist die Aufgabe hingegen zu schwer, ist nicht festzustellen, ob die physiologische Veränderung auf den Inhalt der Aufgabe oder auf „Nebenwirkungen“ der Aufgabenschwierigkeit (z.B. Frustration, erhöhte Aufmerksamkeit, Fehlermonitoring) zurückgehen.
- (2) Für den interindividuellen Vergleich wird die Ähnlichkeit in der funktionellen Anatomie zwischen den Versuchspersonen vorausgesetzt. Die individuelle Variabilität ist hier aber häufig sehr hoch.
- (3) Es ist ein Vergleich zwischen einer Experimental- und einer Kontrollgruppe notwendig, wobei sich die Kontrollbedingung nur in der relevanten Hirnregion von der Experimentalbedingung unterscheiden sollte. Diese Anforderung ist praktisch nicht haltbar.
- (4) Aufgrund der Dauer der BOLD-Antwort (einige Sekunden) ist man in der Auswahl der Aufgaben eingeschränkt. Die BOLD-Antwort ist zudem altersabhängig: Mit zunehmendem Lebensalter sinkt das Signal/Rauschen-Verhältnis.

- (5) Schließlich sind die für den Probanden auftretenden Unannehmlichkeiten (fMRT: Lärm, Enge, unbequeme Lage) zu erwähnen, die nicht zuletzt auch die Qualität der Daten beeinflussen können, weshalb ein Vergleich der vom Probanden erzielten Ergebnisse mit denen einer Kontrollgruppe, welche die entsprechende Aufgabe außerhalb des Scanners absolviert hat, nötig ist. Besonders bei Untersuchungen mittels fMRT fallen ca. 5-10% der Stichprobe durch Probleme wie Klaustrophobie weg.
- (6) Zudem gibt es eine Reihe von Artefakten (durch Bewegung, Atmung, Blutgefäße), die entweder zu Datenverlust führen können oder eine anschließende Filterung der Daten notwendig machen.
- (7) Zu guter Letzt sind die großen Datenmengen zu nennen, die eine simple Anwendung bildgebender Verfahren ohne konkrete Hypothese nicht rechtfertigen: „[...] experimental studies should attempt to address important questions; they should be designed to test specific hypotheses and generate additional hypotheses that can be tested in the future; they should capitalize on the capabilities of this technology and interpret findings in the context of its limitations” (Reiman et al., 2000, S. 93).

Kritisch ist zu bemerken, dass in vielen bisherigen Studien ein ausschließlicher Fokus auf der Lokalisation kognitiver Funktionen liegt. Erforderlich wäre dagegen die Kombination verschiedener Verfahren zur Untersuchung dynamischer Aspekte von kognitiven Prozessen (vgl. Uttal, 2001). Bildgebende Verfahren stellen immer noch ein großes, teilweise noch unausgeschöpftes Potenzial dar. Neuere Anwendungen, z.B. die Möglichkeit, während einer Messung „online“ Feedback über die Aktivierung an den Probanden rückzumelden (vgl. deCharms et al., 2004) oder die Kombination von EEG- und fMRT-Untersuchungen, bieten interessante Möglichkeiten für kognitionspsychologische Studien.

### **Pupillometrie**

Als Pupille bezeichnet man die in der Regel schwarz erscheinende Öffnung in der pigmentierten und damit weitgehend undurchsichtigen Regenbogenhaut (Iris) des Augapfels. Bei Lichteinfall verengt sich die Pupille auf minimal 1,5 mm, um sich bei Dunkelheit wieder (auf maximal 8 mm) zu erweitern. Unter üblichen Beleuchtungsbedingungen beträgt der Spielraum zwischen 2 und 6 mm Durchmesser. Die parasymphatisch und sympathisch gesteuerte Pupillengröße ist also stark von externen Einflüssen (vor allem Lichteinfall)

abhängig (Alexandridis, Leendertz & Barbur, 1991; Loewenstein & Loewenfeld, 1969); die Pupille verkleinert sich aber bspw. auch bei Nah-Akkommodation und Konvergenz der Pupillen. Daneben spielen auch Reizeigenschaften (Variiertheit des Reizmusters, Schwierigkeit der Aufgabe) eine Rolle. Eine ausführliche Zusammenschau der in der Literatur beschriebenen Einflüsse auf den Durchmesser der Pupille findet sich bei Rößger (1997).

Die Reaktion der Pupille, insbesondere ihre Größenzunahme (Dilatation), wird in vielen denkpsychologischen Studien als abhängige Variable betrachtet. In den letzten 40 Jahren hat sich für die Untersuchung der Pupillengröße im Kontext kognitiver Aufgaben der Begriff der „kognitiven Pupillometrie“ eingebürgert (Beatty & Lucero-Wagoner, 2000). Im Mittelpunkt dieses Abschnitts stehen folglich kognitionspsychologische Studien, in denen der Pupillendurchmesser als abhängiges Maß verwendet wurde. Zunächst soll jedoch ein kurzer Überblick über Messmethoden des Pupillendurchmessers gegeben werden.

#### Messmethoden

Zur Bestimmung der Pupillengröße wird üblicherweise ein Pupillendilatationsindex aus dem Ausgangsdurchmesser der Pupille und der (nach einer Latenzzeit erreichten) maximalen Dilatation (sog. peak dilation) gebildet. In jedem interessierenden Auswertungsintervall können mittlere und maximale Dilatation und die Latenzzeit bis zum Erreichen der maximalen Dilatation bestimmt werden. Eine Übersicht über Messmethoden für die Erfassung der Pupillengröße geben Alexandridis, Leendertz und Barbur (1991) sowie Beatty und Lucero-Wagoner (2000). Mit fast jedem hochauflösenden, infrarot-sensitiven, videobasierten Blickbewegungsmessgerät lässt sich auch die Pupillengröße als zusätzliches Maß zur Position der Pupille erfassen (siehe auch Nguyen & Stark, 1993).

Im Mittelpunkt der psychologischen Auseinandersetzung mit der Pupillendilatation steht die Vermutung, dass die Größe der Pupille ein valides Maß für die kognitive Beanspruchung (cognitive load) während einer Aufgabe darstellt und die Annahme, dass Verhaltensdaten somit durch physiologische Maße stützbar sind. Die Messung der lichtunabhängigen Pupillendilatation erfolgte bereits in Untersuchungen im 18. Jh. (Beatty & Lucero-Wagoner, 2000; Loewenfeld, 1958). Dass die Pupillengröße ein Maß kognitiver Aktivität sein könnte, wurde erstmals von Hess und Polt (1960) postuliert. Eine Zusammenfassung der frühen Untersuchungen von Hess und Kollegen findet sich bei Janisse (1977).

### Empirische Befunde

Eine Vielzahl von neueren Untersuchungen deutet darauf hin, dass die Pupillengröße als Maß mentaler Beanspruchung betrachtet werden kann. Studien, die den Zusammenhang von Pupillengröße und kognitiven Prozessen untersuchen, postulieren zumeist, dass der Pupillendurchmesser ein sensitives Maß der sich während einer Aufgabe ständig verändernden Arbeitsgedächtnisbelastung bzw. kognitiven Beanspruchung allgemein und somit auch ein Maß für die Aufgabenschwierigkeit darstellt. Zudem wird häufig ein Zusammenhang zwischen dem Pupillendurchmesser und den im Kontext der Aufgabe auftretenden (bzw. durch diese hervorgerufenen) Emotionen gesehen. In den folgenden Abschnitten werden exemplarisch Studien zu beiden Bereichen vorgestellt. Nicht näher eingegangen werden soll auf die Studien zum Zusammenhang zwischen dem Pupillendurchmesser und der Aufmerksamkeit (Beatty, 1989; Hoeks & Levelt, 1993), der Sprache (Schluroff et al., 1986), Psychophysik (Kahneman & Beatty, 1967) und Motorik (z.B. Vorbereitung einer motorischen Reaktion, Richer & Beatty, 1985).

Bezüglich der Kurzzeit- und Arbeitsgedächtnisleistung kommen verschiedene Studien zu dem Ergebnis, dass eine Vergrößerung des Pupillendurchmessers mit der im KZG gespeicherten Anzahl von Einheiten, mit einer höheren Arbeitsgedächtnisbelastung, mit der steigenden Schwierigkeit einer Aufgabe und während der Suche im LZG linear ansteigt (Ahern & Beatty, 1978; Beatty & Kahneman, 1966; Beatty & Lucero-Wagoner, 1978; Hess & Polt, 1964). In einer Untersuchung zum Pupillendurchmesser bei Kopfrechenaufgaben fand Bradshaw (1968) einen größeren Pupillendurchmesser bei schwierigen in Vergleich zu leichten Aufgaben. Der größte Pupillendurchmesser wurde im Moment der Aufgabenlösung gemessen; nach der Lösung verkleinerte sich der Pupillendurchmesser wieder. Zu ähnlichen Befunden kommen Hess und Polt (1960) bei einfachen Multiplikationsaufgaben. Die Autoren schlussfolgern, dass mentale Aktivität mit der Problemschwierigkeit korreliert ist und dass die Pupillengröße entsprechend mit der Problemschwierigkeit zunimmt. Untersuchungen zur Überbeanspruchung der Verarbeitungskapazität (bei beschleunigter Darbietung der Aufgabe) zeigten jedoch, dass nach Beenden der Aufgabe die Pupille nicht unbedingt in Reaktion auf die Aufgabe kleiner wird und zumindest nicht unter die Ausgangsgröße sinkt (Juris & Velden, 1977). Es gibt jedoch interindividuelle Unterschiede bezüglich des Effekts der Aufgabenschwierigkeit auf den Pupillendurchmesser. So weisen Personen mit höheren Werten bei Mathematik- und Sprachverständnisaufgaben des „Scholastic Aptitude Tests“ mit

steigendem Testscore weniger starke Pupillenvergrößerungen bei kognitiver Beanspruchung auf (Ahern & Beatty, 1978).

Die Pupillengröße wird jedoch nicht nur als Maß kognitiver Beanspruchung verwendet. Pupillendilatation lässt sich auch in Reaktion auf perzeptuelle, motorische und emotionale Beanspruchung feststellen (Arima & Wilson, 1972; Beatty, 1982; Goldwater, 1972; Janisse, 1977; Johnson, 1971; White & Maltzman, 1978). Einige Autoren (z.B. Johnson, 1971) nehmen an, dass emotionale Reaktionen, wie z.B. Angst vor einer schwierigen Aufgabe, die Vergrößerung des Pupillendurchmessers hervorrufen, da Probanden in Untersuchungen der oben genannten Art meistens von vornherein über die Schwierigkeit der folgenden Aufgabe informiert waren. Kahneman (1973) argumentiert gegen diese Vermutung, indem er darauf hinweist, dass der Pupillendurchmesser während der Durchführung einer schwierigen Aufgabe größer ist als davor und danach, was gegen eine Pupillenvergrößerung durch Angst spricht.

### Fazit

Obwohl zahlreiche Untersuchungen belegen, dass die Pupillengröße ein Indikator für die mentale Beanspruchung darstellen kann, ist die praktische Anwendung dieses Augenparameters nicht unproblematisch: (a) Größenveränderungen der Pupille haben eine Latenzzeit von bis zu 1000 msek (Loewenfeld, 1958), (b) kognitiv bedingte Größenveränderungen der Pupille können nur unter kontrollierten Versuchsbedingungen (z.B. Konstanthaltung des Umgebungslichts) von reflexhaften, optisch bedingten Veränderungen der Pupillengröße unterschieden werden, (c) Größenveränderungen sind aufgrund ihres geringen Ausmaßes (ca. 0.5 mm) nur mit speziellen Verfahren messbar, (d) Gesundheitszustand, Alter und kognitive Voraussetzungen des Probanden können Einfluss auf die Pupillengröße haben bzw. den Effekt der Aufgabenschwierigkeit auf die Pupillengröße moderieren, und schließlich (e) stellen Pupillenveränderungen immer nur ein Korrelat kognitiver Prozesse dar. Es handelt sich also um einen nicht kausal verknüpften, indirekten „marker“ (Beatty & Lucero-Wagoner, 2000; Cacioppo & Tassinari, 1990; Just & Carpenter, 1993). Da für meisten der genannten Probleme Lösungen existieren, halten wir die kontrollierte Verwendung pupillometrischer Daten für sinnvoll.

## ***2.3. Auswertungsverfahren***

Geht es um Methoden zur Auswertung erhobener Daten, stehen der Denk- und Problemlöseforschung verschiedene Verfahren zur Verfügung. Die (a) Kognitive

Modellierung erlaubt es innerhalb der jeweiligen Formalismen (Produktionssystem, konnektistischer Ansatz, multinomiales Modell), bestimmte Prozessannahmen in lauffähige Programme zu überführen und zu Datenschätzungen zu kommen, die mit realen Daten abgeglichen werden können. Diesem Vorgehen verwandt ist die Technik (b) Synthetischer Versuchspersonen, deren Verhalten ebenfalls mit realen Daten verglichen werden kann. Die Technik der (c) Markov-Analyse erlaubt es, aus den Übergangswahrscheinlichkeiten zwischen bestimmten Ereignissen Schlussfolgerungen über Merkmale des dahinter stehenden Systems zu ziehen. Schliesslich kann die (d) Latente Semantische Analyse herangezogen werden, wenn riesige Log-File-Datensätze hinsichtlich ihrer Dimensionalität in reduzierter Form abgebildet werden sollen.

### 2.3.1. Kognitive Modellierung

Kognitive Modellierung stellt den Versuch dar, postulierte kognitive Prozesse von Menschen so genau zu beschreiben, dass sie auf einer Maschine nachvollzogen werden können. Dazu ist eine kognitive Architektur in Form von Strukturen und Operationen zu spezifizieren, mit denen die Informationen bereitgestellt werden können, die zur Bearbeitung eines Problems sowie zur Beschreibung von Zwischenzuständen benötigt werden. Die kognitive Modellierung soll sowohl vom Ablauf wie vom Ergebnis her die Vorgänge menschlicher Problemlöser widerspiegeln. Hierzu sind eine Reihe von Kriterien entwickelt worden, mit denen die Modellpassung überprüft werden kann (vgl. Opwis & Spada, 1994; Schmid & Kindsmüller, 1996; Wallach, 1998).

Zwei wesentliche Zugangsformen zur kognitiven Modellierung haben sich etabliert: Regelbasierte Modelle, die Kognition durch Symbolsysteme und die Anwendungen zugehöriger Regelsysteme beschreiben, sowie konnektionistische Modelle, die Kognition auf grundlegende Prinzipien neuronaler Prozesse (Aktivationsausbreitung, Vernetzungsmuster etc.) zurückführen. Neben diesen Modellierungsvarianten, mit denen Individualdaten modelliert werden können, lassen sich mit dem Verfahren der multinomialen Modellierung Annahmen über kognitive Prozesse auf Aggregatebene überprüfen.

Die hier gegebene Darstellung von regelbasierten und konnektionistischen Modellen fällt sehr knapp aus, da der Beitrag von Schmid (in diesem Band) ausführlicher auf formale Modelle im Bereich von Denken und Problemlösen eingeht.

## **Produktionssysteme**

Taatgen und Wallach (2002) haben die Frage verfolgt, mit welchen Regeln sich die Steuerung des Szenarios „Sugar Factory“ von Berry und Broadbent (1984) abbilden lässt. Sie verglichen ein einfaches instanzen-basiertes ACT-R-Modell mit einem Modell von Dienes und Fahey (1995) hinsichtlich der Vorhersage empirischer Daten. Da beide Modelle diesbezüglich gleich gut abschnitten, konnten die Autoren insofern einen Erfolg für ihr ACT-R-Modell verbuchen, als dieses mit weniger Modellannahmen auskam und somit als sparsameres Modell bezeichnet werden durfte.

## **Konnektionistische Verfahren**

Leighton und Dawson (2001) haben eine konnektionistische Modellierung der in der Urteilsforschung so beliebten „Wason-Selection-Task“ vorgenommen. Sie plädieren dafür, die Schwierigkeit von Problemen daran festzumachen, wie viel „hidden units“ (zwischen Input- und Output-einheiten liegende Einheiten eines konnektionistischen Modells) zur Repräsentation benötigt werden.

Grundlegende Bewertungen des konnektionistischen Vorgehens findet man etwa bei Broadbent (1985), Levelt (1991), McCloskey (1991) oder Strube (1990). Einfache Hilfsmittel zur konnektionistischen Modellierung mittels der Tabellenkalkulations-Software Microsoft Excel finden sich bei Macho (2002a; 2002b).

## **Multinomiale Modellierung**

Die multinomiale Modellierung stellt eine Modellierungstechnik dar, die einen Mittelweg zwischen einem rein experimentellen Zugang und einem rein theoretischen Zugang versucht. Ziel des Verfahrens ist es, die Messung unbeobachtbarer, latenter, kognitiver Prozesse aufgrund von beobachtbaren Daten zu leisten und dies mit minimalen, aber substantiellen Annahmen über kognitive Zustände (für einen Überblick siehe Batchelder & Riefer, 1999). Das Vorgehen besteht darin, Situationen zu schaffen, in denen verschiedene postulierte kognitive Prozesse zu diskreten, messbaren Zuständen führen, und anschließend die Passung (den Fit) zwischen Modell und Daten festzustellen.

Drei Aussagen charakterisieren das Verfahren: (1) einer diskreten, endlichen Menge postulierter kognitiver Zustände können ganz bestimmte beobachtbare Verhaltensweisen zugeordnet werden; (2) die Zuordnung von Verhalten zu kognitivem Zustand ist nicht immer eindeutig, wohl aber gibt es eine klare Zuordnung von Zustand zu Verhalten; (3) die meisten

Verhaltensweisen können aus unterschiedlichen Zuständen heraus zustande kommen, aber ein angenommener kognitiver Zustand führt immer zu eindeutigem Verhalten.

Ein Beispiel aus Beckmann (1994) soll diese abstrakte Darstellung anhand eines Problemlöse-Kontexts illustrieren. Beckmann interessierte sich für die Wissenserwerbsprozesse beim komplexen Problemlösen seiner Szenarien „Kirschbaum“ und „Maschine“. Seine Vpn mussten im Verlauf des Experiments mehrfach Kausaldiagramme über die vermuteten Relationen zwischen Systemvariablen erstellen. Eine richtige Antwort wie z.B. „positive Relation“ in diesem Wissenstest (= Verhalten) kann auf vier verschiedenen Wegen zustande kommen, wie man aus Abbildung MUL erkennen kann. Diese Abbildung zeigt den Ereignisbaum, der die Reizkategorie „positive Relation“ mit den entsprechenden Verhaltensäußerungen (dem Antwortverhalten) verknüpft.

---

Hier bitte Abbildung MUL einfügen

---

Wie Abbildung MUL zeigt, steht das Antwortverhalten „positive Relation“ am Ende von drei Pfaden. Die Antwort kommt zum ersten zustande, wenn die Existenz der Relation und das Vorzeichen korrekt erkannt wurden mit Wahrscheinlichkeit  $\underline{E}_{r+} \times \underline{E}_{v+} \times 1$ . Zum zweiten kann die Antwort durch Raten produziert werden mit Wahrscheinlichkeit  $(1 - \underline{E}_{r+}) \times \underline{r} \times \underline{p}$ . Schließlich kommt diese Antwort zustande, wenn die positive Relation erkannt wurde, bei der Vorzeichenerkennung aber richtig geraten wurde; die Wahrscheinlichkeit dafür beträgt  $\underline{E}_{r+} \times (1 - \underline{E}_{v+}) \times \underline{p}$ . Wie man sieht, wird hier das Zustandekommen einer bestimmten Antwort durch eine Reihe von Parametern beschrieben. Das von Beckmann entwickelte Prozessmodell für drei Reizkategorien (positive Relationen, negative Relationen, Nichtrelationen) und fünf Ereigniskategorien („positive Relation“, „negative Relation“, „Relation mit unbekanntem Vorzeichen“, „keine Relation“, „ich weiss nicht“) verfügt somit über  $3 \times 5 = 15$  empirische Häufigkeiten, was bei insgesamt neun zu schätzenden Parametern identifizierbar ist.

Es zeigt sich bei der Modellvorhersage aufgrund der geschätzten Modellparameter nicht nur ein guter Fit zu den Daten, sondern auch der Verlauf einzelner Schätzwerte für die Wissenserwerbsphase liefert interessante Informationen. So steigt etwa die Wahrscheinlichkeit, eine bestehende Relation als solche zu erkennen, im Laufe des Versuchs kontinuierlich an, wohingegen die Wahrscheinlichkeit, eine Nichtrelation zu erkennen, über die gesamte Zeit auf einem sehr niedrigen Niveau verbleibt. Das Erkennen einer Nicht-Relation ist damit deutlich schwieriger als das einer bestehenden Relation. Auch die Rateparameter werden als weitgehend konstant über den gesamten Verlauf geschätzt. Auf

weitere Steigerungsmöglichkeiten der Aussagekraft seines Modells (etwa durch Restriktionen in Form von Parameter-Gleichsetzungen) kann hier nicht ausführlicher eingegangen werden (siehe dazu Beckmann, 1994, S. 158-172).

### Stärken und Schwächen der MM

Die Stärken multinomialer Modellierung liegen in der Separierung verschiedener kognitiver Prozesse und der Bestimmung ihrer jeweiligen Anteile. Diese klare Modellierung wird zudem ergänzt durch eine theoriegeleitete Modellprüfung mittels experimenteller Variationen, bei denen die angenommenen Modellparameter in spezifischer Weise beeinflusst werden sollen (= Validierung der Parameter).

Diesen Stärken stehen auch Schwächen gegenüber. So ist etwa die geforderte Unabhängigkeit einzelner Vpn-Antworten über mehrere aufeinander folgende Durchgänge nicht leicht zu gewährleisten; außerdem führen bei großen Stichproben bereits kleine Modellabweichungen zum Scheitern der Modellgeltungstests. Die „Bäume“ suggerieren darüber hinaus einen sequenziellen Verarbeitungsprozess, der nicht notwendigerweise in der jeweils gezeigten Abfolge existiert (es geht nur um die Kombination, nicht die Sequenz von Prozessen). Schließlich ist es kaum möglich, zu einer Parameterschätzung auf Individuen-Ebene zu kommen – die Schätzungen beziehen sich durchgängig auf Vpn-Aggregate (Ansätze zu individuellen Parameter-Schätzungen finden sich bei Batchelder, 1998).

### 2.3.2. Synthetische Versuchspersonen

Die Idee, mit synthetischen Vpn zu arbeiten, ist bereits in der Arbeit von Dörner (1982) beschrieben. Dort heißt es (S. 65): Das Simulationsprogramm zur Modellierung der Absichtsregulation „erzeugte Abläufe, die man in ihrer Struktur mit echten Abläufen vergleichen kann. Man kann das Programm zur Prognose solcher Abläufe bei Menschen einsetzen“. Die neueren Arbeiten vor dem Hintergrund synthetischer Vpn sind einerseits im „Bauplan für eine Seele“ (Dörner, 1999), andererseits im Projekt „Psi“ der Bamberger Arbeitsgruppe zu finden (Dörner et al., 2002).

Was ist der Grundgedanke dieses Ansatzes? Hat man ein komplexes Annahmefüge über psychische Abläufe, wie es in der Psychologie typisch ist (in der Physik ist – so Dörner et al., 2002, S. 16 – das Fallgesetz eine vergleichsweise einfache Zusammenhangsbehauptung), ist eine Prüfung der zahlreichen Wechselwirkungen, die sich aus dem Zusammenspiel eines großen Bündels von Zusammenhangsbehauptungen ergeben, kaum anders möglich als durch

Simulation. Synthetische Versuchspersonen sind nichts anderes als Bündel von Zusammenhangsbehauptungen, die durch spezielle Parametrisierungen „individualisiert“ werden. So kann z.B. eine Zusammenhangsaussage, in der Auswirkungen im Zusammenhang mit dem Arbeitsgedächtnis formuliert werden, dadurch individualisiert werden, dass in der einen Simulation ein größeres, in der anderen ein kleineres Arbeitsgedächtnis angenommen wird. Auf diese Weise entstehen „Persönlichkeiten“ – Schaub (2001) bezeichnet die Parameter des Informationsverarbeitungssystems in diesem Sinne stringent als Persönlichkeitsfaktoren. Damit wird jede kognitive Modellierung, die Individuenparameter enthält, zu einer synthetischen Versuchsperson.

An dieser Stelle sei eine Bemerkung zum Verhältnis von Allgemeiner und Persönlichkeitspsychologie erlaubt: Mit der Aufnahme von Individuenparametern verlässt man keinesfalls das Gebiet der Allgemeinen Psychologie, wie man vielleicht nach den eben gemachten Ausführungen vermuten könnte. Dass die Allgemeine Psychologie nach allgemeingültigen Gesetzen suche, die Persönlichkeits- und Differentielle Psychologie dagegen die individuellen Unterschiede beachte (so z.B. Amelang & Bartussek, 2001, S. 34-36), ist eine sehr vergrößernde und irreführende Sicht. Viele Gesetze enthalten variable Parameter, mit denen bei Gültigkeit eines allgemeinen Gesetzes den individuellen Unterschieden Rechnung getragen werden kann (siehe dazu auch Erdfelder & Funke, 2004). So gilt etwa Stevens Potenzgesetz der Wahrnehmung in seiner universellen Form auch dann, wenn z.B. zwei Personen unterschiedliche Exponenten für Helligkeitswahrnehmung aufweisen. Auch das Fallgesetz ist durch die jeweilige Gravitationskonstante „individualisiert“ und bleibt trotzdem allgemein gültig.

Fazit. Synthetische Vpn stellen eine komplexe Variante kognitiver Modellierung dar, bei der nicht nur ein bestimmter kognitiver Teilprozess, sondern das Zusammenspiel einer Vielzahl von Modulen nachgebildet wird. Dies stellt einen erheblichen Gewinn gegenüber der Entwicklung von oft sehr speziellen Modellen für ein bestimmtes Phänomen dar. Eine experimentelle Überprüfung solcher Systeme fällt allerdings schwer, da die theoretischen Konstrukte einerseits sehr flüchtig (d.h. sich rasch ändernd) sind, andererseits auch kaum Verankerungen in der Empirie aufweisen. Wie in kaum einem anderen Bereich der Psychologie ist die Gefahr von „Glasperlenspielen“ hier besonders hoch.

### 2.3.3. Markov-Analysen

Markov-Analysen beziehen sich auf die Wahrscheinlichkeiten, mit denen ein System seine Zustände wechselt. Die Markov-Annahme besagt, dass der nächste Zustand des Systems ausschließlich von gegenwärtigen, nicht aber von vergangenen Zuständen abhängt. Der Übergang zwischen sonnigen und regnerischen Tagen kann etwa auf diese Weise als Markov-Kette abgebildet werden. Danach könnte der beste Prädiktor für das morgige Wetter das heutige sein, wenn die Übergänge von Sonne zu Sonne und von Regen zu Regen jeweils hoch wären.

Als Beispiel für eine denkpsychologisch motivierte Markov-Analyse wird hier auf die bei Dörner und Wearing (1995, S. 74f.) beschriebene Untersuchung von Systemverlaufparametern – vor allem die beiden Verhaltensmerkmale „Fragen“ und „Entscheidungen“ – beim Bearbeiten des Szenarios „Lohhausen“ eingegangen. Das Frageverhalten ist insofern informativ, als die Informationsbeschaffung differenziert werden kann hinsichtlich Zustandsfragen („Wie viele Arbeitslose haben wir heute?“), Dependenzfragen („Wovon hängt Arbeitslosigkeit ab?“), Effektanzfragen („Was bewirkt Arbeitslosigkeit?“), Komponentenfragen („Was gehört dazu?“), Subordinationsfragen („Was sind untergeordnete Aspekte?“) oder Superordinationsfragen („Was sind übergeordnete Aspekte?“). Neben der Art der Fragen kann man auch den Ort der Frage bestimmen, d.h. den Punkt, auf den sich die jeweilige Frage richtet. Auch das Entscheidungsverhalten lässt sich nach Ort und Dosierung unterscheiden. Da sowohl Fragen als auch Entscheidungsmerkmale kontextabhängig zu interpretieren sind, wird eine automatische Diagnostik durch Erhebung von Einzelindices erschwert. Aber natürlich kann man versuchen, die Interaktion von Fragen und darauf folgenden Entscheidungen näher zu bestimmen. Bei Dörner und Wearing (1995, S. 74f.) findet sich eine interessante Darstellung der Übergänge zwischen verschiedenen Interaktionsformen. Die in Abbildung MARKOV (a) und (b) dargestellten Zustandsübergänge zwischen fünf verschiedenen Arten der Informationsverarbeitung machen Unterschiede zwischen erfolgreichen und weniger erfolgreichen Versuchspersonen bei „Lohhausen“ deutlich.

---

Hier bitte Abbildung MARKOV einfügen

---

In dieser Abbildung werden die relativen Übergangshäufigkeiten zwischen allgemeiner Orientierung (AO, z.B. „Gibt es ein Theater in der Stadt?“), gerichteter Orientierung (GO, z.B. „Wie hoch ist die Einkommensteuer?“), oder Exploration (z.B. „Was wären die

Konsequenzen einer generellen Steuererhöhung?“), sowie der Information, die der Versuchsleiter gibt, und dem Treffen einer Entscheidung gezeigt. Die Unterschiede zwischen erfolgreichen und weniger erfolgreichen Probanden liegen vor allem darin, dass die Erfolgreichen zwischen den verschiedenen Stationen hin und her springen, während die weniger Erfolgreichen länger an den einzelnen Stationen verweilen (Übergang AO-AO 8.64% versus 18.75%; Übergang GO-GO 15.5% versus 19.97%). Interessant ist auch, dass die erfolgreichen Probanden nach einer Entscheidung Informationen einholen (4.24%), während die weniger erfolgreichen Probanden dies kaum tun (0.75%).

In ähnlicher Weise werten Howie und Vicente (1998) Daten zu ihrem Simulationsszenario „DURESS II“ (Dual Reservoir System Simulation) aus, bei dem die Vpn eine thermo-hydraulische Prozesskontrolle in Echtzeit zu leisten hatten. Die über einen Zeitraum von sechs Monaten anfallenden Daten von sechs Ingenieuren (insgesamt jeweils 224 Versuche) wurden nach einem Vorschlag von Moray, Lootsteen und Pajak (1986) in Form sogenannter „action transition graphs“ repräsentiert. Dabei wird jede der 12 möglichen Aktionen als Knoten betrachtet, die kreisförmig aufgetragen werden. Die Häufigkeit der Übergänge zwischen je zwei Knoten wird durch entsprechend dicke Linien repräsentiert, so dass ein Geflecht unterschiedlich starker Verbindungen zwischen den einzelnen Aktionsmöglichkeiten entsteht (Transitionsdaten). Interessant ist dann z.B. ein Vergleich der Graphen am Anfang und am Ende der sechsmonatigen Trainingsphase, aus dem sich eine deutliche Steigerung der Expertise als Vereinfachung der Übergänge ergibt. Neben solchen eher qualitativen Inspektionen der Graphen lässt sich auch ein quantitatives Maß der Komplexität über die Anzahl der Verbindungen bestimmen. Auch auf dieser Ebene zeigt sich mit zunehmender Expertise eine Abnahme der Aktionskomplexität. Bei Howie und Vicente finden sich noch weitere spezielle Auswertungsvorschläge, auf die hier nicht näher eingegangen werden kann.

Das von Howie und Vicente (1998) eingesetzte Szenario „DURESS II“ weist übrigens Ähnlichkeiten mit der bei Gonzales, Lerch und Lebiere (2003) verwendeten „Water Purification Plant“ auf und repräsentiert einen Problemtyp, der bereits von Crossman und Cooke (1974) untersucht wurde. Für alle diese Aufgaben bietet sich damit ein vereinheitlichter Auswertungsrahmen an, der aber prinzipiell auch auf andere Problemtypen übertragen werden kann.

Cañas, Quesada, Antoli und Fajardo (2003) haben dies für das Szenario „Fire Chief“ von Omodei und Wearing (1995) getan. Ihre Auswertungsidee soll hier ganz kurz skizziert

werden, da sie die Daten in eine komplexere Auswertungsstrategie eingebunden haben, die aus vier Schritten besteht:

- (1) Für jeden Versuchsdurchgang einer Vp wird aus dem Protokollfile eine Matrix der empirisch ermittelten Übergänge erstellt.
- (2) Es werden prototypische Strategien definiert und in Simulationsläufen realisiert. Daraus ergeben sich „theoretische“ Übergangsmatrizen für jede der Strategien.
- (3) Die empirischen Übergangsmatrizen der Vpn werden mit den theoretischen, je nach Strategie unterschiedlichen Übergangsmatrizen korreliert.
- (4) Vor dem Hintergrund der so ermittelten Ähnlichkeiten werden Vpn hinsichtlich der bei ihnen zu vermutenden Strategien klassifiziert (Vp-Übergangsmatrix korreliert hoch mit Übergangsmatrix gemäß Strategie A, aber niedrig mit derjenigen gemäß Strategie B).

Mit diesem Vorgehen wird ein quantitatives Verfahren zur Strategie-Identifikation vorgeschlagen, das über eine qualitative Einstufung hinausgeht. Inwiefern dieser Weg neue Erkenntnisse liefert, bleibt abzuwarten – die beiden theoretisch definierten Strategie-Prototypen (MOVE-DROP: „Fahre zum nächsten Feuer und lösche es!“, CONTROL: „Finde das nächste Feuer und erzeuge Schneisen in der Nähe!“) sind zwei einfache Vorgehensweisen, deren Hauptvorteil darin besteht, dass sie voneinander unabhängig sind (die durch sie erzeugten Übergangsmatrizen korrelieren nicht). Inwiefern sie das Etikett „Strategie“ verdienen, ist allerdings fraglich. Nachteile der beschriebenen Vorgehensweise sind zudem die Unvollständigkeit der Strategien, die fehlende Beachtung von Rateprozessen sowie die hochgradige Spezifität der zwei Vorgehensweisen, die nicht auf andere Problemstellungen übertragen werden können.

Fazit. Die Analyse von Zustandsübergängen entspricht der Forderung nach Prozessanalyse: Nicht mehr ein einzelnes Datum wird analysiert, sondern die Verbindungen zwischen mindestens je zwei Parametern wird als Untersuchungsgegenstand angesehen. Es mag vielleicht übertrieben klingen, die Abfolge von nur zwei Ereignissen als Prozess zu bezeichnen. Jedoch ist dies der Anfang einer Prozessorientierung, den wir nicht hoch genug bewerten können, findet diese doch erst allmählich Anwendung bei komplexen Problemen.

### 2.3.4. Latent Semantic Analysis

Das Verfahren der „Latent Semantic Analysis“ (LSA) wurde ursprünglich dazu entwickelt, die Ähnlichkeit von Wortbedeutungen und Textpassagen in großen Textbeständen zu ermitteln (Deerwester, Dumais, Furnas, Landauer & Harshman, 1990; Landauer & Dumais, 1997). Nach Art einer Faktorenanalyse wird eine sehr große Datenmatrix, die aus einzelnen Wörtern gekreuzt mit ihren Kontexten besteht, so reduziert, dass eine Wortbedeutung typischerweise auf einen 100- bis 500-dimensionalen Vektor abgebildet wird. Kintsch (2001) beschreibt, wie sich mittels LSA die Interpretation von Metaphern („Mein Rechtsanwalt ist ein Haifisch“) oder die Disambiguierung von Homonymen („Ich gehe zur Bank“) vornehmen lässt. Sogar die automatische Qualitätsbeurteilung studentischer Essays scheint nach Kintsch damit erfolgreich möglich.

Was kann dieses Verfahren für die Denk- und Problemlöseforschung leisten? Quesada, Kintsch und Gomez (2002) haben die LSA angewendet auf Daten aus dem Szenario „Firechief“ von Omodei und Wearing (1995). Ausgehend von einem Korpus von 360.199 Aktionen (in Analogie zu den Wörtern) aus insgesamt 3.441 Versuchsdurchgängen (in Analogie zum Wortkontext) haben sie eine Dimensionsreduzierung vorgeschlagen. Am Beispiel kleinerer Datensätze zeigen sie die Möglichkeiten von Ähnlichkeitsbestimmungen verschiedener Eingriffssequenzen, der Ermittlung von Übergangswahrscheinlichkeiten sowie der Gütebestimmung im Vergleich LSA versus menschliche Beurteiler.

Fazit. Als Vorteil des vorgeschlagenen Vorgehens kann dessen breite Verwendbarkeit im Rahmen von Log-File-Analysen herausgehoben werden. Allerdings braucht man dafür riesige Datensätze. Außerdem ist es problematisch, dass innerhalb eines Versuchsdurchgangs keine Reihenfolge-Effekte betrachtet werden können, da die LSA alle möglichen Abfolgen innerhalb eines Trials als äquivalent ansieht. Abhilfe könnte hier eine Verkleinerung der Triallänge bieten.

## 3. Schlussbemerkungen

Die hier gelieferte Zusammenstellung von Methoden der Denk- und Problemlöseforschung weist eine bemerkenswerte Vielfalt auf. Dem schwierigen Gegenstand entsprechend umfasst die Liste ein reichhaltiges Repertoire äußerst verschiedenartiger Zugänge. Wir wollen abschließend noch einmal versuchen, diese Zugänge den in der Zeitskala menschlicher

Handlungen genannten Zeiträumen (Newell, 1990, vgl. auch Abschnitt 2.2.2) zuzuordnen, die bereits vorher erwähnt wurden (vgl. Abbildung FINAL).

---

Hier bitte Abbildung FINAL einfügen

---

Die Darstellung der Verfahren in Abbildung FINAL zeigt vom biologischen zum sozialen Band (von oben nach unten) eine abnehmende Kontrolle über die Wirkung des Stimulus durch den Versuchsleiter, gleichzeitig aber eine Zunahme der Kontrolle bzw. Freiheit für den Beobachter. Von links nach rechts (subjektive Verfahren bis physiologische Verfahren) nimmt die Objektivität der Verfahren zu, gleichzeitig nimmt mit einer zunehmenden Spezifität der untersuchbaren Fragestellung das Ausmaß an Generalisierbarkeit der gefundenen Ergebnisse ab. Ausgehend von der Zielsetzung, Denken, Erleben und Verhalten möglichst reliabel, valide und objektiv zu beschreiben, ohne dabei das natürliche Verhalten der Vp einzuschränken, also möglichst natürliche Untersuchungsbedingungen zu wahren, ist sowohl eine feinkörnige Analyse als auch eine Übertragung der gefundenen Ergebnisse auf das rationale und soziale Band notwendig. Ein sinnvoller Schritt auf dem Weg zur Weiterentwicklung von Verfahren scheint zunächst die Kombination verschiedener Methoden aus unterschiedlichen Zeit- und Objektivitätskategorien.

So schließen wir diesen Überblick mit der Erwartung, dass die aufgezeigten Lücken sicher noch gefüllt werden und durch kreative Weiter- bzw. Neuentwicklung von Verfahren auch unser theoretisches Verständnis von Denken und Problemlösen wachsen wird. Allerdings werden fundamentale Probleme wohl bestehen bleiben: Wie verbindet man z.B. Gehirnzustände, die man beobachten und messen kann, mit Geisteszuständen, die zum jetzigen Zeitpunkt nur mitgeteilt werden können? Zur Beantwortung dieser und anderer Fragen müssen die Grenzen unserer Disziplin überschritten werden.

## 4. Literatur

- Afflerbach, P. (2000). Verbal reports and protocol analysis. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 163-179). Mahwah, NJ: Lawrence Erlbaum.
- Ahern, S. & Beatty, J. (1978). Pupillary responses during information processing vary with scholastic aptitude test scores. *Science*, *205*, 1289-1292.
- Albert, D. & Gundlach, H. (Eds.). (1997). *Apparative Psychologie: Geschichtliche Entwicklung und gegenwärtige Bedeutung*. Lengerich: Pabst Science Publishers.
- Alexandridis, E., Leendertz, J. A. & Barbur, J. L. (1991). Methods for studying the behaviour of the pupil. *Journal of Psychophysiology*, *5*, 223-239.
- Amelang, M. & Bartussek, D. (2001). *Differentielle Psychologie und Persönlichkeitsforschung. Fünfte, aktualisierte und erweiterte Auflage*. Stuttgart: Kohlhammer.
- Anderson, J. R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science*, *26*, 85-112.
- Anderson, J. R., Bothell, D. & Douglas, S. (2004). Eye movements do not reflect retrieval processes: Limits of the eye-mind hypothesis. *Psychological Science*, *15*, 225-231.
- Anzai, Y. & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review*, *86*, 124-140.
- Arima, J. K. & Wilson, G. E. (1972). *Situational stress, anxiety, and the pupillary response*. Paper presented at the APA 80th Annual Convention.
- Atherton, M., Zhuang, J., Bart, W. M., Hu, X. & He, S. (2003). A functional MRI study of high-level cognition. I. The game of chess. *Cognitive Brain Research*, *16*, 26-31.
- Backlund, L., Skånér, Y., Montgomery, H., Bring, J. & Strender, L.-E. (2003). Doctors' decision processes in a drug-prescription task: The validity of rating scales and think-aloud reports. *Organizational Behavior and Human Decision Processes*, *91*, 108-117.
- Baker, E. L. & O'Neil, H. F. (2002). Measuring problem solving in computer environments: Current and future states. *Computers in Human Behavior*, *18*, 609-622.
- Ballard, D. H., Hayhoe, M. M. & Pelz, J. (1995). Memory representations in natural tasks. *Cognitive Neuroscience*, *7*, 66-80.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K. & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*, 723-767.
- Barker, A. T., Jalinous, R. & Freeston, I. L. (1985). Non-invasive magnetic stimulation of human motor cortex. *Lancet*, *1*(8437), 1106-1107.
- Bartlett, F. C. (1958). *Thinking. An experimental and social study*. New York: Basic Books.

- Batchelder, W. H. (1998). Multinomial processing tree models and psychological assessment. *Psychological Assessment, 10*, 331-344.
- Batchelder, W. H. & Riefer, D. M. (1999). Theoretical and empirical review of multinomial processing tree modeling. *Psychonomic Bulletin & Review, 6*, 57-86.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*, 276-292.
- Beatty, J. (1989). Pupillometric signs of selective attention in man. In E. Donchin, G. C. Galbraith & M. L. Kietzman (Eds.), *Neurophysiology and psychophysiology: Basic mechanisms and clinical applications* (pp. 138-143). New York: Academic Press.
- Beatty, J. & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science, 5*, 371-372.
- Beatty, J. & Lucero-Wagoner, B. (1978). Pupillometric signs of brain activation vary with level of cognitive processing. *Science, 199*, 1216-1218.
- Beatty, J. & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142-162). New York: Cambridge University Press.
- Beckmann, J. F. (1994). *Lernen und komplexes Problemlösen. Ein Beitrag zur Konstruktvalidierung von Lerntests*. Bonn: Holos.
- Berry, D. C. & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology, 36A*, 209-231.
- Berry, D. C. & Broadbent, D. E. (1987a). The combination of explicit and implicit learning processes in task control. *Psychological Research, 49*, 7-15.
- Berry, D. C. & Broadbent, D. E. (1987b). Explanation and verbalisation in a computer assisted search task. *Quarterly Journal of Experimental Psychology, 39A*, 585-609.
- Berry, D. C. & Broadbent, D. E. (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology, 79*, 251-272.
- Best, J. B. (1990). Knowledge acquisition and strategy action in "Mastermind" problems. *Memory & Cognition, 18*, 54-64.
- Bhaskar, R. & Simon, H. A. (1977). Problem solving in semantically rich domains: An example from engineering thermodynamics. *Cognitive Science, 1*, 193-215.
- Biggs, S. F., Rosman, A. J. & Sergenian, G. K. (1993). Methodological issues in judgment and decision-making research: Concurrent verbal protocol validity and simultaneous traces of process. *Journal of Behavioral Decision Making, 6*, 187-206.
- Boos, M., Morguet, M., Meier, F. & Fisch, R. (1990). Zeitreihenanalysen von Interaktionsprozessen bei der Bearbeitung komplexer Probleme in Expertengruppen. *Zeitschrift für Sozialpsychologie, 21*, 53-64.

- Boos, M., Scharpf, U. & Fisch, R. (1991). Eine Methode zur Analyse von Interaktionsprozessen beim Problemlösen und Entscheiden in Sitzungen. *Zeitschrift für Arbeits- und Organisationspsychologie*, 35, 115-121.
- Bourne, L. E., Ekstrand, B. R. & Dominowski, R. L. (1971). *The psychology of thinking*. Englewood Cliffs, NJ: Prentice-Hall.
- Bradshaw, J. L. (1968). Pupil size and problem solving. *Quarterly Journal of Experimental Psychology*, 20, 116-122.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433-436.
- Broadbent, D. E. (1985). A question of levels: Comments on McClelland and Rumelhart. *Journal of Experimental Psychology: General*, 114, 189-192.
- Broadbent, D. E. & Aston, B. (1978). Human control of a simulated economic system. *Ergonomics*, 21, 1035-1043.
- Buchner, A. (1999). Komplexes Problemlösen vor dem Hintergrund der Theorie finiter Automaten. *Psychologische Rundschau*, 50, 206-212.
- Buchner, A. & Funke, J. (1992). On the use of finite state automata in problem solving research. *International Journal of Psychology*, 27, 135.
- Buchner, A. & Funke, J. (1993). Finite state automata: Dynamic task environments in problem solving research. *Quarterly Journal of Experimental Psychology*, 46A, 83-118.
- Buchner, A., Funke, J. & Berry, D. C. (1995). Negative correlations between control performance and verbalizable knowledge: Indicators for implicit learning in process control tasks? *Quarterly Journal of Experimental Psychology*, 48A, 166-187.
- Burr, D. C., Morrone, M. C. & Ross, J. (2001). Separate visual representations for perception and action revealed by saccadic eye movements. *Current Biology*, 11, 798-802.
- Cacioppo, J. T. & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45, 16-28.
- Cacioppo, J. T., Tassinary, L. G. & Berntson, G. G. (2000). Psychophysiological science. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (2nd ed., pp. 3-23). Cambridge: Cambridge University Press.
- Call, J. & Tomasello, M. (1995). The use of social information in the problem-solving of orangutans (*Pongo pygmaeus*) and human children (*Homo sapiens*). *Journal of Comparative Psychology*, 109(3), 308-320.
- Cañas, J. J., Quesada, J. F., Antolí, A. & Fajardo, I. (2003). Cognitive flexibility and adaptability to environmental changes in dynamic complex problem solving tasks. *Ergonomics*, 46, 482-501.
- Card, S. K., Moran, T. P. & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Carpenter, R. H. (1988). *Movements of the eyes* (2nd ed. ed.). London: Pion.
- Celli, G. (1986). *Der letzte Alchemist. Betrachtungen über Komik und Wissenschaft*. Stuttgart: ComMedia & Arte.
- Channon, S. & Crawford, S. (1999). Problem-solving in real-life-type situations: The effects of anterior and posterior lesions on performance. *Neuropsychologia*, 37, 757-770.
- Chase, W. G. & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Clancey, W. J. (1997). *Situated cognition: On human knowledge and computer representations*. Cambridge: Cambridge University Press.
- Clark, A. (1997). *Being there. Putting brain, body, and world together again*. Cambridge, MA: MIT Press.
- Clarke, H. H. & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, 3, 472-517.
- Clearwater, S. H., Huberman, B. A. & Hogg, T. (1991). Cooperative solution of constraint satisfaction problems. *Science*, 254, 1181-1183.
- Collewyn, H. (1999). Eye movement recording. In R. H. S. Carpenter & J. G. Robson (Eds.), *Vision research: A practical guide to laboratory methods* (pp. 245-285). Oxford: Oxford University Press.
- Cooper, L. A. & Shepard, R. N. (1973a). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing* (pp. 75-176). New York: Academic Press.
- Cooper, L. A. & Shepard, R. N. (1973b). The time required to prepare for a rotated stimulus. *Memory & Cognition*, 1, 246-250.
- Cornelissen, F. W., Peters, E. M. & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, 34, 613-617.
- Cowey, A. & Walsh, V. (2001). Tickling the brain: Studying visual sensation, perception and cognition by transcranial magnetic stimulation. *Progress in Brain Research*, 134, 411-425.
- Craik, F. I. M. & Lockhart, R. J. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Crane, H. D. & Steele, C. M. (1985). Generation-V dual-Purkinje-image eye tracker. *Applied Optics*, 24, 527-537.
- Crossman, E. R. F. & Cooke, J. E. (1974). Manual control of slow-response systems. In E. Edwards & F. P. Lees (Eds.), *The human operator in process control* (pp. 51-66). London: Taylor & Francis.
- de Groot, A. D. (1978). *Thought and choice in chess*. The Hague: Mouton.

- deCharms, R. C., Christoff, K., Glover, G. H., Pauly, J. M., Whitfield, S. & Gabrieli, J. D. (2004). Learned regulation of spatially localized brain activation using real-time fMRI. *Neuroimage*, *21*, 436-443.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. & Harshman, R. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, *41*, 391-407.
- Deffner, G. (1984). *Lautes Denken - Untersuchungen zur Qualität eines Datenerhebungsverfahrens*. Frankfurt: Lang.
- Deffner, G. (1987). Eye movement patterns as an indicator of solution strategy in solving n-term series tasks. In G. Lüer & U. Lass (Eds.), *European conference on eye movements*. Göttingen: Hogrefe.
- Deubel, H. (1994). Visuelle Verarbeitung und kognitive Faktoren bei der Generierung sakkadischer Augenbewegungen. In W. Prinz & B. Bridgeman (Eds.), *Wahrnehmung (=Enzyklopädie der Psychologie, Themenbereich Kognition, Band 1)* (pp. 189-253). Göttingen: Hogrefe.
- Dickson, J., McLennan, J. & Omodei, M. M. (2000). Effects of concurrent verbalization on a time pressured dynamic decision task. *Journal of General Psychology*, *127*, 217-228.
- Dienes, Z. & Fahey, R. (1995). Role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 848-862.
- Dillon, R. F. (1985). Eye movement analysis of information processing under different testing conditions. *Contemporary Educational Psychology*, *10*, 387-395.
- Dominowski, R. L. (1998). Verbalization and problem solving. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 25-45). Mahwah, NJ: Lawrence Erlbaum Associates.
- Donders, F. C. (1969/1868). On the speed of mental processes. *Acta Psychologica*, *30*, 412-431.
- Dörner, D. (1981). Über die Schwierigkeiten menschlichen Umgangs mit Komplexität. *Psychologische Rundschau*, *32*, 163-179.
- Dörner, D. (1982). Wie man viele Probleme zugleich löst - oder auch nicht. *Sprache & Kognition*, *1*, 55-66.
- Dörner, D. (1989). Die kleinen grünen Schildkröten und die Methoden der experimentellen Psychologie. *Sprache & Kognition*, *8*, 86-97.
- Dörner, D. (1992). Über die Philosophie der Verwendung von Mikrowelten oder "Computerszenarios" in der psychologischen Forschung. In H. Gundlach (Ed.), *Psychologische Forschung und Methode: Das Versprechen des Experiments. Festschrift für Werner Traxel* (pp. 53-87). Passau: Passavia-Universitäts-Verlag.
- Dörner, D. (1999). *Bauplan für eine Seele*. Reinbek: Rowohlt.

- Dörner, D., Bartl, C., Detje, F., Gerdes, J., Halcour, D., Schaub, H., et al. (2002). *Die Mechanik des Seelenwagens. Eine neuronale Theorie der Handlungsregulation*. Bern: Hans Huber.
- Dörner, D., Kreuzig, H. W., Reither, F. & Stäudel, T. (1983). *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Hans Huber.
- Dörner, D. & Reither, F. (1978). Über das Problemlösen in sehr komplexen Realitätsbereichen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 25, 527-551.
- Dörner, D. & Wearing, A. J. (1995). Complex problem solving: Toward a (computersimulated) theory. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 65-99). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Duncker, K. (1935/1974). *Zur Psychologie des produktiven Denkens*. Berlin: Julius Springer.
- El Yagoubi, R., Lemaire, P. & Besson, M. (2003). Different brain mechanisms mediate two strategies in arithmetic: Evidence from Event-Related brain Potentials. *Neuropsychologia*, 41(7), 855-862.
- Endres, J. & Putz-Osterloh, W. (1994). Komplexes Problemlösen in Kleingruppen: Effekte des Vorwissens, der Gruppenstruktur und der Gruppeninteraktion. *Zeitschrift für Sozialpsychologie*, 25, 54-70.
- Epelboim, J. & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research*, 41, 1561-1574.
- Erdfelder, E. & Funke, J. (Eds.). (2004). *Allgemeine Psychologie und deduktivistische Methodologie*. Göttingen: Vandenhoeck & Ruprecht.
- Ericsson, K. A. & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.
- Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (2nd ed.). Cambridge, MA: MIT Press.
- Eveland, W. P. & Dunwoody, S. (2000). Examining information processing on the World Wide Web using think aloud protocols. *Media Psychology*, 2, 219-244.
- Fincham, J. M., Carter, C. S., van Veen, V., Stenger, V. A. & Anderson, J. R. (2002). Neural mechanisms of planning: A computational analysis using event-related fMRI. *Proceedings of the National Academy of Science*, 99, 3346-3351.
- Findlay, J. M. & Gilchrist, I. D. (2003). *Active vision. The psychology of looking and seeing*. Oxford: Oxford University Press.
- Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J. & Mazziotta, J. C. (1997). *Human brain function*. San Diego, CA: Academic Press.

- Franzen, U. & Merz, F. (1988). Einfluß des Verbalisierens auf die Leistung bei Intelligenzprüfungen: Neue Untersuchungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 8, 117-134.
- Frensch, P. A. & Rüniger, D. (2003). Implicit learning. *Current Directions in Psychological Science*, 12, 13-17.
- Friedman, I. A. (1996). Deliberation and resolution in decision-making processes: A self-report scale for adolescents. *Educational and Psychological Measurement*, 56, 881-890.
- Friston, K. J., Holmes, A. P., Poline, J.-B., Price, C. J. & Frith, C. D. (1996). Detecting activations in PET and fMRI: Levels of inference and power. *Neuroimage*, 4, 223-235.
- Fritz, A. & Funke, J. (1990). Superhirn trotz Teilleistungsschwäche? *Acta Paedopsychiatrica*, 53, 146-162.
- Fritz, A. & Funke, J. (1995). Übersicht über vorliegende Verfahren zur Planungsdiagnostik. In J. Funke & A. Fritz (Eds.), *Neue Konzepte und Instrumente zur Planungsdiagnostik* (pp. 47-78). Bonn: Deutscher Psychologen Verlag.
- Funke, J. (1982). TAILOR-24. Ein BASIC-Programm zur Simulation einer Schneiderwerkstatt [Computer program]. Trier: Fachbereich I - Psychologie - der Universität Trier.
- Funke, J. (1983). Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Prädiktor? *Diagnostica*, 29, 283-302.
- Funke, J. (1984). Diagnose der westdeutschen Problemlöseforschung in Form einiger Thesen. *Sprache & Kognition*, 3, 159-172.
- Funke, J. (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle. *Zeitschrift für Psychologie*, 193, 435-457.
- Funke, J. (1986). *Komplexes Problemlösen - Bestandsaufnahme und Perspektiven*. Heidelberg: Springer.
- Funke, J. (1992). *Wissen über dynamische Systeme: Erwerb, Repräsentation und Anwendung*. Berlin: Springer.
- Funke, J. (1993). Microworlds based on linear equation systems: A new approach to complex problem solving and experimental results. In G. Strube & K.-F. Wender (Eds.), *The cognitive psychology of knowledge* (pp. 313-330). Amsterdam: Elsevier Science Publishers.
- Funke, J. (1998). Computer-based testing and training with scenarios from complex problem-solving research: Advantages and disadvantages. *International Journal of Selection and Assessment*, 6, 90-96.
- Funke, J. (2003). *Problemlösendes Denken*. Stuttgart: Kohlhammer.

- Funke, J. & Buchner, A. (1992). Finite Automaten als Instrumente für die Analyse von wissensgeleiteten Problemlöseprozessen: Vorstellung eines neuen Untersuchungsparadigmas. *Sprache & Kognition*, 11, 27-37.
- Funke, J. & Gerdes, H. (1993). Manuale für Videorekorder: Auswahl von Textinhalten unter Verwendung der Theorie endlicher Automaten. *Zeitschrift für Arbeitswissenschaft*, 47, 44-49.
- Funke, J. & Krüger, T. (1995). „Plan-A-Day“: Konzeption eines modifizierbaren Instruments zur Führungskräfte-Auswahl sowie erste empirische Befunde. In J. Funke & A. Fritz (Eds.), *Neue Konzepte und Instrumente zur Planungsdiagnostik* (pp. 97-120). Bonn: Deutscher Psychologen Verlag.
- Funke, J. & Steyer, R. (1985). Komplexes Problemlösen als Konstruktion und Anwendung von Kausalmodellen. In D. Albert (Ed.), *Bericht über den 34. Kongreß der DGfPs in Wien 1984* (pp. 264-267). Göttingen: Hogrefe.
- Garber, P., Alibali, M. W. & Goldin-Meadow, S. (1998). Knowledge conveyed in gesture is not tied to the hands. *Child Development*, 69, 75-84.
- Garber, P. & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in adults and children. *Cognitive Science*, 26, 817-831.
- Gauggel, S., Deckersbach, T. & Rolko, C. (1998). Entwicklung und erste Evaluation einer Skala zur Beurteilung von Handlungs-, Planungs- und Problemlösestörungen. *Zeitschrift für Neuropsychologie*, 9, 3-17.
- Geilhardt, T. & Mühlbradt, T. (Eds.). (1995). *Planspiele im Personal- und Organisationsmanagement*. Göttingen: Verlag für Angewandte Psychologie.
- Genovese, C. R., Lazar, N. A. & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15, 870-878.
- Gigerenzer, G. (1981). *Messung und Modellbildung*. München: Reinhardt.
- Gigerenzer, G. (1988). Woher kommen Theorien über kognitive Prozesse? *Psychologische Rundschau*, 39, 91-100.
- Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in Cognitive Psychology. *Psychological Review*, 98, 254-267.
- Gigerenzer, G. & Selten, R. (Eds.). (2001). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.
- Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Science*, 3, 419-429.
- Goldman, S. R., Zech, L. K., Biswas, G. & Noser, T. (1999). Computer technology and complex problem solving: Issues in the study of complex cognitive activity. *Instructional Science*, 27, 235-268.

- Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin*, 77, 340-355.
- Gonzales, C., Lerch, J. F. & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27, 591-636.
- Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences*, 16, 1-14.
- Gouzoulis-Mayfrank, E., Thimm, B., Rezk, M., Hensen, G. & Daumann, J. (2002). Memory impairment suggests hippocampal dysfunction in abstinent ecstasy users. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 27, 819-827.
- Grafman, J. (1994). Neuropsychology of the prefrontal cortex. In D. W. Zaidel (Ed.), *Neuropsychology* (pp. 159-181). San Diego, CA: Academic Press.
- Grafman, J. (1995). Similarities and distinctions among current models of prefrontal cortical functions. In J. Grafman, K. J. Holyoak & F. Boller (Eds.), *Structure and functions of the human prefrontal cortex* (pp. 337-368). New York: New York Academy of Science (=Annals of the New York Academy of Science, Vol. 769).
- Grant, E. R. & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, 14, 462-466.
- Graumann, C. F. (1964). Phänomenologie und deskriptive Psychologie des Denkens. In R. Bergius (Ed.), *Allgemeine Psychologie. I. Der Aufbau des Erkennens. 2. Halbband: Lernen und Denken* (pp. 493-518). Göttingen: Hogrefe.
- Graumann, C. F. (1965). Denken und Denkpsychologie. In C. F. Graumann (Ed.), *Denken* (pp. 13-43). Köln: Kiepenheuer & Witsch.
- Graumann, C. F. (1997). Zur Ökologie des Gedächtnisses. In G. Lürer & U. Lass (Eds.), *Erinnern und Behalten. Wege zur Erforschung des menschlichen Gedächtnisses* (pp. 269-286). Göttingen: Vandenhoeck & Ruprecht.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics. Vol. 3: Speech acts* (pp. 41-58). New York: Academic Press.
- Groeben, N. (1986). *Handeln, Tun, Verhalten als Einheiten einer verstehend-erklärenden Psychologie: Wissenschaftstheoretischer Überblick und Programmentwurf zur Integration von Hermeneutik und Empirismus*. Tübingen: Francke.
- Groner, R. & Fraisse, P. (1982). *Cognition and eye movements*. New York: North-Holland.
- Groner, R. & Groner, M. T. (1989). Attention and eye movement control: An overview. *European Archives of Psychiatry & Neurological Sciences*, 239, 9-16.
- Haider, H. (1992). Implizites Wissen und Lernen. Ein Artefakt? *Zeitschrift für Experimentelle und Angewandte Psychologie*, 39, 68-100.
- Haider, H. (1993). Was ist implizit am impliziten Wissen und Lernen? *Sprache & Kognition*, 12, 44-52.

- Haider, H. & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *25*, 172-190.
- Hauser, M. D. (2001). *Wilde Intelligenz. Was Tiere wirklich denken*. München: C.H. Beck.
- Hayes, J. R. & Simon, H. A. (1977). Psychological differences among problem isomorphs. In N. J. Castellan, D. B. Pisoni & G. R. Potts (Eds.), *Cognitive theory* (Vol. 2, pp. 21-41). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hayhoe, M. M., Shrivastava, A., Mruczek, R. & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, *3*, 49-63.
- Heineken, E., Arnold, H.-J., Kopp, A. & Soltysiak, R. (1992). Strategien des Denkens bei der Regelung eines einfachen dynamischen Systems unter verschiedenen Totzeitbedingungen. *Sprache & Kognition*, *11*, 136-148.
- Hershey, D. A., Walsh, D. A., Read, S. J. & Chulef, A. S. (1990). The effects of expertise on financial problem solving: Evidence for goal-directed, problem-solving scripts. *Organizational Behavior and Human Decision Processes*, *46*, 77-101.
- Hess, E. H. & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science*, *132*, 349-350.
- Hess, E. H. & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*, 1190-1192.
- Hodgson, T. L., Bajwa, A., Owen, A. M. & Kennard, C. (2000a). The strategic control of gaze direction in the Tower of London task. *Journal of Cognitive Neuroscience*, *12*, 894-907.
- Hodgson, T. L., Bajwa, A., Owen, A. M. & Kennard, C. (2000b). The strategic control of gaze direction in the Tower of London task. *Journal of Cognitive Neuroscience*, *12*(5), 894-907.
- Hodgson, T. L., Tiesman, B., Owen, A. M. & Kennard, C. (2002). Abnormal gaze strategies during problem solving in Parkinson's disease. *Neuropsychologia*, *40*, 411-422.
- Hoeks, B. & Levelt, W. J. M. (1993). Pupillary dilation as a measure of attention: A quantitative system analysis. *Behavior Research Methods & Instrumentation*, *25*, 16-26.
- Hoffman, J. E. & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*, 787-795.
- Hopcroft, J. E. & Ullman, J. D. (1988). *Einführung in die Automatentheorie, formale Sprachen und Komplexitätstheorie*. New York: Addison-Wesley.
- Howie, D. E. & Vicente, K. J. (1998). Measures of operator performance in complex, dynamic microworlds: Advancing the state of the art. *Ergonomics*, *41*, 85-150.

- Huber, G. L. & Mandl, H. (Eds.). (1994). *Verbale Daten: Eine Einführung in die Grundlagen und Methoden der Erhebung und Auswertung* (2. Aufl.). Weinheim: Beltz.
- Huber, O. (im Druck). Entscheiden unter Risiko: Aktive Risiko-Entschärfung. *Psychologische Rundschau*.
- Huber, O., Wider, R. & Huber, O. W. (1997). Active information search and complete information presentation in naturalistic risky decision tasks. *Acta Psychologica*, 95, 15-29.
- Hübner, R. (1987). Eine naheliegende Fehleinschätzung des Zielabstandes bei der zeitoptimalen Regelung dynamischer Systeme. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 34, 38-53.
- Hussy, W. (1975). Information processing and human sequential predictive behavior. *Acta Psychologica*, 39, 351-367.
- Hussy, W. (1982). Transinformations-Komponenten-Analyse: zur Formalisierung von Informationsverarbeitungsqualität bei sequentiellen Vorhersageproblemen. *Archiv für Psychologie*, 134, 137-159.
- Hussy, W. (1989). Strategien zur Bewältigung umfänglicher, problemrelevanter Informationsangebote im Altersvergleich. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 21, 24-39.
- Hussy, W. (1991). Eine experimentelle Studie zum Intelligenzkonzept "Verarbeitungskapazität". *Diagnostica*, 37, 314-333.
- Irving, R. W. (1978). Towards an optimum mastermind strategy. *Journal of Recreational Mathematics*, 11, 81-87.
- Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 18, 307-317.
- Irwin, D. E. & Gordon, R. D. (1998). Eye movements, attention and trans-saccadic memory. *Visual Cognition*, 5, 127-155.
- Jahanshahi, M. & Rothwell, J. (2000). Transcranial magnetic stimulation studies of cognition: An emerging field. *Experimental Brain Research*, 131, 1-9.
- Janisse, M. P. (1977). *Pupillometry*. Washington, DC: Hemisphere.
- Jansson, A. (1994). Pathologies in dynamic decision making: Consequences or precursors of failure? *Sprache & Kognition*, 13, 160-173.
- Jeck, S. (1997). *Planen und das Lösen von Alltagsproblemen*. Lengerich: Pabst Science Publishers.
- Johnson, D. A. (1971). Pupillary responses during a short-term memory task: Cognitive processing, arousal, or both? *Journal of Experimental Psychology*, 90, 311-318.

- Jung-Beeman, M., Bowden, E., Haberman, J., Frymiare, J., Arambel-Liu, S., Greenblatt, R., et al. (2004). Neural activity when people solve verbal problems with insight. *PLoS Biology*, 2(4).
- Juris, M. & Velden, M. (1977). The pupillary response to mental overload. *Physiological Psychology*, 5, 421-424.
- Just, M. A. & Carpenter, P. A. (1976). The role of eye-fixation research in cognitive psychology. *Behavior Research Methods & Instrumentation*, 8, 139-143.
- Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- Just, M. A. & Carpenter, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, 92, 137-172.
- Just, M. A. & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, 47, 310-339.
- Kagan, N. & Kagan, H. (1991). Interpersonal process recall. In P. W. Dowrick (Ed.), *A practical guide to using video in the behavioral sciences* (pp. 221-230). New York: Wiley.
- Kahneman, D. (1973). *Attention and effort*. Englewood-Cliffs, NJ: Prentice Hall.
- Kahneman, D. & Beatty, J. (1967). Pupil diameter and load on memory. *Science*, 154, 1583-1585.
- Kersting, M. (1999). *Diagnostik und Personalauswahl mit computergestützten Problemlöseszenarien? Zur Kriteriumsvalidität von Problemlöseszenarien und Intelligenztests*. Göttingen: Hogrefe.
- Kintsch, W. (2001). Predication. *Cognitive Science*, 25, 173-202.
- Klahr, D. (1981). Untersuchungen zum Problemlösen bei Kindern. In R. H. Kluwe & H. Spada (Eds.), *Studien zur Denkentwicklung* (pp. 240-289). Bern: Hans Huber.
- Klein, G. A., Calderwood, R. & MacGregor, D. (1989). Critical Decision Method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, 19, 462-472.
- Kliegel, M., Martin, M., McDaniel, M. A. & Einstein, G. O. (2002). Complex prospective memory and executive control of working memory: A process model. *Psychologische Beiträge*, 44, 303-318.
- Kluwe, R. H. (1988). Methoden der Psychologie zur Gewinnung von Daten über menschliches Wissen. In H. Mandl & H. Spada (Eds.), *Wissenspsychologie* (pp. 359-385). München: Psychologie Verlags Union.

- Kluwe, R. H. (1995). Single case studies and models of complex problem solving. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 269-291). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Knoblich, G. (1999). Allokation von Aufmerksamkeit und Metakognition beim Problemlösen mit Einsicht. In I. Wachsmuth & B. Jung (Eds.), *KogWis99: Proceedings der 4. Fachtagung der Gesellschaft für Kognitionswissenschaft, Bielefeld, 28.9.-1.10.1999* (pp. 134-139). Sankt Augustin: Infix.
- Knoblich, G., Ohlsson, S., Haider, H. & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1534-1555.
- Knoblich, G., Ohlsson, S. & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory and Cognition*, 29, 1000-1009.
- Knoblich, G. & Rhenius, D. (1995). Zur Reaktivität lauten Denkens beim komplexen Problemlösen. *Zeitschrift für Experimentelle Psychologie*, 42, 419-454.
- Köhler, W. (1921). *Intelligenzprüfungen an Menschenaffen*. Berlin: Springer.
- Köller, O. & Strauß, B. (1994). Was mißt der Kompetenzfragebogen? Eine Reanalyse der Kurzform des Kompetenzfragebogens von Stäudel. *Diagnostica*, 40, 42-60.
- König, E. & Zedler, P. (2002). *Qualitative Forschung* (2. Aufl. ed.). Weinheim: Beltz.
- König, F., Liepmann, D., Holling, H. & Otto, J. (1985). Entwicklung eines Fragebogens zum Problemlösen (PLF). *Zeitschrift für Klinische Psychologie, Psychopathologie und Psychotherapie*, 33, 5-19.
- Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
- Kotkamp, U. (1999). *Elementares und komplexes Problemlösen: Über Invarianzeigenschaften von Denkprozessen*. Lengerich: Pabst Science Publishers.
- Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 1-70). Amsterdam: Elsevier.
- Krause, W. (1981). Internal representation and eye fixation. *Zeitschrift für Psychologie*, 189, 1-13.
- Krause, W. (1988). Identifying cognitive structures through eye movements. In G. Lüer & U. Lass (Eds.), *Eye movement research: physiological and psychological aspects*. Göttingen: Hogrefe.
- Land, M. F. (1999). Motion and vision: Why animals move their eyes. *Journal of Comparative Physiology*, 185, 341-352.
- Land, M. F. & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society London B Biological Science*, 352(1358), 1231-1239.

- Land, M. F. & Hayhoe, M. M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*, 3559-3565.
- Land, M. F. & Lee, D. N. (1994). Where we look when we steer. *Nature*, *369*, 742-744.
- Landauer, T. K. & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction and representation of knowledge. *Psychological Review*, *104*, 211-240.
- Lantermann, E.-D., Döring-Seipel, E., Schmitz, B. & Schima, P. (2000). *SYRENE. Umwelt- und Systemlernen mit Multimedia*. Göttingen: Hogrefe.
- Laughlin, P. R., Lange, R. & Adamopoulos, J. (1982). Selection strategies for "mastermind" problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 475-483.
- Lazarus, R. S., Averill, J. R. & Opton, E. M. (1970). Toward a cognitive theory of emotion. In M. B. Arnold (Ed.), *Feelings and emotions* (pp. 207-232). New York: Academic Press.
- Leigh, R. J. & Zee, D. S. (1999). *The neurology of eye movements. Third edition*. Oxford: Oxford University Press.
- Leighton, J. P. & Dawson, M. R. W. (2001). A parallel distributed processing model of Wason's selection task. *Cognitive Systems Research*, *2*, 207-231.
- Leuthold, H., Sommer, W. & Ulrich, R. (1996). Partial advance information and response preparation: inferences from the lateralized readiness potential. *Journal of Experimental Psychology: General*, *125*, 307-323.
- Levelt, W. J. M. (1991). Die konnektionistische Mode. *Sprache & Kognition*, *10*, 61-72.
- Lienert, G. A. (1964). *D-S-T. Denksport-Test. Handanweisung für die Durchführung und Auswertung*. Göttingen: Hogrefe.
- Loewenfeld, I. E. (1958). Mechanisms of reflex dilation of the pupil. Historical review and experimental analysis. *Documenta Ophthalmologica*, *12*, 185-448.
- Loewenstein, O. & Loewenfeld, I. E. (1969). The pupil. In H. Dawson (Ed.), *The eye* (pp. 255-337). New York: Academic Press.
- Lüer, G. (1973). *Gesetzmäßige Denkabläufe beim Problemlösen*. Weinheim: Beltz.
- Lüer, G., Hübner, R. & Lass, U. (1985). Sequences of eye movements in a problem solving situation. In R. Groner, G. W. McConkie & C. Menz (Eds.), *Eye movements and human information processing* (pp. 299-307). Amsterdam: Elsevier.
- Lüer, G., Lass, U. & Shallo-Hoffmann, J. (1988). *Eye movement research: Physiological and psychological aspects*. Toronto: Hogrefe.
- Luo, J. & Niki, K. (2003). Function of hippocampus in "insight" of problem solving. *Hippocampus*, *13*, 316-323.

- Macho, S. (2002a). Cognitive modeling with spreadsheets. *Behavior Research Methods, Instruments, & Computers*, 34, 19-36.
- Macho, S. (2002b). *Kognitive Modellierung mit Neuronalen Netzen. Eine anwendungsorientierte Einführung*. Bern: Hans Huber.
- Mackworth, N. H. & Morandi, A. J. (1967). The gaze selects information details within pictures. *Perception and Psychophysics*, 2, 547-551.
- Marr, D. (1982). *Vision. A computational investigation into the human representation and processing of visual information*. New York: W. H. Freeman.
- Mayring, P. (2002). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (8. Aufl. ed.). Weinheim: Beltz.
- McCloskey, M. (1991). Networks and theories: The place of connectionism in cognitive science. *Psychological Science*, 2, 387-395.
- McLennan, J. & Omodei, M. M. (1996). The role of prepriming in recognition-primed decision making. *Perceptual & Motor Skills*, 82, 1059-1069.
- McPeck, R. M., Maljkovic, V. & Nakayama, K. (1999). Saccades require focal attention and are facilitated by a short-term memory system. *Vision Research*, 39, 1555-1566.
- Merz, F. (1969). Der Einfluss des Verbalisierens auf die Leistung in Intelligenzaufgaben. *Zeitschrift für Experimentelle Psychologie und Angewandte Psychologie*, 16, 114-137.
- Metcalf, J. (1986). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 288-294.
- Metcalf, J. & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, 15, 238-246.
- Moonen, C. T. W. & Bandettini, P. A. (1999). *Functional MRI*. Berlin: Springer.
- Moray, N., Lootsteen, P. & Pajak, J. (1986). Acquisition of process control skills. *IEEE Transactions on Systems, Man, & Cybernetics*, 16, 497-504.
- Müller, H. (1993). *Komplexes Problemlösen: Reliabilität und Wissen*. Bonn: Holos.
- Müller, S. V., von der Fecht, A., Hildebrandt, H. & Münte, T. F. (2000). Kognitive Therapie von Störungen der Exekutivfunktionen. *Neurologie & Rehabilitation*, 6, 313-322.
- Müller-Gethmann, H., Rinkenauer, G., Stahl, J. & Ulrich, R. (2000). Preparation of response force and movement direction: Onset effects on the lateralized readiness potential. *Psychophysiology*, 37, 507-514.
- Nährer, W. (1986). *Schnelligkeit und Güte als Dimensionen kognitiver Leistung*. Heidelberg: Springer.
- Nährer, W. (1988). Problemkonstruieren statt Problemlösen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 35, 441-450.

- Neuwirth, E. (1982). Some strategies for mastermind. *Zeitschrift für Operations Research*, 26, B257-B278.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Newman, S. D., Carpenter, P. A., Varma, S. & Just, M. A. (2003). Frontal and parietal participation in problem solving in the Tower of London: fMRI and computational modeling of planning and high-level perception. *Neuropsychologia*, 41, 1668-1682.
- Nguyen, A. H. & Stark, L. W. (1993). Model control of image processing: Pupillometry. *Computerized Medical Imaging and Graphics*, 17, 21-33.
- Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Nodine, C. F., Kundel, H. L., Polikoff, J. B. & Toto, L. C. (1988). Using eye movements to study decision making of radiologists. In G. Lüer & U. Lass (Eds.), *Eye movement research: Physiological and psychological aspects* (pp. 349-363). Göttingen: Hogrefe.
- Norman, D. A. (1993). Cognition in the head and in the world: An introduction to the special issue on situated action. *Cognitive Science*, 17, 1-6.
- O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, 46, 461-488.
- O'Regan, K., Deubel, H., Clark, J. J. & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, 7, 191-211.
- Omodei, M. M. & McLennan, J. (1994). Studying complex decision making in natural settings: Using a head-mounted video camera to study competitive orienteering. *Perceptual and Motor Skills*, 79, 1411-1425.
- Omodei, M. M. & Wearing, A. J. (1995). The Fire Chief microworld generating program: An illustration of computer-simulated microworlds as an experimental paradigm for studying complex decision-making behavior. *Behavior Research Methods, Instruments, & Computers*, 27, 303-316.
- Opwis, K. & Spada, H. (1994). Modellierung mit Hilfe wissensbasierter Systeme. In T. Herrmann & W. H. Tack (Eds.), *Methodologische Grundlagen der Psychologie (=Enzyklopädie der Psychologie. Themenbereich B: Methodologie und Methoden. Serie I: Forschungsmethoden der Psychologie. Band 1)* (pp. 199-248). Göttingen: Hogrefe.
- Otto, J. H., Döring-Seipel, E., Grebe, M. & Lantermann, E.-D. (2001). Entwicklung eines Fragebogens zur Erfassung der wahrgenommenen emotionalen Intelligenz. Aufmerksamkeit auf, Klarheit und Beeinflussbarkeit von Emotionen. *Diagnostica*, 47, 178-187.

- Parsons, L. M. (2001). Integrating cognitive psychology, neurology, and neuroimaging. *Acta Psychologica*, 107, 155-181.
- Pascual-Leone, A., Houser, C. M., Reese, K., Shotland, L. I., Grafman, J., Sato, S., et al. (1993). Safety of rapid-rate transcranial magnetic stimulation in normal volunteers. *Electroencephalography and Clinical Neurophysiology*, 89, 120-130.
- Plata, G. (2000). *Augenbewegungen beim Lösen raumgeometrischer Denkaufgaben*. Unpublished Diplomarbeit, Carl von Ossietzky Universität, Oldenburg.
- Plötzner, R. & Kneser, C. (1998). Steps towards the acquisition of expertise: Shifting the focus from quantitative to qualitative problem representations during collaborative problem solving. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the 20th Annual Conference of the Cognitive Science Society* (pp. 842-847). Mahwah, NJ: Lawrence Erlbaum Associates.
- Plötzner, R. & Spada, H. (1993). Multiple mental representations of information in physics problem solving. In G. Strube & K.-F. Wender (Eds.), *The cognitive psychology of knowledge* (pp. 285-312). Amsterdam: Elsevier Science Publishers.
- Plötzner, R., Spada, H., Stumpf, M. & Opwis, K. (1990). Learning qualitative and quantitative reasoning in a microworld for elastic impacts. *European Journal of Psychology of Education*, 4, 501-516.
- Posner, M. I. (1978). *Chronometric explorations of the mind*. Hillsdale, NJ: Lawrence Erlbaum.
- Preußler, W. (1996). Zur Rolle expliziten und impliziten Wissens bei der Steuerung dynamischer Systeme. *Zeitschrift für Experimentelle Psychologie*, 43, 399-434.
- Preußler, W. (1998). Strukturwissen als Voraussetzung für die Steuerung komplexer dynamischer Systeme. *Zeitschrift für Experimentelle Psychologie*, 45, 218-240.
- Pritchard, R. (1990). The evolution of introspective methodology and its implications for studying the reading process. *Reading Psychology: An International Quarterly*, 11, 1-13.
- Putz-Osterloh, W. (1981a). *Problemlöseprozesse und Intelligenztestleistung*. Bern: Huber.
- Putz-Osterloh, W. (1981b). Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg. *Zeitschrift für Psychologie*, 189, 79-100.
- Putz-Osterloh, W. (1985). Selbstreflexionen, Testintelligenz und interindividuelle Unterschiede bei der Bewältigung komplexer Probleme. *Sprache & Kognition*, 4, 203-216.
- Putz-Osterloh, W., Bott, B. & Houben, I. (1988). Beeinflußt Wissen über ein realitätsnahes System dessen Steuerung? *Sprache & Kognition*, 7, 240-251.
- Putz-Osterloh, W. & Lürer, G. (1979). Wann produzieren Probanden räumliche Vorstellungen beim Lösen von Raumvorstellungsaufgaben. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 26, 138-156.

- Putz-Osterloh, W. & Lüer, G. (1981). Über die Vorhersagbarkeit komplexer Problemlöseleistungen durch Ergebnisse in einem Intelligenztest. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 28, 309-334.
- Quesada, J. F., Kintsch, W. & Gomez, E. (2002). A computational theory of complex problem solving using Latent Semantic Analysis. In W. D. Gray & C. D. Schunn (Eds.), *24th annual conference of the Cognitive Science Society* (pp. 750-755). Mahwah, NJ: Lawrence Erlbaum Associates.
- Raichle, M. E. (2001). Cognitive neuroscience. Bold insights. *Nature*, 412, 128-130.
- Reeves, A. & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review*, 93, 180-206.
- Reichert, U. & Dörner, D. (1988). Heuristiken beim Umgang mit einem "einfachen" dynamischen System. *Sprache & Kognition*, 7, 12-24.
- Reiman, E. M., Lane, R. D., Van Petten, C. & Bandettini, P. A. (2000). Positron emission tomography and functional magnetic resonance imaging. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (2nd ed., pp. 85-118). Cambridge: Cambridge University Press.
- Reimann, P. (1998). Novizen- und Expertenwissen. In F. Klix & H. Spada (Eds.), *Wissen (=Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie II: Kognition, Band 6)* (pp. 335-365). Göttingen: Hogrefe.
- Reither, F. (1981). About thinking and acting of experts in complex situations. *Simulation & Games*, 12, 125-140.
- Rhenius, D. & Locher, J. (1992). Auswertungsalgorithmus für Folgen von Augenbewegungen während eines Problemlöseprozesses. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 39, 646-661.
- Richer, F. & Beatty, J. (1985). Pupillary dilations in movement preparation and stimulus probability. *Psychophysiology*, 22, 530-534.
- Rigas, G. (2000). Decision making processes and learning in a simulated ecological system. In N. Georgantzis & I. Barreda (Eds.), *Spatial economics and ecosystems: The interaction between economics and the natural environment* (pp. 323-377). Southampton: WIT Press.
- Ritter, F. E. & Larkin, J. H. (1994). Developing process models as summaries of HCI action sequences. *Human-Computer Interaction*, 9, 345-383.
- Robertson, E. M., Theoret, H. & Pascual-Leone, A. (2003). Studies in cognition: The problems solved and created by transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, 15, 948-960.
- Roland, P. E. & Friberg, L. (1985). Localization of cortical areas activated by thinking. *Journal of Neurophysiology*, 53, 1219-1243.

- Rösler, F. (1996). Methoden der Psychophysiologie. In E. Erdfelder, R. Mausfeld, T. Meiser & G. Rudinger (Eds.), *Handbuch Quantitative Methoden* (pp. 491-514). Weinheim: Psychologie Verlags Union.
- Rößger, P. (1997). *Die Entwicklung der Pupillometrie zu einer Methode der Messung mentaler Beanspruchung in der Arbeitswissenschaft*. Sinzheim: Pro Universitate Verlag.
- Roth, T. (1985). Sprachstatistisch objektivierbare Denkstilunterschiede zwischen „guten“ und „schlechten“ Bearbeitern komplexer Probleme. *Sprache & Kognition*, 4, 178-191.
- Roth, T. (1987). Erfolg bei der Bearbeitung komplexer Probleme und linguistischer Merkmale des Lauten Denkens. *Sprache & Kognition*, 6, 208-220.
- Roth, T., Meyer, H. A. & Lampe, K. (1991). Sprachgebrauch, Informationsstrukturierung und Verhalten in einer komplexen Problemsituation. *Sprache & Kognition*, 10, 28-38.
- Rötting, M. (2001). *Parametersystematik der Augen- und Blickbewegungen für arbeitswissenschaftliche Untersuchungen*. Aachen: Shaker.
- Rowe, J. B., Owen, A. M., Johnsrude, I. S. & Passingham, R. E. (2001). Imaging the mental components of a planning task. *Neuropsychologia*, 39, 315-327.
- Russo, J. & Rosen, L. D. (1975). An eye fixation analysis of multialternative choice. *Memory & Cognition*, 3, 267-276.
- Salvucci, D. D. & Anderson, J. R. (2001). Automated eye-movement protocol analysis. *Human-Computer Interaction*, 16, 39-86.
- Sanders, A. F. (1998). *Elements of human performance: Reaction process and attention in human skill*. Mahwah, NJ: Lawrence Erlbaum.
- Sanderson, P. M. (1989). Verbalizable knowledge and skilled task performance: Association, dissociation, and mental models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 729-747.
- Schaub, H. (2001). *Persönlichkeit und Problemlösen. Persönlichkeitsfaktoren als Parameter eines informationsverarbeitenden Systems*. Weinheim: Beltz PVU.
- Schaub, H. & Strohschneider, S. (1992). Die Auswirkungen unterschiedlicher Problemlöseerfahrung auf den Umgang mit einem unbekanntem komplexen Problem. *Zeitschrift für Arbeits- und Organisationspsychologie*, 36, 117-126.
- Schluroff, M., Zimmermann, T. E., Freeman, R. B., Jr., Hofmeister, K., Lorscheid, T. & Weber, A. (1986). Pupillary responses to syntactic ambiguity of sentences. *Brain & Language*, 27, 322-344.
- Schmid, U. & Kindsmüller, M. C. (1996). *Kognitive Modellierung. Eine Einführung in die logischen und algorithmischen Grundlagen*. Heidelberg: Spektrum Akademischer Verlag.

- Schmuck, P. & Strohschneider, S. (1995). Exekutive Kontrolle und Verhaltensstabilität beim Bearbeiten eines komplexen Problems: Eine Replikation. *Diagnostica*, *41*, 150-171.
- Schnotz, W. (1994). Rekonstruktion von individuellen Wissensstrukturen. In G. L. Huber & H. Mandl (Eds.), *Verbale Daten: Eine Einführung in die Grundlagen und Methoden der Erhebung und Auswertung* (pp. 220-240). Weinheim: Psychologie Verlags Union.
- Schooler, J. W., Ohlsson, S. & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, *122*, 166-183.
- Schroiff, H. W. (1987). Zum Stellenwert von Blickbewegungsdaten bei der Mikroanalyse kognitiver Prozesse. *Zeitschrift für Psychologie*, *195*, 189-208.
- Shepard, R. N. & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701-703.
- Short, E. J., Evans, S. W., Friebert, S. E. & Schatschneider, C. (1991). Thinking aloud during problem solving: Facilitation effects. *Learning and Individual Differences*, *3*, 109-122.
- Simon, H. A. & Barenfeld, M. (1969). Information-processing analysis of perceptual processes in problem solving. *Psychological Review*, *76*, 473-483.
- Simons, D. J. & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*, 261-267.
- Springer, C. S., Patlak, C. S., Pályka, I. & Huang, W. (2000). Principles of susceptibility contrast-based functional MRI: The sign of the functional MRI response. In C. T. W. Monnen & P. A. Bandettini (Eds.), *Functional MRI* (pp. 91-102). Berlin: Springer.
- Stäudel, T. (1987). *Problemlösen, Emotionen und Kompetenz*. Regensburg: Roderer.
- Stäudel, T. (1988). Der Kompetenzfragebogen. Überprüfung eines Verfahrens zur Erfassung der Selbsteinschätzung der heuristischen Kompetenz, belastenden Emotionen und Verhaltenstendenzen beim Lösen komplexer Probleme. *Diagnostica*, *34*, 136-148.
- Sternberg, R. J. (1995). Expertise in complex problem solving: A comparison of alternative conceptions. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 295-321). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sternberg, S. (1966). High speed scanning in human memory. *Science*, *153*, 652-654.
- Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, *57*, 421-457.
- Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science: Methods, models, and conceptual issues* (pp. 703-863). Cambridge, MA: MIT Press.
- Stewart, L., Ellison, A., Walsh, V. & Cowey, A. (2001). The role of transcranial magnetic stimulation (TMS) in studies of vision, attention and cognition. *Acta Psychologica*, *107*, 275-291.

- Stoltze, A. (1991). *Konstruktion eines neuropsychologischen Tests zur Messung von Planungsfähigkeiten nach Frontalhirnschädigung*. Unpublished Diplomarbeit, Fachbereich Psychologie der Universität Konstanz, Konstanz.
- Strauß, B. & Kleinmann, M. (Eds.). (1995). *Computersimulierte Szenarien in der Personalarbeit*. Göttingen: Verlag für Angewandte Psychologie.
- Strohschneider, S. (1986). Zur Stabilität und Validität von Handeln in komplexen Realitätsbereichen. *Sprache & Kognition*, 5, 42-48.
- Strohschneider, S. (1994). Strategien beim Umgang mit einem komplexen Problem: Ein deutsch-deutscher Vergleich. *Zeitschrift für Arbeits- und Organisationspsychologie*, 38, 34-40.
- Strohschneider, S. & Güss, D. (1999). The fate of the Moros: A cross-cultural exploration of strategies in complex and dynamic decision making. *International Journal of Psychology*, 34, 235-252.
- Strube, G. (1990). Neokonnektionismus: Eine neue Basis für die Theorie und Modellierung menschlicher Kognition? *Psychologische Rundschau*, 41, 129-143.
- Stuss, D. T. & Alexander, M. P. (2000). Executive functions and the frontal lobes: A conceptual view. *Psychological Research*, 63, 289-298.
- Suppes, P., Cohen, M., Laddaga, R., Anliker, J. & Floyd, R. (1982). Research on eye movements in arithmetic performance. In R. Groner & P. Fraisse (Eds.), *Cognition and eye movements* (pp. 57-73). New York: Elsevier.
- Süß, H.-M. (1999). Intelligenz und komplexes Problemlösen - Perspektiven für eine Kooperation zwischen differentiell-psychometrischer und kognitionspsychologischer Forschung. *Psychologische Rundschau*, 50, 220-228.
- Süß, H.-M. & Faulhaber, J. (1990). Berliner Version der Schneiderwerkstatt. PC Simulationsprogramm. Berlin: Freie Universität Berlin, Fachbereich Erziehungs- und Unterrichtswissenschaften, Institut für Psychologie.
- Taatgen, N. A. & Wallach, D. (2002). Whether skill acquisition is rule or instance based is determined by the structure of the task. *Cognitive Science Quarterly*, 2, 163-204.
- Thalmaier, A. (1979). Zur kognitiven Bewältigung der optimalen Steuerung eines dynamischen Systems. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 26, 388-421.
- Tisdale, T. (1990). Zur Bedeutung selbstreflexiver Prozesse beim Problemlösen. In D. Frey (Ed.), *Bericht über den 37. Kongreß der Deutschen Gesellschaft für Psychologie in Kiel 1990. Band 1: Kurzfassungen* (pp. 151). Göttingen: Hogrefe.
- Toga, A. W. & Mazziotta, J. C. (2002). *Brain mapping: The methods* (2nd ed. ed.). New York: Academic Press.
- Tomasello, M. (2002). *Die kulturelle Entwicklung des menschlichen Denkens*. Frankfurt: Suhrkamp.

- Uttal, W. R. (2001). *The new phrenology*. Cambridge, MA: MIT Press.
- van den Heuvel, O. A., Groenewegen, H. J., Barkhof, F., Lazeron, R. H., van Dyck, R. & Veltman, D. J. (2003). Frontostriatal system in planning complexity: A parametric functional magnetic resonance version of Tower of London task. *Neuroimage*, *18*, 367-374.
- Vester, F. (1983). *Ökolopoly. Umweltspiel*. Ravensburg: Ravensburger Spiele.
- Viviani, P. (1990). Eye movements in visual search: Cognitive, perceptual and motor control aspects. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 353-393). Amsterdam: Elsevier.
- Vollmeyer, R. & Funke, J. (1999). Personen- und Aufgabenmerkmale beim komplexen Problemlösen. *Psychologische Rundschau*, *50*, 213-219.
- Volz, K. G., Schubotz, R. I. & von Cramon, D. Y. (2003). Predicting events of varying probability: Uncertainty investigated by fMRI. *Neuroimage*, *19*, 271-280.
- von Cramon, D. Y., Matthes-von Cramon, G. & Mai, N. (1991). Problem-solving deficits in brain-injured patients: A therapeutic approach. *Neuropsychological Rehabilitation*, *1*, 45-64.
- Wagener, D. (2001). Personalauswahl und -Entwicklung mit komplexen Szenarios. *Wirtschaftspsychologie*, *3*, 69-76.
- Wagner, A. C., Uttendorfer-Marek, I. & Weidle, R. (1977). Die Analyse von Unterrichtsstrategien mit der Methode des "Nachträglichen Lauten Denkens" von Lehrern und Schülern zu ihrem unterrichtlichen Handeln. *Unterrichtswissenschaft*, *5*, 244-250.
- Wallach, D. (1998). *Komplexe Regelungsprozesse. Eine kognitionswissenschaftliche Analyse*. Wiesbaden: Deutscher Universitäts-Verlag.
- Walsh, V. (1998). Brain mapping: Faradization of the mind. *Current Biology*, *8*, 8-11.
- Weidle, R. & Wagner, A. C. (1994). Die Methode des lauten Denkens. In G. L. Huber & H. Mandl (Eds.), *Verbale Daten* (pp. 81-103). Weinheim: Psychologie Verlags Union.
- White, G. L. & Maltzman, I. (1978). Pupillary activity while listening to verbal passages. *Journal of Research in Personality*, *12*, 361-369.
- Williamson, J., Ranyard, R. & Cuthbert, L. (2000). A conversation-based process tracing method for use with naturalistic decisions: An evaluation study. *British Journal of Psychology*, *91*, 203-221.
- Wundt, W. (1907). Über Ausfrageexperimente und über die Methoden zur Psychologie des Denkens. *Psychologische Studien*, *3*, 301-360.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.

Zelinsky, G. J. & Loschky, L. (1998). Toward a realistic assessment of visual working memory. *Investigative Ophthalmology & Visual Science*, 39, 224.

**Autorenfussnote**

Für hilfreiche Kommentare zu einer früheren Textfassung bedanken wir uns bei Timm Lochmann. Michael Öllinger hat durch gründliche Lektüre geholfen, zahlreiche Fehler zu beseitigen und Unverständlichkeiten zu reduzieren – ganz herzlichen Dank dafür.

**Tabellen**

Tabelle NEWELL: Zeitskala menschlicher Handlung nach Newell (1990), erweitert durch Ballard et al. (1997).

<b>Abstraktionslevel</b>	<b>Zeitskala</b>	<b>System</b>	<b>Beispiel</b>
Sozial	Monate	Projekt, Lebensbereich	Planung und Koordination, Lernen und Erwerb kognitiver Fähigkeiten
	Wochen		
	Tage		
Rational	Stunden	Problem	Lösen eines komplexen Problems
	10 min	Aufgabe	Lösen einer einfachen Aufgabe
	1 min	Aufgabe	
Kognitiv	10 sek	Aufgabeneinheit	Wählen einer Telefonnummer
Verkörperung (embodiment)	200-300 msek	Körperbewegung	Augen-, Finger- oder Handbewegung
Aufmerksamkeit	50 msek	Einfachste geplante Handlung	Wahrnehmen eines Reizes
Neuronal	10 msek	Neuronaler Kreislauf	Laterale Inhibition
Neuronal	1 msek	Neuronale Feuerung	Einfaches Signal

Tabelle MESS: Überblick über die gängigen Verfahren zur Messung von Augenbewegungen.

Messverfahren	Anatomisch-physiologische Eigenschaften des Auges	Charakteristika der Messverfahren	Vorteile	Nachteile
EOG	Corneo-retinales Potenzial	<ul style="list-style-type: none"> <li>- Genauigkeit: 1.5° (H), 2° (V)</li> <li>- Messbereich: 30°</li> <li>- Zeitliche Auflösung: wird durch Registriersystem (Abtastrate der A/D-Wandlung) bestimmt</li> </ul>	<ul style="list-style-type: none"> <li>- Aufzeichnung der Augenbewegungen bei Brillen- und Kontaktlinsenträgern möglich</li> <li>- Keine Einschränkung des Blickfelds, keine Beeinträchtigung des Pb</li> </ul>	<ul style="list-style-type: none"> <li>- Tageszeitliche Schwankungen des Potenzials</li> <li>- Lidschlag kann vertikale Messung beeinträchtigen</li> <li>- Überlagerung durch Muskelpotenzials der Gesichtsmuskulatur</li> <li>- elektrisches Rauschen, Drift des Messsystems</li> </ul>
Cornea-Reflex-Methode	Reflexionen	<ul style="list-style-type: none"> <li>- Genauigkeit: 0.5-1°</li> <li>- Messbereich</li> <li>- Zeitliche Auflösung: meist 250-500 Hz</li> </ul>		<ul style="list-style-type: none"> <li>- Anfälligkeit gegenüber Körper-, insbesondere Kopfbewegungen (z.B. Stirnrunzeln), daher Kinnstütze erforderlich</li> </ul>
DPI	Reflexionen	<ul style="list-style-type: none"> <li>- Genauigkeit</li> <li>- Messbereich</li> <li>- Zeitliche Auflösung: 500-1000 Hz</li> </ul>	<ul style="list-style-type: none"> <li>- Sehr genaues, hoch auflösendes Verfahren</li> </ul>	<ul style="list-style-type: none"> <li>- Aufwändige Messung, erfordert viel Übung für den Versuchsleiter und geübte Pbn</li> <li>- Beissbrett erforderlich, kann unangenehm sein</li> </ul>
Kontaktlinsenmethode (scleral search coil)	Krümmung der Cornea	<ul style="list-style-type: none"> <li>- Genauigkeit</li> <li>- Messbereich</li> <li>- Auflösung</li> </ul>	<ul style="list-style-type: none"> <li>- Sehr genaues, hoch auflösendes Verfahren</li> </ul>	<ul style="list-style-type: none"> <li>- Erheblicher Vorbereitungsaufwand</li> <li>- Hohe Kosten</li> <li>- Beeinträchtigung und Verletzungsgefahr des Pb</li> <li>- Kurze Tragedauer (bis 30 Minuten)</li> </ul>

Tabelle FINAL: Zuordnung von Verfahren auf den drei diagnostischen Ebenen zu den verschiedenen Zeitspannen nach Newell (1990) bzw. Anderson (2002).

Zeitspanne (Sek)	subjektive Daten	Verhaltensdaten	physiologische Daten
„Biologisches Band“: $10^{-4}$ bis $10^{-2}$ (100 $\mu$ s bis 10 msek)			Elektrophysiologische Methoden (z.B. EEG), TMS
„Kognitives Band“: $10^{-1}$ bis $10^1$ (100 msek bis 10 sek)		Blickbewegungen (Sakkaden, Fixationen), Reaktionszeiten, Mimik & Gestik	Pupillometrie, bildgebende Verfahren (z.B. fMRI)
„Rationales Band“: $10^2$ bis $10^4$ (Minuten, Stunden)	Laut-Denken, Tests, Wissensdiagnostik	Lösungsgüte und Verhaltensprotokolle bei Denksport- aufgaben (z.B. TOH), kryptarithm. Probl., Einsichtsproblemen, Szenarios	
„Soziales Band“: $10^5$ bis $10^7$ (Tage, Wochen, Monate)	Interviews	Lösungsgüte und Protokolle bei Planspielen, Szenarios	

**Abbildungen**

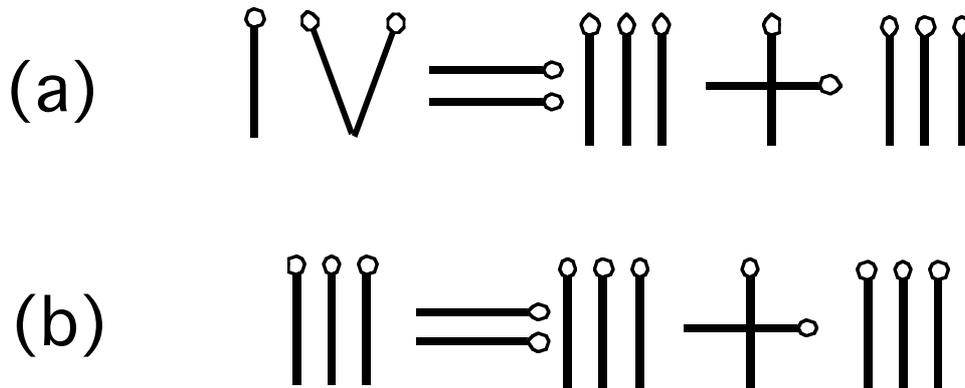


Abbildung STREICH: Beispiel für „Streichholz-Arithmetik“ (nach Knoblich et al., 1999).



Abbildung KRYPT: Zwei kryptarithmetische Problemstellungen mit siebenbuchstabigen Zielwörtern: Welche Ziffer entspricht welchem Buchstaben?

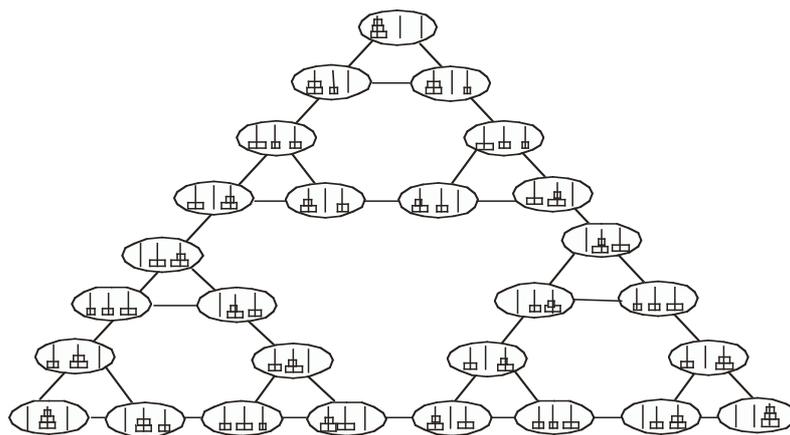


Abbildung WEG: Der Problemraum des „Turm von Hanoi“ für 3 Scheiben. Dargestellt sind alle zulässigen Konfigurationen der drei Scheiben auf den drei Stäben. Der kürzeste Weg verläuft in sieben Schritten an der Aussenkante von oben nach rechts unten. Oben befinden sich die drei Scheiben der Größe nach geordnet auf dem linken Stab (= Ausgangszustand), rechts unten geordnet auf dem rechten Stab (= Zielzustand).

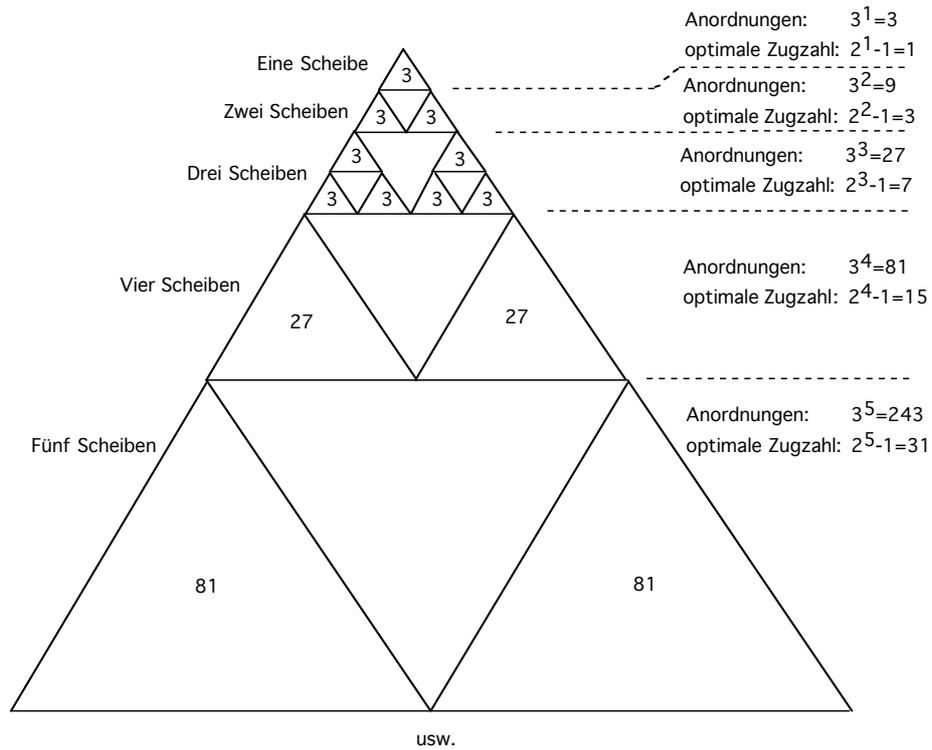


Abbildung TVH\_LÖS: Abstrakte Darstellung des Problemraums beim „Turm von Hanoi“. Deutlich sichtbar ist die Symmetrie des Problemraums, dessen Größe mit jeder weiteren Scheibe um eine Dreier-Potenz wächst.

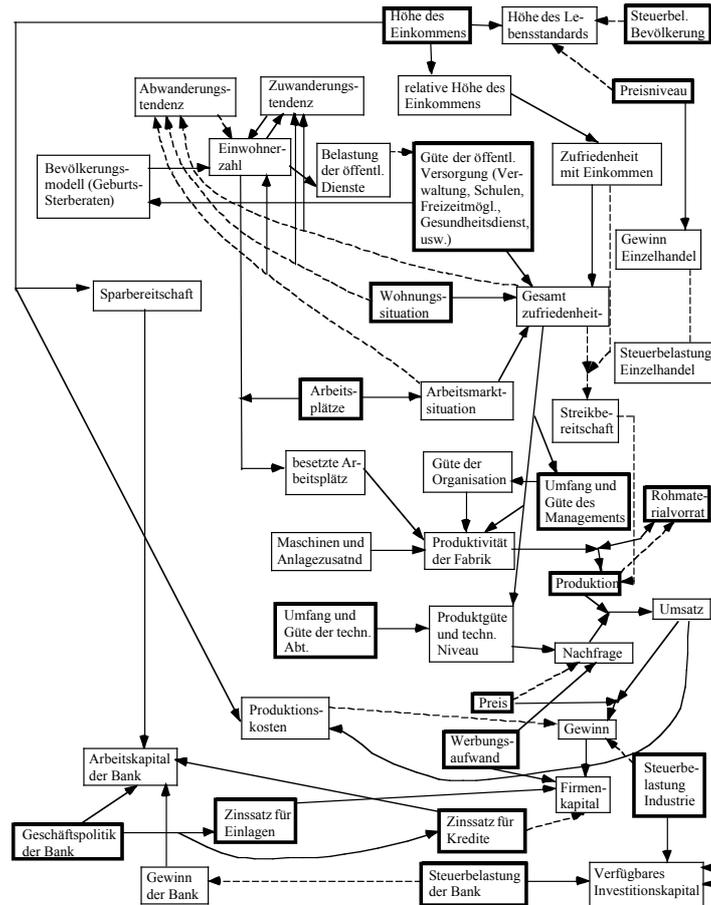


Abbildung LOH: Grobstruktur des LOHHAUSEN-Systems (nach Dörner et al., 1983, S. 110): 44 der rund 2000 Variablen. Fett markiert: Eingriffsmöglichkeiten der Versuchspersonen; gestrichelt: negative Wirkungen.

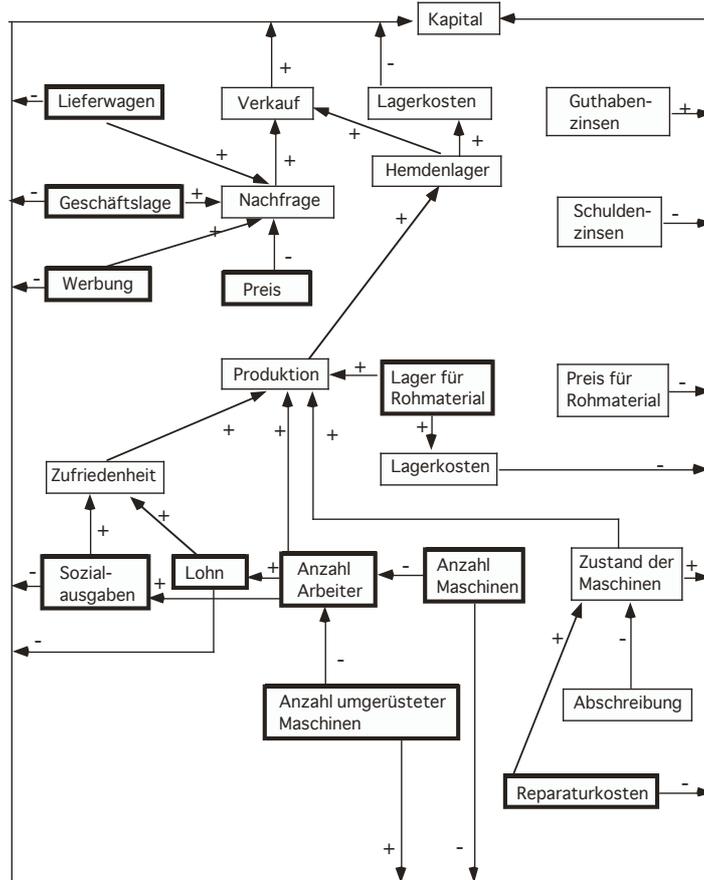


Abbildung SCHNEI: Die Variablen des „Schneiderwerkstatt“-Szenarios und ihre Verknüpfungen. Die dick umrandeten Variablen sind direkt beeinflussbar (N=11). Ein „+“ bedeutet eine positive, ein „-“ eine negative Beziehung (nach Putz-Osterloh, 1981b, S. 85).

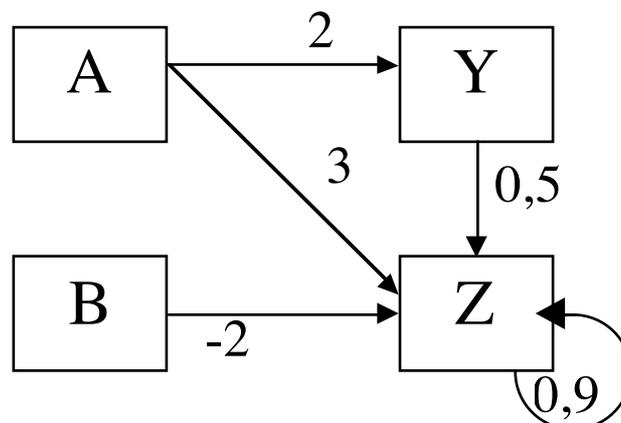


Abbildung LIN\_SYS: Struktur eines einfachen linearen Systems mit zwei exogenen Variablen A und B sowie zwei endogenen Variablen Y und Z, die gemäß der angegebenen gerichteten Kanten (= kausale Verbindungen zwischen den Variablen) und deren Gewichte untereinander verknüpft sind.

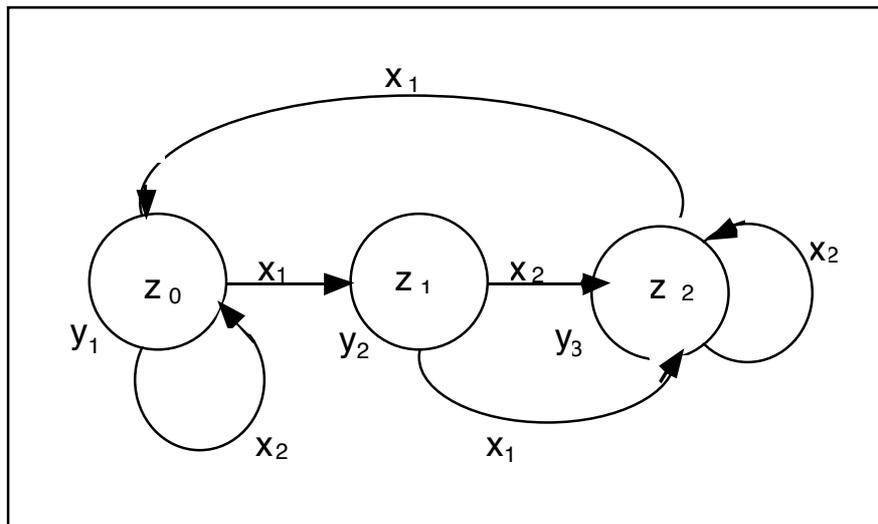


Abbildung FSA: Graphische Repräsentation des finiten Automaten mit drei Zuständen ( $z_0$ ,  $z_1$  und  $z_2$ ) und zwei Eingabemöglichkeiten ( $x_1$ ,  $x_2$ ). Die Ausgabebezeichnungen ( $y_1$ ,  $y_2$ ,  $y_3$ ) sind links neben die zugehörigen Zustandsknoten gesetzt.

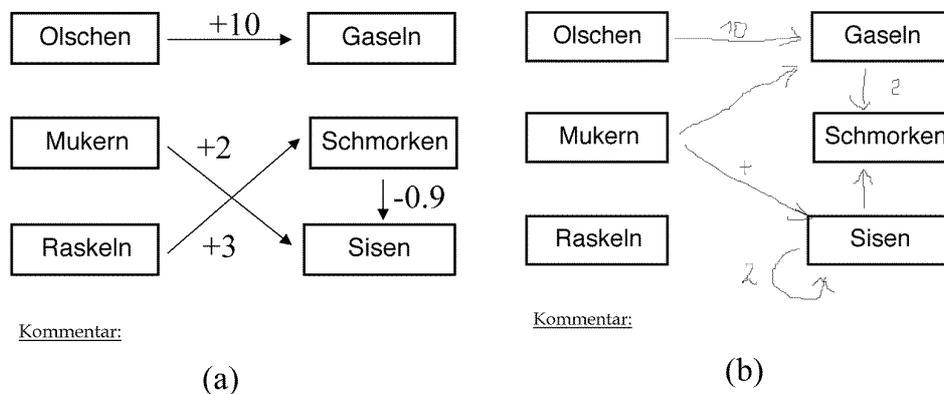


Abbildung KDD: Beispiel für Kausaldiagramm-Diagnostik (KDD): Tatsächlich implementierte Struktur (links) und subjektives Modell einer Vp (rechts).

Aus 8 gleich aussehenden Kugeln ist durch 2maliges Wägen diejenige Kugel herauszufinden, die etwas leichter als die übrigen 7 ist. Wieviel Kugeln müssen jeweils in einer Waagschale sein?

- Bei der ersten Wägung sind in jeder Waagschale 3 Kugeln. (richtige Antwort)
- Bei der ersten Wägung sind in jeder Waagschale 4 Kugeln.
- Bei der ersten Wägung sind in jeder Waagschale 2 Kugeln.
- Bei der ersten Wägung ist in jeder Waagschale 1 Kugel.
- Bei zweimaligem Wiegen ist kein Unterschied festzustellen.

Abbildung DST\_ITEM: Beispiel eines Items aus dem DST von Lienert (1964).

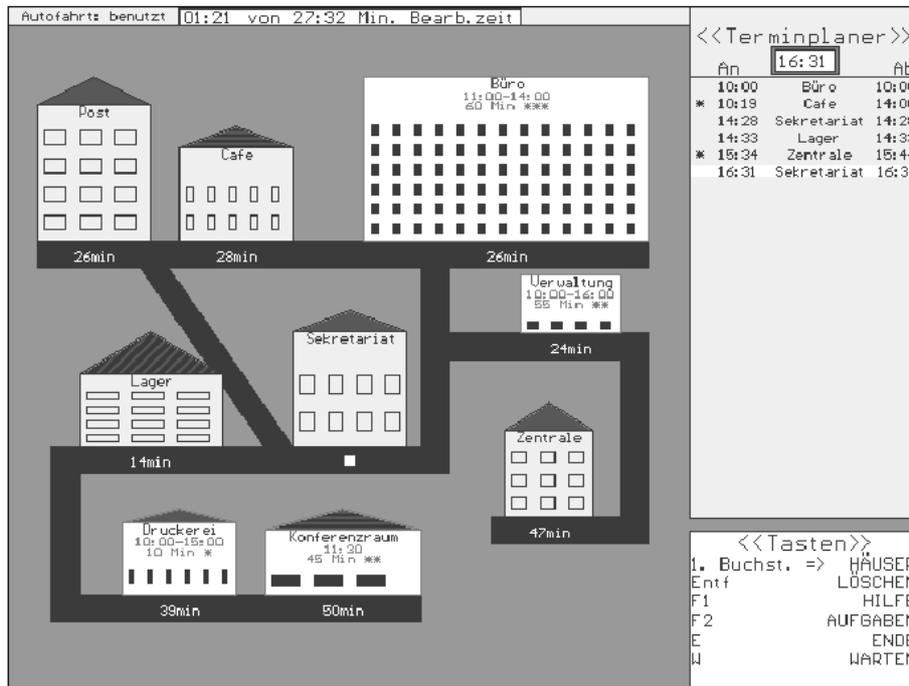


Abbildung PAD: Benutzeroberfläche der Planungsaufgabe „Plan A Day“ von Funke und Krüger (1995). Im linken Teil sieht man den Lageplan und die an bestimmten Orten noch zu erledigenden Aufgaben, rechts sind der bislang erstellte Plan (oben) und die Bedienelemente (unten) zu sehen.

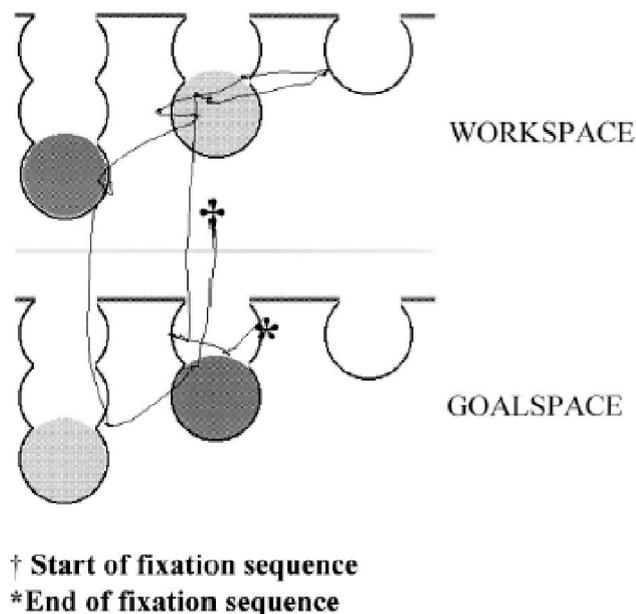


Abbildung TVL: X-Y Blickpositionen für einen Versuchsdurchgang. Die Instruktion der Versuchsperson lautete, das obere Muster („Workspace“) dem unteren („Goalspace“) anzugleichen (aus Timothy L. Hodgson, A. Bajwa, A. M. Owen & Christopher Kennard, 2000b, S. 895).

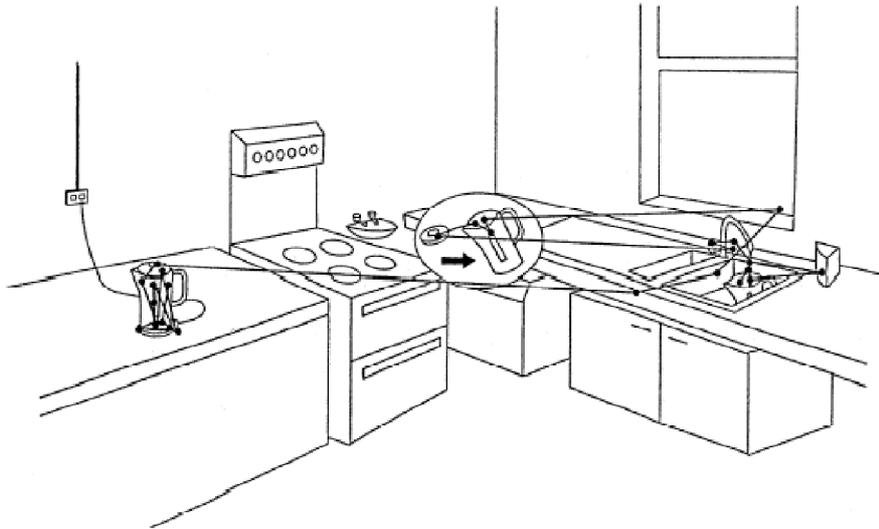


Abbildung TEA: Sakkaden ( $> 1^\circ$ ) und Fixationen beim Teekochen innerhalb einer 10 sek Sequenz (aus Land & Hayhoe, 2001, p. 3560).

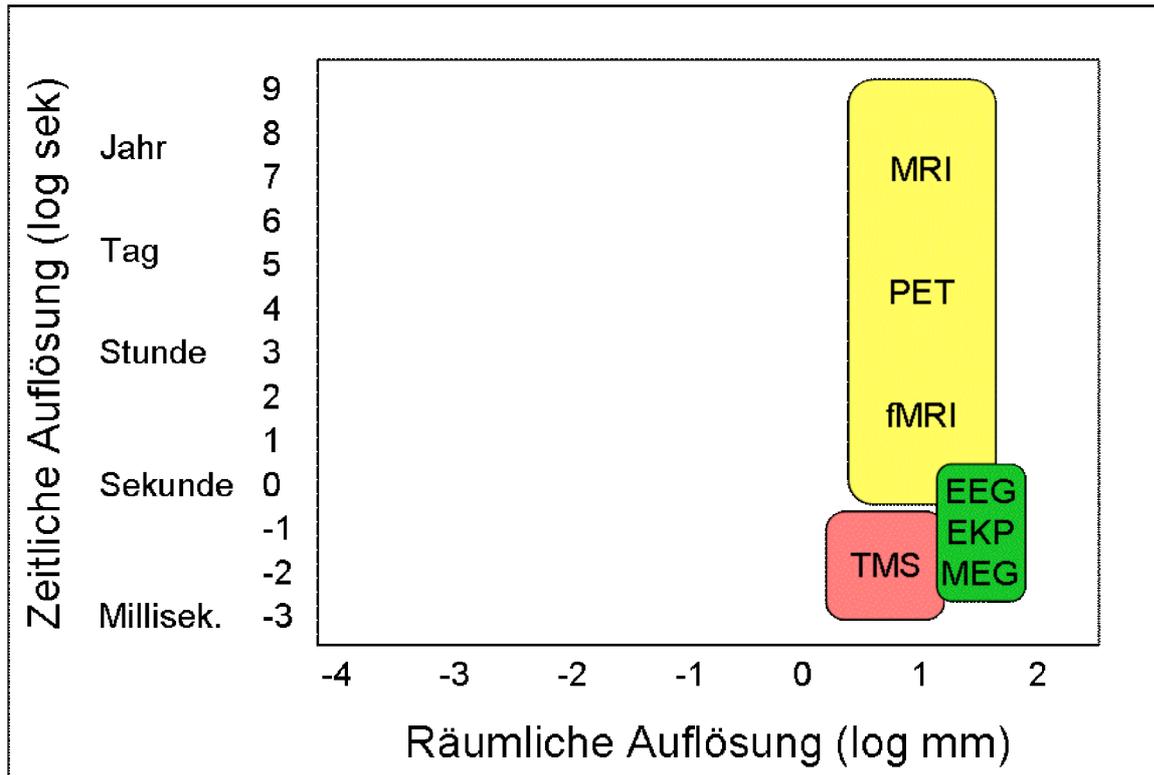


Abbildung RESOLUT: Zeitliche und räumliche Auflösung ausgewählter bildgebender Verfahren (nach Walsh, 1998, p. 9).

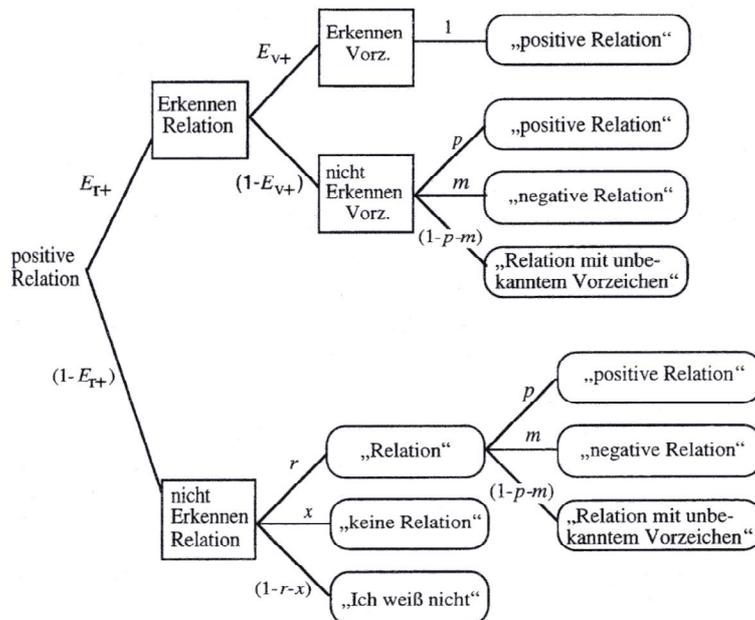


Abbildung MUL: Ereignisbaum zur Beschreibung der vermuteten Prozesse beim Erkennen einer positiven Relation in einem Kausaldiagramm und der Konsequenzen auf das Antwortverhalten (aus Beckmann, 1994, S. 162).

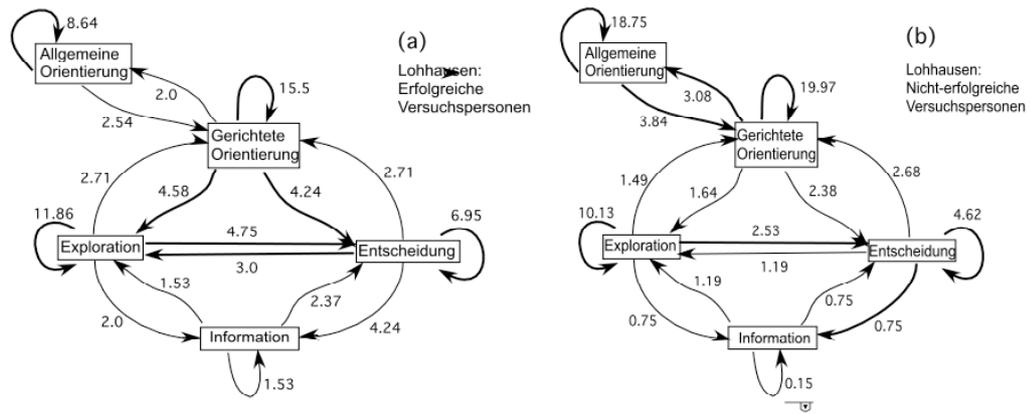


Abbildung MARKOV: Übergänge in Prozent zwischen fünf verschiedenen Arten der Informationsverarbeitung aus der „Lohhausen“-Studie: (a) erfolgreiche Probanden, (b) weniger erfolgreiche Probanden (nach Dörner & Wearing, 1995, p. 75).