

1. Kapitel

Ziele und Aufgaben einer Allgemeinen Methodenlehre der Psychologie

Werner H. Tack

Was verbirgt sich hinter der Vorstellung einer „Allgemeinen Methodenlehre“, die unbeschadet ihrer Allgemeinheit doch eine solche der Psychologie ist, und deren Existenz - oder zumindest deren Möglichkeit - der Titel dieses Kapitels voraussetzt? Sicher kann diese Frage hier nicht voll zufriedenstellend beantwortet werden; erste Versuche einer Explikation, Prolegomena also, sind aber möglich und notwendig. Dieser Beitrag eröffnet in einer Enzyklopädie der Psychologie innerhalb einer Serie über Forschungsmethoden einen Band, der sich mit **methodologischen Grundlagen** beschäftigt. Zur gleichen Serie gehören Bände über Klassen von Methoden, die spezifischen Teilaufgaben im Forschungsprozeß dienen (Datenerhebung, Messen und Testen, Strukturierung und Reduzierung von Daten, Hypothesenprüfung). Bezeichnen wir die dort vorherrschenden Beschäftigungen mit mehr oder weniger eng abgegrenzten Gruppen von Methoden und Techniken als „spezielle“ Methodenlehren, so ist die im Titel dieses Bandes angesprochene Behandlung methodologischer Grundlagen das „allgemeine“ Pendant. Die Frage nach Inhalt und Methodik einer Allgemeinen Methodenlehre der Psychologie ist somit gleichbedeutend mit der Frage, was unter „methodologischen Grundlagen“ zu verstehen und mithin in den weiteren Kapiteln dieses Bandes zu erwarten ist.

1. Versuch einer Abgrenzung

Eine erste Annäherung an die Konzeption einer Allgemeinen Methodenlehre der Psychologie liefern Überlegungen zu Fragen oder Themen, die ihr zuzuordnen sind. Damit sei keineswegs unterstellt, Wissenschaften und deren Teildisziplinen bedürften zu ihrer Abgrenzung eines jeweils spezifischen Gegenstandes. Unbeschadet der Frage nach der **Notwendigkeit einer Spezifikation** durch Gegenstandsangabe gibt es aber die **Möglichkeit der Explikation** der

Konzeption einer Wissenschaft oder eines Teilgebietes durch Verweis auf Gegenstände, mit denen sie sich beschäftigt.

Wir gehen also nicht von der Vorstellung aus, „Allgemeine Methodenlehre der Psychologie“ sei abschließend **definierbar**, sondern lediglich davon, sie sei durch Verweis auf Gegenstände, mit denen sie sich befaßt, **explizierbar**. Dabei können die Gegenstände mehr oder weniger typisch sein. Zur Annäherung an den weder hinreichend klar bekannten noch beschreibbaren Prototypen (einen fiktiven „zentralen“ Gegenstand Allgemeiner Methodenlehre bzw. einer „zentralen“ methodologischen Grundlage) wird durch Rückgriff auf möglichst typische Gegenstände vielleicht der Eindruck einer sehr engen Konzeption erweckt. Diese thematische Enge dient lediglich der Verdeutlichung; wie breit das Gesamtkonzept in der Realität ist, das zeigt die Vielfalt der Themen, Fragestellungen und Ansätze in den nachfolgenden Kapiteln dieses Bandes.

Sind es bei Herrmann im Kapitel 6 dieses Bandes „Forschungsprogramme“, mit denen sich eine Wissenschaft - und mithin auch die Psychologie - beschäftigt, so spielen in diesem Kapitel „Gegenstände“ (der Allgemeinen Methodenlehre) eine analoge Rolle. Werden Programme der Psychologie dort durch „Probleme“ angebbar, so sollen hier Gegenstände einer Allgemeinen Methodenlehre durch „Fragestellungen“ - nämlich durch „methodologische Grundfragen“ beschrieben werden.

„Methodologie“ - das sei vorab festgehalten - ist **wissenschaftliche Beschäftigung mit Methoden**, die hier also nicht lediglich in der Forschung eingesetzt werden, sondern selbst Objekt der Forschung sind. Soll sich wissenschaftliche Tätigkeit nicht mit der Angabe isolierter Befunde bescheiden, sondern darüber hinaus Ergebnisse systematisieren (siehe hierzu Gadene im einführenden Abschnitt zu Kap. 7 dieses Bandes), so muß auch „Methodologie“ nicht lediglich Methoden darstellen, sondern diese analysieren und über sie theoretisieren.

1.1 Allgemeine vs. spezifische Methodenlehre

Zur Spezifikation einer Allgemeinen Methodenlehre der Psychologie ist eine doppelte Abgrenzung sinnvoll. Auf der einen Seite ist zwischen „Allgemeiner“ und „Spezifischer“ Methodenlehre zu differenzieren, auf der anderen Seite zwischen Allgemeiner Methodenlehre „der Psychologie“ und der einer beliebigen anderen Wissenschaft oder sogar der Wissenschaft allgemein. Die erste dieser beiden Unterscheidungen ist relativ einfach, eine Methodenlehre ist **spezifisch**, wenn sie sich mit jeweils einer bestimmten Methode (wie etwa dem Experiment, dem statistischen Schließen, der Messung, der Psychodiagnostik, der Faktorenanalyse, dem hierarchischen Clustering oder ähnlichem) beschäftigt.

Diese Abgrenzung hat zwei wichtige Konsequenzen. Da jede Entwicklung und Darstellung einer Methode spezifisch für eben diese Methode ist, gehören Methodenentwicklung und -darstellung **nicht** zum Arbeitsbereich einer Allgemeinen Methodenlehre. Was bleibt ist die Methoden-Systematisierung und -Analyse, also die **Methodologie**. Zum zweiten folgt, daß im Zentrum einer Allgemeinen Methodenlehre Fragen stehen, die sinnvoll für **verschiedene** Methodenarten gestellt werden können. Derartige Fragen gibt es in größerer Zahl, als ein Blick in die meisten einschlägigen Methoden-Lehrbücher der Psychologie vermuten läßt.

Es wird oft gesagt, es bestehe ein enger Zusammenhang zwischen Forschungsfragen und -methoden; wie aber könnte dieser genauer untersucht werden als durch Vergleich eines breiten Spektrums von Fragestellungen mit einem angemessenen Vorrat unterschiedlicher Methoden? Methoden unterscheiden sich bezüglich der (oft impliziten) Vorannahmen, die man unterstellt, um ihren Einsatz rechtfertigen zu können; wie anders als durch Methodenvergleich soll man die Unterschiedlichkeit solcher Vorannahmen analysieren? Die Möglichkeit, Forschungsergebnisse unterschiedlich zu formulieren, ist nicht unabhängig vom methodenbedingten Ausmaß der Ergebnisspezifikation; auch für Arbeiten zur Bedeutsamkeit von Befunden ist ein Ansatz erforderlich, der über Methodenklassen hinausgreift. Dies sind nur beispielhafte Hinweise, die zeigen, daß der Gegenstandsbereich einer Allgemeinen Methodenlehre alles andere als leer ist.

1.2 Psychologiespezifität

Was macht eine Allgemeine Methodenlehre zu einer solchen **der Psychologie**? Sicherlich eine vorrangige Berücksichtigung jener Fragen und Fragen-Varianten, die für psychologische Forschung besonders wichtig sind. Was das im einzelnen ist, hängt zum großen Teil davon ab, was man für Psychologie hält. Tseëlon (1991) vergleicht ein Methodenarsenal mit einer Sprache, in der man eine vorgegebene Realität nicht lediglich **beschreiben** kann, sondern in der man aufgrund der gegebenen Möglichkeiten Realität stets auch **konstruiert**. Methoden sind keine neutralen Instrumente, durch die hindurch eine objektive Realität mehr oder weniger so erscheint, wie sie ist. Ist die Verbindung zwischen Methoden und Ergebnissen derart eng, dann bestimmt psychologische Forschungsmethodik weitgehend, was als Psychologie geschieht, erfassbar und darstellbar ist. Es ist damit eine methodologische Aufgabe, sich um methodenbedingte Grenzen dessen zu kümmern, was im Fach als Forschungsergebnis möglich ist.

Wenn im folgenden von „Spezifika der Psychologie“ die Rede ist, dann sei damit keineswegs unterstellt, es gebe Eigenschaftsangaben oder - anders for-

muliert - Aussagen, die ausschließlich auf Psychologie zutreffen. Gemeint ist lediglich, daß bestimmte Merkmale des Faches von besonderer Bedeutung für die Entwicklung seines Methodenkanons sind und (vielleicht) in ihrer Kombination ein „Spezifikum“ der Disziplin im engeren Sinne darstellen.

Zum Selbstverständnis der Psychologie gehört zunächst, daß sie eine **Wissenschaft** ist. Akzeptable Methoden müssen mithin **wissenschaftliche Methoden** sein. Man muß nicht unbedingt auf Feyerabend (1976, 1980) verweisen, um zu sehen, daß aktuelle wissenschaftstheoretische Positionen sehr vielfältig sind. Will man dieser Vielfalt gerecht werden, so bleibt nur - als kleinster gemeinsamer Nenner - die Feststellung, daß Wissenschaft ein soziales Unternehmen ist, das die Kontrolle seiner (inhaltlichen) Aussagen inkorporiert hat.

Agnew und Pyke (1969) verdeutlichen dies, indem sie Wissenschaft mit Journalismus vergleichen. Beide berichten über etwas, das sie herausgefunden haben. Dem Journalisten gestehen wir häufig nicht nur bei Leitartikeln, Theaterkritiken und anderen Arten von Kommentaren einen relativ großen Spielraum für Subjektivität zu. Wird dieser Spielraum überzogen, so sprechen wir zwar von schlechtem Journalismus; es ist aber ungewöhnlich, etwa in einem Zeitungsbericht Angaben über das Vorgehen des Berichterstatters bei der Erkenntnisgewinnung zu erwarten, die so detailliert sind, daß andere Personen die Richtigkeit seiner Aussagen überprüfen oder - bei nicht wiederholbaren Ereignissen - aufgrund einer Analyse des Vorgehens bewerten können. Einem Sportreporter geben wir zumindest gelegentlich auch das Recht, in einem Fußballspiel ein Abseits zu sehen, das Schiedsrichter, Linienrichter und 20000 Zuschauer nicht gesehen haben; er ist nicht verpflichtet, möglichst genau mitzuteilen, wie er zu seiner Wahrnehmung gekommen ist. Genau diese Verpflichtung, die **Kontrolle von Aussagen** durch andere erst ermöglicht, gilt aber für **wissenschaftliche** Aussagen.

Die Konsequenzen für ein Unternehmen, das sich - wie die Psychologie - als Wissenschaft begreift, sind offensichtlich. Zum ersten müssen Ergebnisse in einer mehreren Menschen zugänglichen Form - als Bericht objektiviert - vorliegen. Private Erkenntnis ist noch nicht Wissenschaft; Methoden müssen Handwerkszeug für die Herstellung von Berichten enthalten. Zum zweiten ist es erforderlich, nicht nur Ergebnisse mitzuteilen, sondern auch Informationen über den Weg zu vermitteln, auf dem sie gewonnen wurden. Es muß ein Methodenarsenal geben, auf dessen jeweils benutzte Techniken verwiesen werden kann; die in diesem Arsenal verfügbaren Methoden müssen so präzise dargestellt sein, daß sie bis auf unwesentliche Variation von verschiedenen Personen eingesetzt werden können. Hier ist eines der Aufgabenfelder der spezifischen Methodenlehren, die auch untersuchen, was im jeweiligen Kontext als wesentliche Änderung einer Methode und was als unwesentliche Va-

riante gilt. Ein mögliches Kriterium hierfür ist die Frage, welche Abweichungen den Anwendungsbereich der resultierenden Ergebnisse affizieren.

Soviel zur **Wissenschaftlichkeit** der Psychologie und den daraus resultierenden Minimalforderungen an ihre Methoden. Von den zahlreichen Spezifika des Faches seien nun im weiteren zwei herausgegriffen, deren methodologische Konsequenzen nicht nur an vielen Stellen in diesem Buch, sondern auch in anderen Kontexten immer wieder eine Rolle spielen. Das erste ist die Aussage, Psychologie sei eine **empirische** Wissenschaft, das zweite die Behauptung, ihr Gegenstand sei von **hoher Komplexität**.

2. Psychologie als empirische Wissenschaft

Was macht Psychologie zu einer **empirischen** Wissenschaft? Von de Groot (1969, S. 1) stammt die Aussage, eine Wissenschaft sei empirisch, wenn sie Wissen über die Welt bereitstellt, also über die Realität, in der wir leben. Dabei entstehen Theorien, die nach Newell (1990, S. 13) als Wissensbestände angesehen werden können, aus denen man Antworten auf Fragen erhalten kann. Sowohl Vorhersagen als auch Erklärungen oder Handlungsanweisungen sind gleichermaßen mögliche Varianten von Antworten. Die Tatsache, daß damit auch die Wissenschaften selbst als organisierte Wissensbestände interpretierbar sind, bezeichnet bereits Nagel (1961) als etwas, was „zweifelsfrei der Fall ist“ (S. 3, vom Verf. übersetzt). Die Aussagen von de Groot, Newell und Nagel lassen sich problemlos kombinieren, wenn man als Ziel einer **empirischen** Wissenschaft die Bereitstellung von Wissensbeständen ansieht, aus denen man Antworten über die **empirisch erfahrbare Welt** gewinnen kann.

2.1 Empirisch, formal, normativ

Fragt man, welches denn Wissenschaften sind, die auch in einem sehr liberalen Empiriekonzept als „nicht empirisch“ zu gelten haben, so gerät man in Schwierigkeiten. Betrachten wir ein **formales axiomatisches System**. Popper (1966) weist darauf hin, daß ein solches System auf zweifache Weise interpretiert werden kann: „Man kann die Axiome [i] als Festsetzungen betrachten oder [ii] als empirisch-wissenschaftliche Hypothesen“ (S. 42). Faßt man sie als Festsetzungen auf, dann legen sie fest, wie die in ihnen auftretenden Begriffe zu benutzen sind; sie sind implizite Definitionen eben dieser Begriffe. Da diese Begriffe als Variablen angesehen werden können, hat das System selbst den Charakter einer Struktur von Aussagenfunktionen. Es sei nun vereinbart, für den Vektor aller Variablen eines solchen Systems nur Wertekombinationen einzusetzen, bei denen die Aussagen des Axiomensystems gelten. Eine solche

Wertekombination ist ein „Modell“ des Systems. Setzt man nur Modelle ein, dann sind die resultierenden Aussagen in jedem Fall wahr, denn sonst wäre das, was man eingesetzt hat, kein Modell. Aus einem als Festsetzung interpretierten Axiomensystem läßt sich also eine Menge von Aussagen gewinnen, die unbeschadet aller Gegebenheiten in der uns umgebenden Welt stets wahr sind (Popper, 1966, S.42-43). Eine Wissenschaft, die lediglich einen Wissensbestand dieser Art bereitstellt, ist sicher nicht empirisch und soll zunächst als „formal“ gekennzeichnet werden.

2.1.1 Reine und angewandte Disziplin

Die angesprochene doppelte Interpretationsmöglichkeit axiomatischer Systeme führt dazu, daß es zu jeder **formalen Wissenschaft** ein **empirisches Pendant** gibt, das die gleichen Systeme als idealisierende Aussagen über eine äußere Realität behandelt. Diese beiden Geschwister werden oft als „reine“ und „angewandte“ Disziplin bezeichnet. Den Zusammenhang zwischen beiden diskutiert Nagel (1966, S. 215-233) am Beispiel der reinen und der angewandten Geometrie. Er bedient sich dazu eines einfachen Beispiels: Will man die Spitzen zweier Masten miteinander verbinden, die 30 Meter voneinander entfernt stehen, und von denen der eine 20 und der andere 15 Meter hoch ist, so werden die meisten von uns den zu überbrückenden Abstand zwischen den Mastspitzen nicht direkt ausmessen, sondern - unter Nutzung des Pythagoräischen Lehrsatzes - als $\sqrt{30^2 + (20 - 15)^2}$ berechnen, was einen Abstand zwischen den Mastspitzen von etwa 30,41 Metern ergibt. Was rechtfertigt die Durchführung dieser Berechnung? Die Gültigkeit des Pythagoräischen Lehrsatzes folgt „rein formal“ aus den Axiomen der Euklidischen Geometrie. Was aber rechtfertigt den Rückgriff auf diese Axiome?

Wechseln wir das Beispiel und betrachten die - zwar schon etwas beharrte, dafür aber vielleicht gut bekannte - Theorie des individuellen Wahlverhaltens von Luce (1959). Sie ist probabilistisch formuliert; es gelten also zunächst alle Grundannahmen der Wahrscheinlichkeitstheorie. Es werden nun als „Wahlen“ bezeichnete Ereignisse eingeführt, die man durch Ausdrücke der Form $\alpha \dashv \beta$ darstellen kann. Dabei ist für β eine Menge einsetzbar, für α hingegen eine Teilmenge oder ein Element der für β substituierten Menge. Die Einführung von Mengen impliziert die Grundannahmen der Mengenlehre, die aber bereits durch die Einführung von Wahrscheinlichkeiten, die mit dem Reden über Ereignisse Objekte mit Mengeneigenschaften voraussetzen, ohnehin als gültig unterstellt wird. Die Wahl „ $b \dashv B$ “ (gelesen: „ b aus B “) mit $b \in B$ wird interpretiert als Wahl der Alternative b bei gegebener Optionmenge B . Entsprechend steht „ $A \dashv B$ “ mit $A \subset B$ für die Wahl irgendeines Elements aus A bei gegebenem B . Die entscheidende Annahme ist nun, daß für $x \in A \subseteq B$ mit **Prob** (x) $\neq 0$ stets gilt:

$$\text{Prob } \{x \vdash B\} = \text{Prob } \{x \vdash A\} \text{Prob } \{A \vdash B\} \quad (2.1)$$

Führt man eine Gesamt-Optionenmenge Ω als Vereinigung aller im Kontext einer bestimmten inhaltlichen Fragestellung denkbaren Optionenmenge ein, dann folgt im endlichen Fall formal die Existenz von Funktionen

$$v : \Omega \rightarrow \mathfrak{R} \quad (2.2)$$

derart, daß für alle $x \in A \subseteq \Omega$ gilt:

$$\text{Prob } \{x \vdash A\} = \frac{v(x)}{\sum_{a \in A} v(a)}. \quad (2.3)$$

Vergleicht man - bei beliebiger Optionenmenge A , die x und y enthält - die Entscheidungswahrscheinlichkeiten für zwei Optionen x und y , dann ist

$$\frac{\text{Prob } \{x \vdash A\}}{\text{Prob } \{y \vdash A\}} = \frac{v(x)}{v(y)}. \quad (2.4)$$

Das ist die Grundlage der Bradley-Terry-Luce-Skalierung (Bradley & Terry, 1952; Luce, 1959). Führen wir eine weitere Skala ein, indem wir den Logarithmus eines v -Wertes mit u bezeichnen, dann ist

$$\log(\text{Prob } \{x \vdash A\}) - \log[\text{Prob } \{y \vdash A\}] = u(x) - u(y). \quad (2.5)$$

Damit aber sind wir bei einem Ausdruck aus dem Reservoir der log-linearen Modelle angelangt, die in diesem Band von Steyer (Kap. 15) vorgestellt werden, und die sich in der empirischen psychologischen Forschung beachtenswerter Beliebtheit erfreuen.

Nicht nur in der Geometrie, auch in der Psychologie gibt es also Anwendungen formaler Systeme. Dabei sei nicht unterstellt, daß die formale Einführung der „Anwendung“ zeitlich voraus geht. In den meisten Fällen ist es umgekehrt. Die von ägyptischen Landvermessern benutzten Techniken der Neuvermessung des Landes nach Nil-Überschwemmungen sind wesentlich älter als die „Elemente“ des Euklid, probit- und logit-Analysen gab es auch schon vor der Luceschen Konzeption individuellen Wahlverhaltens, Vertrauensbereiche für individuelle Testwerte und Reliabilitäten wurden bereits berechnet, bevor Lord und Novick (1968) das formale Konzept der „propensity distribution“ einführen.

Formale Wissenschaft liefert häufig im nachhinein Begründungen für das Vorgehen empirischer Forschung und alltäglicher Praxis. Sie schlägt Grundannahmen vor und weist nach, daß sich die Richtigkeit bestimmter Vorgehensweisen daraus ableiten läßt. Handelt es sich bei den Vorgehensweisen um Forschungsmethoden, so geht es also um die Analyse der „hidden assumptions“, die

diesen zugrundeliegen. Die Rede von den „hidden assumptions“ ist allerdings leicht mißverständlich. Da die gleiche Methode in durchaus unterschiedlichen formalen Systemen begründbar ist, darf man nicht der Vorstellung verfallen, es gebe zu Forschungsmethoden jeweils **die** verborgenen Grundannahmen, die zu akzeptieren jeder verpflichtet ist, der die entsprechende Methode benutzt, und die zu akzeptieren man ihm mithin auch unterstellen kann.

Die Vorteile der Analyse von Grundannahmen sind offensichtlich. Da formale Annahmensysteme Aussagen enthalten, die für alle Modelle trivial wahr sind, kann man Phänomenbereiche daraufhin untersuchen, ob sie als Modelle des jeweiligen Systems in Frage kommen. Wenn beispielsweise bei bestimmten Arten von Entscheidungen zwischen je zwei Alternativen bereits aufgrund von Plausibilitätsüberlegungen mit intransitiven Präferenzen zu rechnen ist (siehe etwa van Acker, 1977), dann liefern diese Entscheidungsbereiche keine Modelle für Theorien des Wahlverhaltens, deren Annahmen Präferenz-Transitivität implizieren. Bei dieser Art des Folgerns sind oft Methodenanalyse und psychologisch-inhaltliche Fragestellung kaum sinnvoll auseinanderzuhalten. Die Untersuchung der Grenzen eindimensionaler Präferenz-Skalierung wirft die Frage auf, wodurch sich denn ein- und mehrdimensionale Präferenzsysteme unterscheiden; zu deren Beantwortung bedarf es inhaltlicher Hypothesen, etwa über die Wahrnehmung und Analyse von Optionen und/oder über den Prozeß der Bildung von Präferenzurteilen.

Die Abstraktion von einer konkreten Forschungsmethode zu einem formalen Annahmensystem liefert überdies die Möglichkeit, von einem ursprünglich „üblichen“ Anwendungsbereich, der oft sogar die Terminologie des methodischen Apparates geprägt hat, zu einem anderen überzugehen. Ein Beispiel dafür ist die Verwendung der „Signal-Entdeckungs-Theorie“, die in ihrem Namen ihre Herkunft aus der Nachrichtentechnik (Peterson, Birdsall & Fox, 1954; van Meter & Middleton, 1954) nicht verleugnet, in der psychologischen Gedächtnisforschung (Egan, 1958). Die Verschränkung von formaler Analyse und empirischen Befunden verdeutlichen Mausfelds Ausführungen zur Messung im 13. Kapitel dieses Bandes; zur Nutzung formaler Theorien aus anderen Bereichen der Psychologie finden sich Beispiele in seiner Darstellung der Psychophysik im 4. Kapitel.

Empirische Forschung hat also oft einen formalen Partner; das Selbstverständnis der Psychologie als einer **empirischen Wissenschaft** begründet mithin keine **Beschränkung** ihres Methodenarsenals auf **empirische Methoden**, sondern fordert sogar, auch **formale Methoden** zumindest zur Begründung ihres empirischen Vorgehens einzusetzen. Wie steht es nun mit der Abgrenzung von einer Wissenschaft, die man „normativ“ nennen kann?

2.1.2 Normativität

Im Umfeld der Psychologie taucht die Rede vom **Normativen** häufig im Zusammenhang mit Rationalitätskonzepten auf, bei denen zwischen einer deskriptiven, einer präskriptiven und einer normativen Nutzung unterschieden wird (Bell, Raiffa & Tversky, 1988). Dabei wird normative Nutzung oft so aufgefaßt, daß hiermit der Anspruch verknüpft sei, jeder Mensch möge sich entsprechend der jeweiligen Konzeption verhalten. So verstandene normative Wissenschaft liefert also Handlungsanweisungen; ihre Systematisierung besteht aus dem Aufbau von Systemen, in denen sich diese Anweisungen ableiten lassen. In der mathematisch orientierten Spiel- und Entscheidungstheorie, die sich ausgiebig mit Rationalitätsvorstellungen beschäftigt, herrscht dagegen eine andere Sichtweise vor. Prämissen (für Systeme von Handlungsanweisungen) werden als Normen im Sinne von **Maßstäben** angesehen; sie definieren implizit eine Klasse von Handlungen als „normgerecht“ im Sinne der jeweiligen Prämissen. Mit formalen Methoden kann man dann untersuchen, welche Eigenschaften derartigen Handlungen und ihren Konsequenzen zukommen und wie sie zu finden sind.

Selten (1989) unterscheidet zwischen „praktischer“ und „idealer“ Normativität. **Praktische Normativität** sieht die durch einen Satz von Prämissen beschriebene Norm als Handlungsanweisung, **ideale Normativität** versucht, unbeschadet der Möglichkeiten und Grenzen realer Menschen zur Klärung des jeweiligen Normkonzeptes beizutragen. Die hierbei eingesetzten Methoden sind die der formalen Analyse; da damit stets Implikationen der Prämissen - also analytisch wahre Aussagen - gewonnen werden, nennt Aumann (1988) auch die zugehörigen Methoden „analytisch“. In verschiedenen Kontexten hat sich offensichtlich ein unterschiedlicher Sprachgebrauch eingebürgert; die „analytischen“ Methoden sind genau die gleichen, die wir zuvor (siehe S. 8) als „formal“ bezeichnet haben. Auch die Systeme selbst sind oft formalisiert; die Prämissen werden als Axiome dargestellt, „formale“ Systeme und „normative“ Systeme sehen gleich aus. **Ein** Unterschied aber bleibt. Hatten wir es zuvor mit Grundannahmen zu tun, die auch als Festsetzungen oder als idealisierende Aussagen über eine **gegebene** äußere Realität angesehen werden konnten, so haben wir es im Bereich des Normativen mit Aussagen zu tun, die als Festsetzungen oder als Aussagen über **mögliches** menschliches Handeln interpretierbar sind.

Natürlich hindert nichts daran, auch Prämissen eines normativen formalen Systems als empirische Hypothesen über real vorfindbares Verhalten zu interpretieren. Dieses Vorgehen wird als „deskriptive Nutzung“ normativer Systeme bezeichnet; diese aber hat sich - zumindest im Bereich der Rationalitätskonzepte - als wenig brauchbar erwiesen. Es gibt jedoch mindestens zwei andere Verwendungsmöglichkeiten. Wir können real beobachtbare **Handlun-**

gen durch Vergleich mit verschiedenen Normsystemen beschreiben und klassifizieren. Hier wird „Norm“ als „Vergleichsmaßstab“ aufgefaßt. Für jedes System läßt sich dann untersuchen, unter welchen Bedingungen normgerechtes Handeln (im Sinne eines jeweils spezifischen Satzes von Prämissen) zu erwarten ist, und unter welchen nicht. Zur Erklärung dabei auftretender Befunde kann man dann (beispielsweise) der Frage nachgehen, welche kognitiven oder motivationalen Mechanismen zum normgerechten Handeln führen. Eine zweite Möglichkeit ist, **Situationen** daraufhin zu untersuchen, welche Möglichkeiten normgerechten Handelns sie überhaupt bieten. Es wäre überraschend, wenn Situationsarten, die sich auf diese Weise unterscheiden lassen, nicht auch durch Unterschiede des in ihnen auftretenden Realverhaltens gekennzeichnet wären. Normative Systeme können also die Grundlage für Methoden der Situationsbeschreibung und -klassifikation auf der Basis der Analyse situationsspezifisch erfüllbarer bzw. nicht erfüllbarer Handlungsnormen bieten. Tack (1991) exemplifiziert ein derartiges Vorgehen am Beispiel der Normen kooperativer Rationalität.

Diese Überlegungen zum Verhältnis von empirischer, formaler und normativer Wissenschaft zeigen, daß aus dem Selbstverständnis der Psychologie als einer empirischen Wissenschaft zunächst lediglich folgt, daß sie neben einer empirischen Methodik stets auch der formalen bzw. analytischen Methoden bedarf, um Teile ihres Vorgehens in größeren Kontexten begründen und normative Vorstellungen sowohl zur Handlungsbeschreibung als auch zur Situationsanalyse nutzen zu können. Aus der Empirizität der Psychologie folgt **keine Beschränkung** auf empirische Methoden, sondern bestenfalls, daß diese **unverzichtbar** sind. Wenn also „empirische Methoden“ nicht schlicht als „Methoden einer empirischen Wissenschaft“ umschrieben werden können, dann müssen wir nach anderen Abgrenzungs-Gesichtspunkten suchen.

2.2 Empirische und rationale Erkenntnis

Bei der Frage, was Psychologie zu einer empirischen Wissenschaft macht, sind wir - in Anlehnung an de Groot (1969) - davon ausgegangen, sie stelle Wissen über die Welt, in der wir leben, zur Verfügung. Im vorausgegangenen Abschnitt haben wir gezeigt, daß daraus aufgrund der doppelten Interpretierbarkeit von Aussagensystemen kein Verzicht auf formale Methoden, sondern sogar deren Notwendigkeit folgt. Es dürfte aber auch klar geworden sein, daß formale Methoden allein ebenfalls nicht ausreichen, da sie keine Grundlagen für die Behauptung und Kontrolle der **faktischen Richtigkeit** oder - zumindest - **Brauchbarkeit** von Tatbestandsaussagen bereitstellen.

Notwendig sind also Methoden, mit denen man Aussagen **als faktisch gültig rechtfertigen** kann, indem man Beweismaterial für ihre Geltung beibringt. Be-

trachtet man Erkenntnis jeder Art, also auch wissenschaftliche Erkenntnis, aus einer individualistischen Position als Überzeugung von einzelnen, dann führen solche Überlegungen zunächst auf einen Regreß, der nach Chalmers (1989, S. 116) mindestens seit Platon bekannt ist. Zur Rechtfertigung einer Aussage werden andere Aussagen herangezogen; um diese zu rechtfertigen wieder andere und so fort. Um diesen Regreß zu stoppen, braucht man Aussagen, die selbst keiner weiteren Rechtfertigung bedürfen, weil sie „in gewissem Sinne durch sich selbst gerechtfertigt sind“ (Chalmers, 1989, S. 116). Der „klassische Rationalismus“ geht davon aus, daß solche Grundlagen der Erkenntnis durch Nachdenken gewonnen werden können, da sie sich als eindeutig, unabhängig und selbstverständlich erweisen, während der „Empirismus“ die Grundlagen der Erkenntnis durch Sinneswahrnehmung zu sichern bemüht ist.

2.2.1 Klassisch rationales Vorgehen

Descartes (1637) wird oft als der erste klassische Rationalist der Neuzeit bezeichnet. Da es für uns nur schwer vorstellbar ist, daß man durch bloßes Nachdenken zu faktisch wahren Aussagen kommt, sei ein Beispiel skizziert, das gleichzeitig die Tücken einer solchen Methodik veranschaulicht. Damerow, Freudenthal, McLaughlin und Renn (1992, S. 30 ff.) analysieren unter anderem eine Tagebuchnotiz Descartes' aus dem Jahre 1618, in der dieser den Zusammenhang zwischen Zeit und Weg beim freien Fall herauszufinden versucht. Zur Notiz gehört eine Skizze, die in Abbildung 1 dargestellt ist.

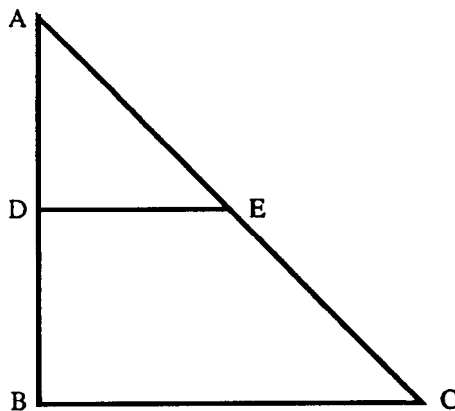


Abb. 1: Skizze zu einer Tagebuchnotiz Descartes' von 1618

Descartes notiert, ein ausnehmend kluger Mann (womit Beekman gemeint ist) habe ihn mit einem Problem konfrontiert, das - stark verkürzt - im Kern

fragt: „Ein Stein fällt in einer Stunde von **A** nach **B** und wird von der Erde mit einer gleichbleibenden Kraft angezogen; in welcher Zeit wird **AD** durchfallen?“ Descartes sagt, er habe das Problem gelöst. Fläche **ABC** steht für die „Bewegung“, der Unterschied zwischen dem Punkt **A** und der Strecke **BC** zeigt die Ungleichheit der Bewegung. Somit wird Strecke **AD** in der durch Fläche **ADE** gegebenen Zeit durchlaufen, und **DB** in der durch **DEBC** repräsentierten Zeit. Dabei müsse man allerdings beachten, daß die kleinere Fläche für langsamere Bewegung steht. Nun ist **ADE** ein Drittel von **DEBC**, und **AD** wird somit dreimal so langsam durchfallen wie **AB**.

Diese Argumentation wirkt heute etwas chaotisch. Flächen werden als Kehrwerte des Zeitbedarfs interpretiert. Ist die zugehörige Fallstrecke in mehreren Fällen gleich (wie bei **AD** und **DB**), so können wir diese Strecke als Einheit wählen; die Flächen sind dann proportional den mittleren Geschwindigkeiten. Dabei erhalten die Längen der basisparallelen Strecken zwischen **AB** und **AC** den Charakter von Geschwindigkeiten, die somit proportional der jeweils zurückgelegten Fallstrecke sind (und nicht - wie man heute bereits in der Schule lernt - proportional zur Fallzeit). Damerow et al. (1992) untersuchen in ihrem Buch die Möglichkeiten einer konsistenten Interpretation der damals benutzten Sprache, und sie analysieren, wann und wie Galileo, Descartes, Beekman und andere erfaßt haben, daß die Geschwindigkeit eines frei fallenden Körpers nicht proportional zur durchfallenen Strecke, sondern zur verflissenen Zeit ist. Immerhin: Dieser Wechsel fand statt. Eine Fehlannahme über die uns umgebende Welt wurde durch Nachdenken korrigiert und nicht durch Beobachtung und/oder Messung.

Zwei Fragen dazu. Die Erste: Warum versucht man im 17. Jahrhundert die Frage nach einem faktisch geltenden Gesetz durch Nachdenken zu lösen, wo doch empirische Forschung, die nachschaut und registriert, zu dieser Zeit großen Aufschwung nimmt. Olaus Römer mißt die Lichtgeschwindigkeit, Evangelista Toricelli und Otto von Guericke experimentieren mit dem Luftdruck, Robert Boyle legt empirische Daten zur Nicht-Existenz der klassischen vier Elemente vor, Antony van Leeuwenhoek sieht zum ersten Mal unter dem Mikroskop Bakterien und Thomas Hobbes propagiert eine „soziale Physik“, deren Erkenntnisgrundlage Selbstbeobachtung durch Introspektion ist (Störig, 1970, S. 299-360).

Eine mögliche Antwort ist rein pragmatisch. Praktisch realisierbare Fallhöhen sind so klein und Meßtechniken so ungenau, daß man allein mit technisch unterstützter Beobachtung keine Daten erhält, aus denen sich das Fallgesetz ergibt. Die uns umgebende Luft bremst statt zu beschleunigen; empirische Daten werden nicht einmal symmetrisch um ideale Werte streuen. Die in der Psychologie so beliebte Mittelwerts-Strategie hilft also auch nicht. Könnte man extrem genau messen und gäbe es extrem lange Fallstrecken, dann ließe

sich das Gesetz des freien Falls in empirischen Daten auch nicht wiederfinden, da während eines jeden Falles eine Annäherung an den Erdmittelpunkt stattfindet, womit die (beim Fallgesetz als konstant unterstellte) Erdanziehung zunimmt.

Damit sind wir bei einer **zweiten** Antwort. Der freie Fall ist eine **Idealisierung**, eine kontrafaktische Aussage darüber, was geschehen würde, wenn bestimmte Annahmen erfüllt wären, deren Erfüllung praktisch nicht zu sichern ist. Andererseits kann man über die „Störfaktoren“ und deren Effekte ebenfalls Aussagen machen, so daß zunehmende Berücksichtigung von immer mehr „sonstigen“ Einflüssen eine zunehmend bessere Anpassung empirischer Daten ermöglicht. Gegenstand der rationalen Erkenntnissuche ist also etwas, das als Idealisierung in der uns umgebenden Welt durchaus vorhanden ist, sich aber wegen der Überlagerung von Effekten einer Bedingungsvariation, die im Idealfall nicht vorgesehen ist, der direkten Beobachtung und Registrierung entzieht. Die Brücke zur Empirie besteht darin, daß man unter expliziter Berücksichtigung zusätzlicher Bedingungsfaktoren etwas darüber aussagen kann, wie sich die Diskrepanz zwischen Realdaten und Idealphänomen ändert, wenn man eben diese zusätzlichen Faktoren kontrolliert.

2.2.2 Metaprinzipien

Nun zu der Frage: Was ist es, das in rationalen Analysen letztendlich als so eindeutig, unabhängig und selbstverständlich angesehen wird, daß man es als Erkenntnisgrundlage nutzt, ohne es weiter in Frage zu stellen, und das wir ein „Metaprinzip“ nennen wollen? Im Beispiel des freien Falls ist dies ein **Einfachheits-Prinzip**. Wenn die Geschwindigkeit beim freien Fall nicht konstant ist, sondern offensichtlich zunimmt, dann ist die einfachste Möglichkeit eine Proportionalität der Geschwindigkeit zu einer Größe, deren Zunahme während des Falles trivial ist. Als solche Größen stehen Fallstrecke und Fallzeit zur Verfügung. Descartes' Pech war, daß er zunächst auf die falsche Karte setzte, nämlich auf die Fallstrecke.

Ein anderes Prinzip ist **Optimalität**. Schoemaker (1991) bezeichnet es als sehr erfolgreiches und flexibles wissenschaftliches Metaprinzip und verweist auf Nutzenmaximierung in der Ökonomie, Aufwandsminimierung in der Physik, Entropie in der Chemie und das Überleben der am besten angepaßten genetischen Ausstattung in der Biologie. **Optimalitätsprinzip** ist allgemein die Annahme, ein Verhalten oder ein anderes empirisch faßbares Phänomen verlaufe so, daß eine geeignete Zielfunktion unter Berücksichtigung gegebener Randbedingungen maximiert oder minimiert wird. Verschiedene Zielfunktionen entsprechen unterschiedlichen **Optimalitätsstandards**. Er weist darauf hin, daß bei Coombs, Dawes und Tversky (1970) Beispiele dafür zu finden sind, daß

sich auch in der Psychologie beispielsweise im Kontext der Signalentdeckungstheorie, der Erforschung des Lernens und insbesondere der operanten Konditionierung sowie in der Entscheidungsforschung Formalisierungen finden, denen Optimalisierungsannahmen zugrundeliegen (Schoemaker, 1991, S. 207).

Am Beispiel der Optimalität läßt sich zeigen, wie Wissenschaft heute mit Meta-Prinzipien umzugehen pflegt. Der klassische Rationalismus ging davon aus, daß letzte Prinzipien, die als eindeutig, unabhängig und selbstverständlich jeder gegen sie gerichteten Argumentation standhalten, stets richtig sind, und daß mithin aus ihnen abgeleiteten Aussagen über die äußere Welt faktische Richtigkeit zukommt. Schoemaker (1991) bezeichnet hingegen Meta-Prinzipien als „Heuristiken“. Eine Heuristik führt zwar oft und einfach, aber **niemals immer** zur richtigen Lösung eines Problems. Da ist die Gefahr einer Fehlattribution. Kitcher (1985) weist darauf hin, daß allein die Erklärbarkeit von Verhalten in einem Optimierungsmodell noch kein hinreichender Grund dafür ist, Optimierung als einzig mögliches Erklärungsprinzip und damit als in der äußeren Welt gegeben zu betrachten. Eine überzogene Bestätigungstendenz kann dazu führen, daß für eine Optimalitätsannahme Daten aus Bereichen angehäuft werden, die mit hoher Wahrscheinlichkeit die jeweilige Annahme stützen, also nicht „kritisch“ sind. Schoemaker (1991, S.212) führt diese Fehlermöglichkeit explizit an; sie dürfte allerdings kaum spezifisch für den Umgang mit Meta-Prinzipien sein, sondern jeden plausibel erscheinenden Erklärungsansatz treffen. Überziehung des (angenommenen) Geltungsbereichs und zu schwache empirische Kontrolle wegen allzu eindrucksvoller Plausibilität sind weitere Gefahren. Meta-Prinzipien liefern somit lediglich Aussagen mit hypothetischem Charakter, die der Kontrolle durch Validierung bedürfen. Das Methodenarsenal des klassischen Rationalismus wird hier als Werkzeug zur Generierung und Elaboration (hypothetischer) Theorien benutzt. Über Heuristiken der Theorie- und Hypothesenkonstruktion findet sich mehr in einem Beitrag von Dörner (Kap. 8 dieses Bandes).

Was sind nun die Kriterien der Validierung rational begründeter Theorien und Hypothesen? Schoemaker (1991, S. 238) führt fünf Kriterien an, von denen das erste die Genauigkeit ist, mit der existierende Daten erklärt oder vorhergesagt werden können, und das zweite die Generierung neuer Hypothesen und deren Validierbarkeit. Damit ist offensichtlich, daß - im Gegensatz zum klassischen Rationalismus - die rationale Begründung nicht als hinreichend gesehen wird, sondern als möglicherweise zu Fehlschlüssen führende Heuristik, die der Kontrolle durch Anwendung empirischer Methoden bedarf. Ansonsten ist der Kriterienkatalog eine von vielen Varianten der Zusammenstellung unterschiedlicher Aspekte der Theoriebewertung, mit denen sich in diesem Buch Gadenne (Kap. 9 dieses Bandes) ausführlicher beschäftigt.

Beispiele für die Bedeutung derartiger Überlegungen im Rahmen psychologischer Forschung bieten in breiter Auswahl die Kommentare zu Schoemakers (1991) Artikel, von denen mehr als die Hälfte von Psychologen stammen. Dabei zeigt sich die Spannweite von einer eng empirizistischen Position, die allein Daten gelten läßt und rational begründete Prinzipien für überflüssig hält, über eine Einstellung, derzufolge rationale Systeme die Bedeutung globaler „Nullhypothesen“ haben, bis hin zu Beispielen dafür, daß die Nutzung von Meta-Hypothesen neue Einsichten und Erklärungsmöglichkeiten eröffnet. Die „Nullhypothesen-Position“ wird oft bezüglich des Modells vom rationalen Entscheider vertreten. Danach sind **es** die **Abweichungen** vom rationalen Entscheidungsverhalten, die psychologisch relevant und erklärungsbedürftig sind. Eine andere Einstellung vertreten etwa Baron (1991) und Tack (1991), die die **Realisierung** rationalen Verhaltens für genauso erklärungsbedürftig halten. Die Tatsache, daß es sowohl reales rationales Verhalten als auch Abweichungen davon gibt, wird - zunächst wieder rein rational und unter Plausibilitäts Gesichtspunkten - durch die Existenz von Heuristiken erklärt, die unter bestimmten und im Alltag häufigen Bedingungen erfolgreich sind, während sie unter anderen Bedingungen Abweichungen liefern. Die Identifikation solcher Heuristiken und die Überprüfung der aus ihnen ableitbaren Hypothesen über die Situationsabhängigkeit realisierter Rationalität ist es, die dann mit empirischer Methodik zu untersuchen ist. Diese Sichtweise entspricht jener, die Anderson (1991) bezüglich der Optimalität des menschlichen Gedächtnisses vertritt.

Rationale Theorie- und Hypothesenbildung mit der zugehörigen Methodik fällt - sofern die jeweiligen Prinzipien und deren Varianten hinreichend formalisiert sind - mit der zuvor angesprochenen formalen Analyse zusammen. Wie in den Bemerkungen zur Validierung erwähnt, wird - damit von empirischer Wissenschaft die Rede sein kann - darüber hinaus noch eine Methodenart benötigt, die Aussagen mit Beobachtungen verknüpft. Diese Methodenart ist gemeint, wenn wir im folgenden von **empirischen Methoden** reden.

2.3 Empirische Methoden

Empirische Methoden liefern Aussagen, die Beobachtungsergebnisse möglichst direkt wiedergeben, und deren faktische Richtigkeit sich ohne weiteren Begründungszwang aus dem jeweils beobachteten Sachverhalt ergibt. Man kann zweifeln, ob es überhaupt Vorgehensweisen gibt, die dies leisten können. Vorausgesetzt wird sowohl die Veridikalität von Beobachtungen als auch die Möglichkeit einer Sprache, die Beobachtungsergebnisse unverzerrt zu repräsentieren vermag. Eine solche Sprache ist in der Wissenschaftstheorie unter dem Namen „Beobachtungssprache“ (siehe Stegmüller, 1970, S. 293-399) be-

kannt. Es gibt viele Argumente, mit denen man die Möglichkeit einer derartigen Sprache in Zweifel ziehen kann; nahezu jede wissenschaftstheoretische Position hat hierzu eine eigene Meinung geäußert und begründet (siehe die entsprechenden Absätze bei Chalmers, 1989). Das gleiche gilt für die Frage nach der Möglichkeit unbelasteter und wahrheitsgetreuer Beobachtung. Hierzu findet man gelegentlich den Hinweis, am ehesten fehlerfrei feststellbar seien Raum-Zeit-Koinzidenzen; die zuverlässigsten Beobachtungsaussagen seien also solche, in denen festgestellt wird, daß sich Objekt **A** zum Zeitpunkt t am Punkt mit den Ortskoordinaten $\langle x, y, z \rangle$ befindet bzw. befunden hat. Da Veränderungen von Raum-Zeit-Koinzidenzen Bewegungen sind, sind auch diese problemlos feststellbar.

In der Psychologie hat eine derart enge Konzeption empirischer Letztbegründung wissenschaftlicher Aussagen außer in extremen Spielarten des Behaviorismus kaum eine Rolle gespielt. Soweit sich Psychologie mit Verhalten beschäftigt, hat sie es mit beobachtbaren Phänomenen zu tun, die bestenfalls in manipulativ extrem kontrollierten Situationen durch hinreichend einfach beschreibbare Klassen von Raum-Zeit-Koinzidenzen oder deren Veränderungen erfaßt werden können. Hinzu kommt, daß nach einer ebenfalls zur Psychologie gehörenden Tradition, für die etwa William James (1890) steht, Gedanken und Gefühle zu den grundlegenden Daten des Faches gehören, für die man entweder eine besondere Art von Beobachtung - „Introspektion“ genannt - postulieren muß, oder eine Zugänglichkeit durch Interpretation beobachtbarer Sachverhalte.

Die Charakterisierung empirischer Methoden bedarf daher der Abschwächung. Verlangt werde nun lediglich, daß empirische Methoden Aussagen liefern, von denen allgemein **unterstellt** wird, sie gäben Beobachtungsergebnisse möglichst direkt wieder, und für deren Richtigkeit keine weiteren Begründungen **gefordert** werden. Das gibt dem Ganzen einen konventionalistischen Anstrich, der durchaus Sinn macht. Wenn Psychologie - wie jede Wissenschaft - ein soziales Unternehmen ist, dann muß eine Aussage, zu der keine weiteren Begründungen verlangt werden, für andere so verständlich sein, daß die zugrundeliegende Feststellung (Beobachtung) nachvollzogen werden kann. Dies setzt sozialen Konsens darüber voraus, was als „in sich selbst begründet“ gelten möge. Andererseits ist die abgeschwächte Charakterisierung aber auch nicht ausschließlich konventionalistisch. Der allgemeine Verzicht auf weitere Begründung bei bestimmten Aussagenarten kann jederzeit in Frage gestellt werden. Diskussionen über Erhebungsartefakte oder Kontrollen der Inter-Rater-Reliabilität von Beobachtungsdaten sind Beispiele dafür.

Zur weiteren Explikation empirischer Methodik sollen nun zwei Aspekte betrachtet werden, in denen sich empirische Methoden unterscheiden, nämlich das Ausmaß der „Datenrestriktion“ und der „Manipulationsgrad“.

2.3.1 Datenrestriktion

Mit „**Datenrestriktion**“ ist gemeint, wie stark und auf welche Weise der im Prozeß einer wie auch immer gearteten Beobachtung bzw. Registrierung stattfindende Übergang von einem Phänomen zu einem objektiviert vorliegenden Datum die Menge der erfaßten Phänomenaspekte einschränkt. Das eine Extrem - mit minimaler Restriktion - wäre beispielsweise die Video-Aufzeichnung eines Verhaltens mit möglichst vielen Kameras aus verschiedenen Perspektiven, das andere Extrem die rechnergesteuerte Erhebung eines Testwertes, bei der lediglich eine Gesamtzahl richtiger Antworten ausgewiesen wird und nicht einmal mehr das Antwortmuster erkennbar ist.

Beide Extreme - wie auch die dazwischen liegenden Möglichkeiten - haben ihre spezifischen Anwendungsbereiche. Eine möglichst geringe Datenrestriktion ist das Ideal einer **interpretierenden Wissenschaft**. Soll Interpretation, wie von Soeffner und Hitzler im 3. Kapitel dieses Bandes ausgeführt, die Vielfalt der Deutungsmöglichkeiten eines Sachverhaltes erschließen, dann ist zunächst unbekannt, welche Möglichkeiten es gibt und welche Aspekte des Sachverhalts zur Begründung und Kontrolle benötigt werden. Die ideale Interpretationsbasis kommt mithin einer vollständigen Repräsentation der zu deutenden Gegebenheit, die diese als eine Art Kopie über die Zeit hinweg bewahrt, möglichst nahe. Dem dient ein empirischer Zugang durch möglichst freie und umfassende Beobachtung.

Nun zum anderen Extrem. Eine starke Datenrestriktion setzt voraus, daß im vorhinein bekannt ist, zu welchen Aspekten beobachtbarer Gegebenheiten welche Informationen benötigt werden. Das ist beim streng **theorie- und hypothesesengeleiteten Vorgehen** oft der Fall, wenn die benutzten Hypothesen derart spezifiziert sind, daß sie die Art der zu ihrer Kontrolle benötigten Informationen festlegen. Soll eine Hypothese zur ‚perceptual defense‘ geprüft werden, die unterschiedliche mittlere Erkennungszeiten für Wörter verschiedener emotionaler Tönung bei Personen mit unterschiedlicher Neuroseform behauptet, dann benötigt man als Daten Worterkennungszeiten, eine Kennzeichnung der emotionalen Tönung der benutzten Wörter, klinisch-psychologische Diagnosen der Versuchspersonen und sonst nichts.

Es gibt noch einen anderen guten Grund für Datenrestriktion, der in der alltäglichen Forschungspraxis oft eine Rolle spielt. Als sogenannte „Theorie“ oder „Hypothese“ liegt lediglich eine unscharfe Vermutung über einen nur wenig spezifizierten Zusammenhang vor (wie bei der Katharsis-Theorie, der Frustrations-Aggressions-Theorie und ähnlichem mehr). Weitere Elaboration führt auf Annahmen über Zusammenhänge, die als Korrelationen ungleich Null verstanden werden, über Mittelwertsunterschiede Verschiedenster Art, die man als varianzanalytische Haupt- und/oder Wechselwirkungseffekte deu-

tet, über Ähnlichkeitsbeziehungen, die faktoren- oder clusteranalytischen Konzepten entsprechen, und so fort. Diesen und ähnlichen Elaborationen ist gemeinsam, daß sie eine bestimmte **Strategie der Datenauswertung** nahelegen. Jede Entscheidung für eine bestimmte Technik der Auswertung ist aber zugleich die Festlegung auf Daten, deren Eigenschaften die gewünschte Auswertung überhaupt ermöglichen.

Damit die in Erhebungen generierten Daten für bestimmte Auswertungen brauchbar sind, benötigt man einen konzeptuellen Rahmen, in dem sich entsprechende Datenformate darstellen und analysieren lassen. Dabei muß gezeigt werden, aufgrund welcher Regeln bei der Datengenerierung Formate mit Eigenschaften resultieren, die Voraussetzung für unterschiedliche Möglichkeiten der Weiterverarbeitung sind. Eine Basis hierzu bietet die **Datentheorie** von Coombs (1964), deren weitere Entwicklung etwa bei Roskam (1983) dargestellt wird. Allgemein spricht man dann, wenn Restriktionen zu Daten in einem vorher festgelegten Format führen, auch von einer „**datenförmigen Empirie**“.

2.3.2 Manipulationsgrad

Der **Manipulationsgrad** kann unabhängig von der Datenrestriktion variieren. Dies wird oft übersehen, da in der einschlägigen Literatur häufig **Felduntersuchungen und Laboruntersuchungen** als Extrema eines Kontinuums unterschiedlicher Lebensnähe betrachtet werden (siehe Bortz, 1984, S. 33), wobei manipulative Kontrolle und Datenrestriktion als Reduktionen dieser „Lebensnähe“ nicht immer klar auseinanderzuhalten sind.

Betrachten wir eine verkehrspsychologische Untersuchung, bei der an verschiedenen gestalteten Ampelkreuzungen durch Lichtschranken registriert wird, wie lange nach einem Wechsel von Grün auf Rot wieviele Autos noch in den Kreuzungsbereich einfahren. Manipulative Kontrolle ist so gut wie nicht vorhanden; die Autofahrer sind in der gleichen alltäglichen Situation, in der sie auch wären, wenn keine Registrierung stattfinden würde. Die Datenrestriktion ist allerdings massiv; die Erhebungstechnik liefert Zeitverteilungen und sonst nichts. Die Untersuchungssituation ist lebensnah, die Daten sind hingegen ausgesprochen realitätsfern abstrahierend. Übrigens wäre es sinnvoll, sich in diesem Kontext den Ausdruck „Lebensnähe“ abzugewöhnen und besser von „Alltagsnähe“ zu sprechen, denn auch in der Laborsituation ist die Versuchsperson selbst unter extremer manipulativer Kontrolle immer noch ein lebender Mensch. Das Labor, in dem Versuchspersonen nur noch durch präparierte Teile vertreten sind, gibt es in der Psychologie nicht. Die Konzeption einer „lebensfernen“ Laborsituation sollten wir den Wissenschaften überlassen, in denen diese wirklich vorkommt.

Zeigt das verkehrspsychologische Beispiel die Vereinbarkeit eines geringen Manipulationsgrades mit einer starken Datenrestriktion, so wäre ein Gegenstück die Untersuchung einer Versuchsperson in einer perfekt abgedunkelten camera silens, bei der mit Tonband und Infrarot-Kamera registriert wird, wie sich im Laufe der Zeit Äußerungen zu Halluzinationen und Indikatoren eines Verlusts des emotionalen Gleichgewichts einstellen. Die Situation ist stark manipuliert und entsprechend alltagsfern, die Daten sind kaum restringiert und entsprechend realitätsnah und anschaulich.

2.3.3 Interne und externe Validität

Eng verknüpft mit dem Manipulationsgrad ist die Konzeption einer „**internen**“ und einer „**externen Validität**“, die alles andere als klar ist. Die Angabe, eine Untersuchung sei intern valide, „... wenn ihre Ergebnisse eindeutig interpretierbar sind (Bortz, 1984, S.29)“, besagt solange wenig, wie nicht klar gestellt ist, was als Ergebnis und was als Interpretation gelten soll; die Gleichsetzung externer Validität mit Generalisierbarkeit über die Bedingungskonstellation einer Untersuchungssituation und den Kreis der Versuchspersonen hinaus ist ohne Präzisierung des Generalisierungsbegriffs ebenso unklar.

Zur Klärung beschränken wir uns auf empirische Methoden, die Angaben darüber liefern, daß bei bestimmten Bedingungskonstellationen bestimmte Phänomene auftreten. Die Bedingungskonstellationen seien durch Werte endlich vieler Variablen charakterisiert, die „**unabhängige Variablen***“ heißen, während die Beschreibung der resultierenden Phänomene durch Wertangabe endlich vieler anderer Variablen erfolgt, die „**abhängig**“ genannt werden.

Sind unabhängige und abhängige Variablen getrennt, dann ist zu fragen, ob eine **neue** Beobachtung mit einer bereits früher realisierten Kombination von Werten der **unabhängigen** Variablen mit der gleichen Wertekombination **abhängiger** Variablen verknüpft ist wie zuvor. Wenn ja, dann ist die durchgeführte Untersuchung - und insbesondere die benutzte empirische Methode - „valide“. Dabei läßt sich das Prinzip der vollständigen Übereinstimmung von Wertekombinationen nach Einführung einer Unterschiedlichkeits-Metrik auf verschiedene Weise sinnvoll probabilistisch abschwächen.

In der Psychologie stammt die Bezeichnung „Validität“ aus der Theorie psychodiagnostischer Verfahren und deren Anwendung; sie kennzeichnet dort die Brauchbarkeit eines Diagnostikums für Vorhersagen in neuen Situationen. Die Vorstellung, daß Validität stets eine Art **prognostischer Brauchbarkeit** ist, bleibt beim hier skizzierten Präzisierungsversuch erhalten. Liegen valide empirische Daten vor, dann können wir - nach üblichem Sprachgebrauch - auch davon sprechen, die durch unabhängige Variablen beschriebene Bedingungs-

kombination „bewirke“ die durch abhängige Variablen erfaßten Phänomenaspekte. Daß **A** ein **B** bewirkt, besagt, daß **A** - wie immer es auch zustandekommen möge - stets mit **B** einhergeht, daß also der Befund, **A** trete mit **B** kombiniert auf, valide ist.

Ein zusätzliches Problem soll nicht unerwähnt bleiben: Wenn Validität eine Art prognostischer Eignung ist, dann wird dazu nicht nur eine empirische Erhebungsmethode benötigt, die Einzelaussagen liefert, sondern zusätzlich eine Methode zur Generalisierung dieser Aussagen, um aus deren Ergebnis eine Individualaussage über einen neuen Fall ableiten zu können. Der Weg von Einzelbeobachtungen zu generalisierenden Aussagen ist aber keineswegs trivial; hierzu bedarf es einer Methode, die durch vorliegende Einzeldaten „vernünftig begründete“ allgemeine Aussagen ermöglicht und damit eine Lösung des „schwachen Induktionsproblems“ (siehe Westermann & Gerjets in Kap. 10 dieses Bandes) anbietet.

Dieser Versuch einer Präzisierung des Validitätsbegriffes ist keineswegs der einzig vertretbare. Eine Alternative stellen Erdfelder und Bredenkamp im 14. Kapitel dieses Bandes vor. Sie gehen ebenfalls davon aus, daß Validität mit der Beziehung zwischen einem empirisch gewonnenen Datensatz und einer allgemeineren Aussage zu tun hat, die als Hypothese aufgefaßt wird. Funktion des Datensatzes ist die Prüfung dieser Hypothese, die „streng“ und „fair“ sein soll. Validität ist bei Erdfelder und Bredenkamp eine Funktion von „Strenge“ und „Fairneß“. Diese beiden Konzepte werden in Kapitel 14 dieses Bandes genauer spezifiziert.

Wie steht es nun im hier angedeuteten Ansatz mit dem Unterschied zwischen interner und externer Validität? **Interne Validität** geht davon aus, daß die neuen Vergleichsfälle „unter Untersuchungsbedingungen“ realisiert werden, das heißt, daß alle nicht kontrollierten Variablen genau so variieren wie bei den Ausgangsbeobachtungen. Somit ist interne Validität in einer streng manipulierten Laborsituation, die nicht kontrollierte Variablen auf einen möglichst engen Variationsbereich einschränkt, sicher größer als bei alltagsnahen Beobachtungen in Situationen, in denen sehr viele Variablen breit variieren.

Faßt man die eingangs erwähnte Umschreibung interner Validität als möglichst eindeutiger Ergebnisinterpretation so auf, daß damit auch eine geringe Wahrscheinlichkeit unkontrollierter konfundierter Variablen gemeint ist, so ist dieser Aspekt durch den hier vorgestellten Präzisierungsentwurf nicht abgedeckt. Es ist wenig einsichtig, daß manipulative Kontrolle die Konfundierungswahrscheinlichkeit verringern sollte, sofern man nicht eine sehr spezifische Kontrollvariante unterstellt, die alle nicht registrierten Bedingungsaspekte auf geheimnisvolle Weise konstant zu halten versteht und so eine ideale Begründung für ceteris-paribus-Klauseln liefert. Wir werden bei Überlegungen

zur Komplexität des Phänomenbereichs der Psychologie auf diesen Punkt zurückkommen (siehe S. 26-30).

2.3.4 Laborbedingungen und Alltagssituationen

Wird zur Prüfung **interner** Validität ein Vergleichsfall unter Untersuchungsbedingungen herangezogen, **so** geht es bei **externer Validität** um den Vergleich mit Daten aus anders gearteten Situationen, wobei im allgemeinen an Alltagssituationen gedacht ist. Dabei tauchen zunächst mindestens zwei Schwierigkeiten auf: Die unter Untersuchungsbedingungen realisierte Bedingungskonstellation kann in Alltagssituationen überhaupt nicht vorkommen, oder sie kommt vor, das Ergebnis entspricht **nicht** dem Validitätskriterium, ist aber dennoch mit den jeweiligen Untersuchungsbefunden vereinbar.

Der erstgenannte Fall liegt vor, wenn Versuchspersonen in einer Untersuchungssituation etwas tun, was sie im Alltag nie tun. Tachistoskopisch dargebotene Wörter zu erkennen ist im Alltag genau so unüblich wie ein vollständiger Paarvergleich oder eine weitgehend leere Landkarte, in die man mehr oder weniger gut bekannte Städte einzeichnet. Bedeutet dies, daß Befunde aus solchen Experimenten für praktische Entscheidungshilfen wertlos sind? Offensichtlich nicht. Man kann die Beschreibungsebene wechseln, also Darbietungszeiten im Tachistoskopversuch als Indikatoren erschwerter Wahrnehmbarkeit, Paarvergleiche als spezifische Konkretisierung von Entscheidungen bei wechselnden und überlappenden Optionenmengen, und leere Landkarten als Komponenten einer Indikatorsituation für subjektive Repräsentationen räumlicher Beziehungen auffassen. Solche Übertragungen von Aussagen von einem auf einen anderen Gegenstandsbereich unter Nutzung eines geeigneten Abstraktionsniveaus sind Analogieschlüsse. Zu einer Methodologie, die sich um bessere Aufklärung der Zusammenhänge zwischen **Forschungs- und Anwendungsmethodik** bemüht, gehört also sicher auch die Analyse von **Analogisierungen**.

Im zweiten Fall sind sowohl Verhaltensweisen der Versuchspersonen als auch Bedingungskonstellationen aus der Untersuchungssituation im Alltag wiederzufinden; dennoch sind die Kriterien externer Validität nicht erfüllt. Beim assoziativen Lernen weitgehend sinnlosen Materials im Labor findet man oft Daten, die den Angaben eines einfachen stochastischen Modells mit linearen Operatoren, dem sogenannten „a-Modell“ (Bush & Mosteller, 1951), gut entsprechen. Betrachtet man ähnliche Lernprozesse im Alltag, etwa das Lernen von Vokabeln einer noch weitgehend unbekanntes Sprache, dann sehen die Daten völlig anders aus. Hier gibt es Plateaus in den Lernkurven, die im Labor registrierte negativ beschleunigte Exponentialfunktionen nicht kennen. Dennoch wäre es aberwitzig zu behaupten, die in den Labordaten zum Aus-

druck kommenden psychischen Mechanismen seien im Alltag schlicht abwesend; stattdessen wird man Erklärungen suchen, denen zufolge andere Prozesse (im Beispielfall Superierung und verwandte Formen der kognitiven Reorganisation von Lernmaterial) zu Überlagerungen einer basalen Einprägungsmechanik führen, aus denen sich die beobachtbaren Abweichungen ergeben. Das auf Dingler (1928, 1955) zurückgehende und in der Psychologie vor allem von Holzkamp (1964) propagierte Verfahren des Exhaurierens, das ursprünglich eine Theorie durch Einführung „störender Bedingungen“ gegen unpassende Daten verteidigt, ist entsprechend verwendbar, um Daten aus einer manipulativ kontrollierten Laborsituation gegen Alltagserfahrungen zu verteidigen. Dabei darf man allerdings nicht übersehen, daß auch die Behauptung störender Bedingungen wiederum der empirischen Kontrolle bedarf.

Externe Validität ist somit kein Gütekriterium an sich; sie sagt lediglich etwas darüber, aus, wie weit von Befunden **direkt** auf Aussagen über Verhalten in Alltagssituationen geschlossen werden kann. **Anwendung von Wissenschaft auf Alltagsfragen** setzt voraus, daß man die jeweils zu problematisierende Situation auf Komponenten untersucht, zu denen wissenschaftliche Aussagen verfügbar sind, und daß man die so gewonnenen Aussagen derart in Beziehung setzt, daß eine partielle Repräsentation der Alltagssituation entsteht. Diese Repräsentation kann dann als Aussagensystem genutzt werden, aus dem Entscheidungshilfen ableitbar sind. Die derzeit übliche Psychologie hat oft einen überstark vereinfachend praxeologischen Zug, der wissenschaftliche Aussagen ohne Anpassung an die Eigenheiten der jeweiligen Alltagssituation direkt anzuwenden bemüht ist. Ein Physiker würde kaum untersuchen, wie weit er Ablösungszeitpunkt und Fallbahn eines Ahornblattes im Herbstwind vorherzusagen vermag, und auch in der Biologie umfaßt die Untersuchung des Verhaltens in freier Wildbahn nur einen recht geringen Teil der Forschungslandschaft. Viele Psychologen streben dagegen nach einer Forschung, deren wissenschaftliche Ergebnisse im Alltag nach Art einer Ahornblatt-Physik **direkt** anwendbar sind.

In diesem Zusammenhang ist auch die altherwürdige Kritik am **Experiment** zu erwähnen. Das Experiment setzt voraus, daß man recht genau weiß, welche Bedingungsvariation erforderlich ist. Das ist notwendig zur Realisierbarkeit manipulativer Kontrolle. Ist die Fragestellung einer Untersuchung derart vorstrukturiert, dann folgt meistens auch, welche Verhaltensaspekte zu registrieren sind. Das aber bietet die Möglichkeit zur Datenrestriktion. Hoher Manipulationsgrad und starke Datenrestriktion sind im psychologischen Experiment kombiniert; dies mag einer der Gründe dafür sein, daß sie in der Literatur nicht immer klar getrennt werden.

Hermann (1990) setzt sich mit Argumenten auseinander, die immer wieder gegen das Experiment in der Psychologie ins Feld geführt werden. Dabei zeigt

sich, daß solche Argumente oft erst dadurch möglich werden, daß man an das Experiment Forderungen stellt, die seinem eigenen (oft wesentlich bescheideneren) Anspruch nicht angemessen sind. Auch die Rede von der externen Validität ist von dieser Art; sie unterstellt eine direkte Anwendbarkeit von Befunden im Alltag, ohne nach den Methoden einer angemessenen Analyse von Alltagssituationen und einer entsprechenden Synthese wissenschaftlicher Befunde zu einem praktisch nutzbaren Aussagensystem zu fragen.

Wir lassen offen, wie man überhaupt von einer Substanztheorie ausgehend zu Aussagen kommen kann, die einer direkten empirischen Kontrolle zugänglich sind, wie man also empirische Daten in einem größeren wissenschaftlichen Zusammenhang verwenden kann. Hiermit beschäftigt sich Erdfelder im 2. Kapitel dieses Buches. Auch in der Psychologie zur Zeit eher ungewöhnliche Arten des Umgangs mit empirischen Daten sollen hier nicht genauer vorgestellt werden. Dazu gehören etwa Formen einer „mittleren Datenrestriktion“, die nicht die volle Spielbreite einer freien Interpretierbarkeit eröffnen, wohl aber unterschiedliche Auswertungsmöglichkeiten im Zusammenhang mit verschiedenen Fragestellungen. Die Konsequenz ist, solche Daten in einem „Corpus“ verfügbar zu machen, der für wechselnde Zwecke genutzt werden kann. Ein Beispiel hierzu liefert die als „Ulmer Textbank“ (Mergenthaler, 1985; Kächele, 1986) bekannte Sammlung psychotherapeutischer Gespräche. Ebenfalls (noch) ungewöhnlich ist eine Form des empirischen Zugangs, der in kognitionswissenschaftlich orientierten Arbeiten an Bedeutung gewinnt. Kaum restringierte verbale Daten (etwa aus lautem Denken oder aufgabenbegleitenden Befragungen) werden als Äußerungen interpretiert, die bestimmte Wissensinhalte voraussetzen. Die resultierenden Interpretate werden in einem stark restringierten Repräsentationsformat niedergelegt und als Daten weiterer Verarbeitung zugeführt, indem sie beispielsweise in die Datenbasis eines wissensbasiert agierenden Modells eingebaut werden. Opwis und Spada zeigen im 5. Kapitel dieses Bandes Möglichkeiten der Modellierung kognitiver Prozesse durch wissensbasierte Systeme, bei denen neben Experimenten herkömmlicher Art auch empirische Zugänge eine Rolle spielen, die interpretierende und datenförmige Ansätze miteinander verknüpfen.

Überlegungen zu den methodischen Konsequenzen der Auffassung von Psychologie als einer **empirischen Wissenschaft** haben wir mit der Unterstellung eingeleitet, **es** müsse **empirische Methoden** geben, deren sie sich bedient. Die Betrachtung formaler und normativer methodischer Ansätze erwies diese allerdings ebenfalls als unverzichtbar; aus der Empirizität der Psychologie folgt keine Beschränkung auf den ausschließlichen Einsatz empirischer Methoden. Der mit analytischen Methoden möglichen rationalen Erkenntnis haben wir dann allerdings im Kontext empirischer Wissenschaft einen **heuristischen** Status zugesprochen, bei dem Ergebnisse nicht als letztbegründet akzeptiert werden und daher einer Überprüfung bedürfen, für die **empirische Methoden** un-

verzichtbar sind. Zusammengefaßt: In der Psychologie als empirischer Wissenschaft ist der Einsatz empirischer Methoden nicht hinreichend, aber **unverzichtbar**. Die Vielfalt der Fragen nach Kriterien akzeptabler empirischer Methoden, die nicht zuletzt wegen der skizzierten zentralen Stellung der Empirie in der Psychologie besonderer Beachtung bedürfen, werden in diesem Kapitel nicht weiter verfolgt, da sie ohnehin Gegenstand der weitaus meisten nachfolgenden Beiträge dieses Bandes sind.

3. Komplexität

Die für Psychologie relevanten methodologischen Fragestellungen haben oft damit zu tun, daß Psychologie als eine empirische Wissenschaft - in einem sehr breiten und für zahlreiche Varianten offenen Sinne des Wortes - aufgefaßt wird. Ihre besondere Tönung erhalten sie durch zusätzliche Spezifika, in denen sich Psychologie von anderen, ebenfalls empirischen Wissenschaften unterscheidet. Hierzu gehört unter anderem die oft beschworene große Komplexität des Phänomenenbereiches.

Außerhalb der Komplexitätstheorie, bei der es um den zur Berechnung berechenbarer Funktionen bzw. zur Entscheidung entscheidbarer Aussagen erforderlichen Aufwand geht (siehe Paul, 1987), ist die Bedeutung von „Komplexität“ vielfältig und unklar. De Groot (1969, S. 327-342) analysiert die Komplexität der zu untersuchenden **Phänomene** wie auch der **Instrumente** (Methoden, Theorien, Modelle), derer sich die Psychologie bedient. Breuer (1977, S. 20) spricht von der „Komplexität der Systembedingungen“ des **Objekts** der Psychologie, die diese unangemessen vereinfache, und nennt Beispiele für derartige Vereinfachungen ohne zu explizieren, was denn Komplexität nun sein möge. Die Endlichkeit menschlichen Wissens - und der für wissenschaftliche Publikationen verfügbaren Seitenzahlen - bringt stets Produkte hervor, die nicht alle Aspekte und Möglichkeiten berücksichtigen, die zur Lösung eines Problems oder zur Analyse eines Sachverhaltes hätten herangezogen werden können. Mangelnde Berücksichtigung von Komplexität ist so ein stets wohlfeiles Argument gegen den jeweils anderen.

3.1 Komplexität von Systemen

Simon (1977) und auch Nelson (1977) bemühen sich um eine Präzisierung der Rede von der Komplexität eines **Systems**. Danach wächst Komplexität mit der Anzahl der Systemkomponenten, mit der Anzahl und dem Ausmaß der Verschiedenheit unterscheidbarer Komponentenklassen und mit der Zahl der Relationen zwischen ihnen. Aus dieser Sicht kommt Komplexität nur einem

System zu, das Ergebnis der Analyse einer Gegebenheit auf einem bestimmten Auflösungsgrad ist; es geht also in der Psychologie weniger um die Komplexität des Menschen als vielmehr um die eines angemessenen Menschenbildes.

3.1.1 Komplexitätsvergleich

Diese Konzeption von Komplexität ist zunächst komparativ; man kann lediglich unter bestimmten Bedingungen bei zwei Systemen ausmachen, welches **komplexer als** das andere ist. Nelson (1989, S.61-73) verdeutlicht dies am Beispiel von „Peter the sleeper“. Peter ist ein relativ einfach funktionierendes Wesen. Es gibt für ihn nur zwei Zustände: Entweder schläft er, oder er ist wach. Es gibt zwei Reize, auf die er reagiert: das Klingeln eines Weckers und ein Stoß zwischen die Rippen. Die Kombinationen ergeben vier für ihn relevante Situationen: „Klingeln mit Stoß“, „Klingeln ohne Stoß“, „Stoß ohne Klingeln“, „weder Klingeln noch Stoß“. Wenn er schläft, kann ihn nur ein Stoß zwischen die Rippen in den Wachzustand überführen. In diesem Fall reagiert er mit Gähnen, sofern er nur einen Rippenstoß erhalten hat, während er beim gleichzeitigen Weckerklingeln zuschlägt. Ist er wach, dann gähnt er und schläft prompt wieder ein, sofern er weder gestoßen noch durch Weckerrasseln wachgehalten wird. Unter allen anderen Bedingungen bleibt er wach und reagiert mit irgendwelchen streng festgelegten Kombinationen seiner Möglichkeiten „Zuschlagen“ und „Gähnen“.

„Peter the sleeper“ ist ein deterministischer endlicher Automat. Damit zumindest einige von Peters Verhaltensmöglichkeiten zum Tragen kommen können, muß er einen Wecker erhalten, der klingeln kann, und einen Zimmergenossen, der ihm einen Rippenstoß versetzen kann. Zwischen den drei Komponenten des so entstehenden Superautomaten müssen bestimmte Beziehungen gegeben sein. man könnte etwa daran denken, daß der Schlag, den Peter manchmal ausführt, den Zimmergenossen trifft, sofern dieser überhaupt im Zimmer ist, und ähnliches mehr. Die Kombination aus Peter, dem Wecker und dem Zimmergenossen ist immer noch ein endlicher deterministischer Automat, aber von höherer Komplexität als „Peter the sleeper“, da neue Komponenten anderer Art und vorher überhaupt nicht mögliche Relationen zwischen den drei kombinierten Komponenten hinzugekommen sind.

3.1.2 Komplexitätsreduktion

Ein solcher Aufbau immer komplexerer Systeme durch Kombination vorhandener Komponenten dient hier lediglich der beispielhaften Erläuterung des Komplexitätskonzepts. Für das Vorgehen einer empirischen Wissenschaft ist

der umgekehrte Weg interessanter. Empirisch vorfindbare Phänomene werden erklärt, indem man das zugrundeliegende System als ein Gefüge aus Komponenten auffaßt. So wird der menschliche Geist zu einer Struktur aus sensorischen Puffern, diversen Lang- und Kurzzeitgedächtnissen, motorischen Programmen, Motivsystemen, Zielstrukturen, Bewertungsfunktionen, Entscheidungseinheiten, und ähnlichem mehr. Erst in der Darstellung als Kombination von weniger komplexen Teilsystemen erhält er im Vergleich mit eben diesen Teilsystemen seine oft beschworene „Komplexität“.

Palmer und Kimchi (1986) betrachten eine solche „Komplexitätsreduktion“ als charakteristisch für jeden Ansatz, der den Menschen als informationsverarbeitendes System auffaßt. Hier wird eine Operation, beispielsweise die Beantwortung einer Frage, als ein Gefüge aus Teiloperationen (syntaktische Analyse, Worterkennung, semantische Analyse, verschiedene Formen des Gedächtniszugriffs, Operationen der Zielverwaltung, Bildung einer Antwortabsicht, Umsetzung in Sprache, Realisierung eines motorischen Programms und so weiter) dargestellt; die Relationen zwischen diesen stehen für Parallelität und Sequentialität, Unterbrechungen, Reaktivierungen und ähnliches mehr. Selbstverständlich kann man auch Teiloperationen wieder in Komponenten zerlegen und so weiter und so fort. Auf diese Weise entsteht eine Kette von Komplexitätsreduktionen. Sie beginnt mit dem zu analysierenden Phänomenbereich, über dessen Komplexität wir zunächst nicht aussagen, und führt zur Reduktion auf immer mehr und immer einfachere Komponenten, deren größere Einfachheit nichts weiter widerspiegelt als das triviale Vergleichsergebnis, daß sie gegenüber dem Ausgangsphänomen als weniger komplex auszeichnet.

Eine solche Sicht der Dinge legt die Frage nahe, ob es „einfachste“ Operationen gebe, bei denen die Komplexitätsreduktion endet. Ein hemmungsloser Reduktionist wird vielleicht darauf hinweisen, daß alle geistigen Prozesse einer neurophysiologischen Basis bedürfen, daß somit jeder Fortschritt der Neurophysiologie ein Schritt näher zum Ende der Psychologie sei, daß Neurophysiologie ihrerseits Vorgänge analysiere, die sich biochemisch erklären lassen, daß Biochemie nur ein Teilbereich der Chemie schlechthin sei, daß deren Aussagen im Endeffekt auf Gesetze der Teilchenphysik reduzierbar seien und daß jede konsequent fortgesetzte Komplexitätsreduktion schließlich hier anlangen müsse. Ernsthaft vertreten wird ein solcher rigoroser Reduktionismus kaum; das reduktionistische Programm ist unrealistisch, weil es nicht nur einen sehr hohen Stand der jeweiligen Basiswissenschaft voraussetzt, sondern auch die Beschäftigung mit den Gefügen von Relationen zwischen Komponenten, die zum Gegenstandsbereich der Basiswissenschaft gehören. Psychologie könnte erst dann durch Neurophysiologie ersetzt werden, wenn diese bestimmte Klassen neuronaler Aktivitätsmuster (von denen die eine vielleicht die Klasse der möglichen physiologischen Implimentationen dessen ist, was wir „Fru-

stration“ nennen) daraufhin untersucht, welche Effekte die Anwesenheit eines Elements einer solchen Klasse auf die Realisierung eines Musters hat, das zu einer anderen Klasse gehört (deren Elemente beispielsweise für individuelle Varianten von „Aggression“ stehen). Dies aber wurde im Endeffekt bedeuten, daß eine derartige Neurophysiologie auch ursprünglich psychologische Konzepte untersuchen wurde; es läge keine Reduktion, sondern lediglich eine neuartige institutionelle Verknüpfung vor.

Diese Argumentation basiert auf der Vorstellung, daß bei (gegenüber ihren Komponenten) komplexeren Systemen Regelhaftigkeiten auftreten können, bei denen Relationengefüge auf Komponenten-Teilmengen relevant sind. Ein solches Relationengefüge ist dann eine „emergente Eigenschaft“ (Nagel, 1961, S. 366-380), auf die nicht verzichtet werden kann, wenn die in Frage stehende Regelhaftigkeit darstellbar sein soll. Damit aber mißlingt das radikale reduktionistische Programm.

Die Frage nach dem möglichen Ende einer Komplexitätsreduktion ist so zu modifizieren, daß die Möglichkeit kleinster Einheiten bei Systemdarstellungen zu erkunden ist, die im Rahmen einer Wissenschaft oder einer Teildisziplin bleiben. Für das Gemischtunternehmen „Psychologie“ liegt hierzu kein überzeugender Antwortentwurf vor. Nun macht zur Zeit innerhalb des Faches der „kognitive“ Ansatz immer mehr von sich reden, dessen fachübergreifende Komponente einer allgemeinen „Kognitionswissenschaft“ (siehe etwa Posner, 1989) zuzurechnen ist. Kognitionswissenschaft stellt den Anspruch, intelligentes Verhalten mit den dazu erforderlichen Strukturen und Prozessen so darzustellen und zu analysieren, daß die über verschiedene konkrete Realisierungen hinweg invarianten Aspekte erfaßt werden. Dabei ist bei den Realisierungen vorwiegend an intelligentes Verhalten von Menschen einerseits und das ebenfalls als „intelligent“ bezeichnete Verhalten von Rechnern andererseits gedacht. Die übergreifend abstrahierende Struktur ist die **kognitive Architektur**, die menschen- und/oder maschinenspezifische Realisierung deren **Implementation**. Diese Differenzierung zwischen kognitiver Architektur und Implementation macht es möglich, nach „kleinsten Einheiten“ zu fragen, die noch nicht Implementationskomponenten sind. Dabei läßt sich diese Frage verschieden wenden. Betrachtet man intelligente Systeme als wissensbasierte Agenten (wie von Opwis und Spada in Kap.5 dieses Bandes genauer dargestellt), so geht es um Eigenschaften **elementarer Wissenseinheiten**. Zerlegt man (wie Palmer und Kimchi, 1986) informationsverarbeitende Operationen in Komponenten-Operationen, **so** geht **es** um **elementary information processes (eips)**’ sensu Newell und Simon (1972).

3.2 Wahrscheinlichkeiten, Fehler, Störvariablen

Im alltäglichen Wissenschaftsbetrieb wird man bei einer systemorientierten Betrachtung die Analyse durch Einführung von Komponentengefügen irgendwann beenden, wohl wissend, daß man weitermachen könnte. Breite Bereiche psychologischer Forschung denken überhaupt nicht an eine systemorientierte Komplexitätsreduktion, sondern betrachten Abhängigkeiten und Zusammenhänge zwischen empirisch manipulierbaren und/oder registrierbaren Variablen, ebenfalls wohl wissend, daß eine feinere Analyse möglich wäre. Dies zeigt eine andere Variante des Komplexitätsproblems. Was auch immer in einer Untersuchung oder einem Forschungsprogramm erfaßt und berücksichtigt wird, es besteht stets Grund zu der Annahme, daß die beobachteten und zu erklärenden Phänomene auch von anderen nicht-kontrollierten Gegebenheiten abhängen oder modifiziert werden, die erst auf einem feineren Auflösungsgrad faßbar wurden. Es gilt also, die Möglichkeit unbekannter relevanter Variablen zu berücksichtigen. Hier kommen Konzeptionen von Wahrscheinlichkeit, Zufall und Fehler ins Spiel.

3.2.1 Störvariablen

Betrachten wir - ohne Verlust der Allgemeinheit und lediglich zur Vereinfachung - eine Untersuchungssituation mit nur **einer** reellwertigen abhängigen Variablen. Wir untersuchen, welche Werte diese Variable annimmt, wenn die Werte einer bestimmten Kombination unabhängiger Variablen systematisch variiert werden. Ohne allzu überrascht zu sein, stellen wir fest, daß bei wiederholter Realisierung der gleichen Wertekombination der unabhängigen Variablen verschiedene Werte der abhängigen Variablen resultieren. Wir führen dies auf den Einfluß **nicht-kontrollierter Variablen** zurück, deren Existenz angesichts der Komplexität, die sich bei weiterer komplexitätsreduzierender Analyse des untersuchten Systems ergeben würde, plausibel ist. Dazu nehmen wir an, daß die aus der skizzierten Variation resultierende Varianz, die wir gern „Fehlervarianz“ nennen, durch Kontrolle zusätzlicher relevanter Variablen verringert werden kann. Letztendlich nutzen wir diese Überlegung methodisch zur Entwicklung von Versuchsplan-Techniken und zur Einführung korrespondierender statistischer Verfahren der Datenauswertung.

Bei einer solchen Art von Begründung statistischer und versuchsplan-technischer Methoden, die von Rehm und Strack im Kapitel 12 dieses Bandes ausführlicher vorgestellt werden, ist zu berücksichtigen, daß die oft als **Störvariablen** bezeichneten unbekannt relevanten Variablen sowohl bei den Beschreibungsgrößen des menschlichen Geistes zu suchen sind wie auch im System der uns umgebenden äußeren Welt. Erstere versucht man gern durch

Versuchspersonenauswahl oder Erhebung als Kovariate zu kontrollieren, letztgenannte durch Manipulation.

Existieren Störvariablen, dann ist bei gleicher Wertekombination der kontrollierten unabhängigen Variablen mit unterschiedlichen Werten der abhängigen Variablen zu rechnen. Gleiche Werte der abhängigen Variablen bei gleicher Wertekonstellation der unabhängigen Variablen kommen kaum noch vor und können für die Einführung eines Validitätskriteriums nicht mehr verwandt werden. An ihre Stelle tritt die **Unterschiedlichkeit** zwischen den Werten abhängiger Variablen bei Realisierung **verschiedener** Werte der unabhängigen Variablen. Könnten wir Paare von Untersuchungssituationen herstellen, bei denen mit Sicherheit alle nicht kontrollierten Bedingungsaspekte - also auch die Werte aller Störvariablen - gleich wären, während sich die Werte der kontrollierten unabhängigen Variablen in bekannter Weise unterscheiden, so wäre zu fordern, daß die resultierenden **Unterschiede** auf der abhängigen Variablen - also die „**Effekte**“ kontrollierter Variation - stets gleich sind. Dies ist eine sehr strenge Version der Vorstellung, daß alle Aussagen über empirisch registrierbare Effekte stets unter dem Vorbehalt stehen, alles nicht kontrollierte Sonstige (cetera omnia) sei gleich (pars). Diese - oft stillschweigende - Zusatzannahme ist die ‚ceteris paribus‘-Klausel in ihrer wohl strengsten Variante.

Nun leidet auch diese Vorstellung darunter, daß sie in vielen Fällen weder realisierbar noch überzeugend approximierbar ist. Es bietet sich daher an, den Begriff des „Sonstigen“ etwas anders zu fassen, und an die Stelle der Gleichheit nicht-kontrollierter Bedingungen die schwächere Forderung nach deren gleicher Verteilung zu setzen. An dieser Stelle braucht man Zusatzannahmen, aus denen prüfbare Aussagen über die als „Fehler“ bezeichneten Abweichungen registrierter Werte vom Mittelwert bei der jeweiligen kontrollierten Bedingungskombination ableitbar sind. Die strengste Möglichkeit wäre, den kombinierten Effekt aller Störvariablen auf die abhängige Variable als unabhängig vom und additiv zum kombinierten Effekt der kontrollierten Bedingungsaspekte anzunehmen. Dann wären die Fehler über alle kontrollierten Kombinationen hinweg stets gleich verteilt. Selbstverständlich sind auch schwächere - und in vielen Fällen realistischere - Annahmen möglich. Weitere Ausführungen zu sinnvoll probabilisierten ceteris paribus-Varianten und daraus resultierenden Fehlermodellen findet man bei Steyer im Kapitel 15 dieses Bandes.

3.2.2 Konfundierung und Effektunterdrückung

Zwei Fragen, die in diesem Zusammenhang bislang offen geblieben sind, seien kurz angesprochen. Erstens: Es ist nie auszuschließen, daß auch schwache Forderungen nach Unabhängigkeit zwischen Störvariablen und kontrollierter

Bedingungsvariation nicht erfüllt sind. Die Differentielle Psychologie ist voll von Nachweisen über Korrelationen zwischen menschlichen Eigenschaften. Wird also bei einer psychologischen Untersuchung Intelligenz, Ängstlichkeit, Neurotizismus, Ich-Stärke oder was auch immer durch entsprechende Probandenauswahl kontrolliert, so müssen wir damit rechnen, daß entsprechend den Korrelationen anderer Persönlichkeitsvariablen mit den kontrollierten Größen abhängig von der jeweils realisierten Bedingungsvariation diese anderen Variablen unterschiedlich verteilt sind. Sind sie - was zunächst unbekannt ist - als Störvariablen wirksam, kann jeder gefundene Effekt statt auf die kontrollierten Größen auf eine Kombination solcher **konfundierter Variablen** zurückgeführt werden. Es wurde bereits erwähnt (S.22), daß auch die Abwesenheit konfundierter Störvariablen als eine der Forderungen interner Validität angesehen wird. Die Erfüllung dieser Forderung kann allerdings nicht effektiv festgestellt werden, weil dazu die Nicht-Existenz einer konfundierten Störvariablen in einer offenen Population definierbarer Bedingungskomponenten nachzuweisen wäre.

Die zweite Frage: Die gesamte Konzeption der „Fehler“ und „Störvariablen“ basiert auf der Annahme, daß nicht kontrollierte Aspekte einer Bedingungsvariation - sofern sie einen Einfluß auf die jeweils abhängigen Variablen haben - stets zu einer **Varianzerhöhung** führen, wobei die Störvariablen mit **Varianzquellen** identifiziert werden. Es ist aber durchaus denkbar, daß nicht-kontrollierte Bedingungsaspekte beispielsweise zu einem Korrekturmechanismus führen, der unter anderen Bedingungen manifeste Effekte überdeckt. So könnte etwa die oft gescholtene Frustrations-Aggressions-Hypothese durchaus richtig sein; in allen Untersuchungen, in denen Daten gegen sie sprechen, führen irgendwelche nicht identifizierten Aspekte, die etwa mit Angst vor den Folgen einer aggressiven Handlung, mit Vorstellungen vom sozial erwarteten Verhalten, mit der Überlagerung durch andere Motivsysteme oder mit ähnlichem zu tun haben, dazu, daß eine durchaus existente frustrationsbedingte Aggressionstendenz in den registrierten Daten nicht manifest wird. Um Mißverständnisse zu vermeiden: Ich behaupte nicht, daß dies so sei, sondern lediglich, daß wir eine solche Möglichkeit im Kontext der Fehler/Störvariablen-Methodik überhaupt nicht berücksichtigen.

Es wäre daher auf den zweiten Blick gar nicht so abwegig, eine statistische Methodik zu entwickeln, deren Verfahren nicht - wie bei der üblichen Signifikanzstatistik - der Entscheidung über eine Hypothese dienen, sondern vielmehr der Entscheidung darüber, ob **fehlerhafte Daten** vorliegen, bei denen störende Variablen nicht zur Varianzerhöhung, sondern zur **Effektüberdeckung** geführt haben. Tradition hätte ein solches Vorgehen. Frühe Anwendungen statistischer Techniken in der Astronomie dienten nicht der Entscheidung über Theorien oder Hypothesen, sondern der Kontrolle der Brauchbarkeit einzelner Daten (Swijtink, 1987). War ein Datum aufgrund der

Erwartung extrem unwahrscheinlich, wurde nicht Erwartung, sondern das Datum als fehlerhaft verworfen.

3.2.3 Begründung probabilistischer Konzepte

Man kann davon ausgehen, daß probabilistische Methoden genutzt werden, um mit mangelndem Wissen über Systemkomponenten fertig zu werden, die bei gegebenem Auflösungsgrad der Analyse nicht bekannt sind, mit deren Existenz aber zu rechnen ist. Man kann Wahrscheinlichkeiten als einen Ausdruck partieller Informiertheit benutzen. Wenn wir die Wahrscheinlichkeit der Entnahme einer roten Kugel aus einer Urne mit 50 roten und 50 schwarzen Kugeln als $\frac{1}{2}$ bezeichnen, setzen wir voraus, daß die Person, für die diese Wahrscheinlichkeit gilt, nicht in die Urne hineinschaut und sieht, daß alle schwarzen Kugeln unten und alle roten Kugeln oben liegen, denn dann wäre - bei Entnahme von oben - die Wahrscheinlichkeit einer roten Kugel wohl besser bei 1 anzusiedeln. Wahrscheinlichkeiten werden also als Relationen zwischen Sachverhalten und Informiertheiten interpretiert.

Diese Sichtweise bezeichnen Gigerenzer und Murray (1987, S.xii) als die **epistemische Interpretation** probabilistischer Ideen. Zufalls- und Wahrscheinlichkeitsannahmen bilden Grenzen unserer Erfahrung ab. Diese Sicht sollte aber keinesfalls dahingehend mißverstanden werden, als seien alle Zufallsvorstellungen in Theorien und Methoden der Psychologie ausschließlich aus der Diskrepanz zwischen Komplexität und eingeschränkter Erfahrung abzuleiten. Man kann auch die Meinung vertreten, es gebe Prozesse, die als nicht weiter reduzierbare natürliche Phänomene inhärent probabilistisch sind, also auch ohne unsere bekannten Wissenslücken nicht anders als durch Wahrscheinlichkeitsangaben beschrieben werden können. Dies ist die **ontische Interpretation** probabilistischer Konzepte.

Schließlich und endlich gibt es auch noch Wahrscheinlichkeitskonzepte, die Relationen zwischen empirischen Aussagen über endlich viele Fälle und allgemeinen Hypothesen, Annahmen und Gesetzen beschreiben, die die Form von Allaussagen über offenen Populationen haben. Solche Wahrscheinlichkeiten haben wenig mit der Beschreibung von Phänomenen oder einem begrenzten Wissen darüber zu tun, um so mehr aber mit Formen und Möglichkeiten induktiver Folgerungen (siehe Westermann und Gerjets in Kap. 10 dieses Bandes). Hierher gehört auch die Verwendung probabilistischer Konzepte bei der statistischen Entscheidung, deren Varianten in diesem Band von Ostmann und Wutke (Kap. 16) dargestellt und analysiert werden.

Komplexität mit den methodischen Konsequenzen der systemorientierten Komplexitätsreduktion und der Verwendung probabilistisch begründbarer Methoden sind nicht die einzigen Psychologie-Spezifika, die eine fachspezi-

fische Allgemeine Methodenlehre berücksichtigen muß. Da gibt es etwa noch die Frage nach den Kriterien der Angemessenheit einer Methode an eine Fragestellung, die ihre Tönung durch die Eigenschaften des Bereichs der Psychologie-üblichen Forschungsprogramme erhält, oder die Frage nach der Eindeutigkeit von Befundaussagen, die analog dem Eindeutigkeitskonzept der Meßtheorie behandelt werden kann, mit dem sich in diesem Buch Mausfeld (Kap. 13) auseinandersetzt. Aus spezifischen Teilbereichen des Faches ergeben sich weitere methodologische Fragestellungen. So hat die Entwicklungspsychologie mit der Abgrenzbarkeit und empirischen Identifizierbarkeit verschiedener Arten von Entwicklungsverläufen zu tun und die Differentielle Psychologie mit der oft nicht unbeträchtlichen Schwierigkeit der Entscheidung darüber, ob ein Befund (nur) Aussagen über Populationen oder (auch) solche über Individuen erlaubt. Sind Faktoren der Intelligenz Beschreibungsgrößen einer Kovarianzstruktur in einer Population oder sind es auch Variablen, die für irgendetwas bei jedem Individuum stehen? Beschreibt die Differenzierungshypothese der Intelligenz eine altersabhängige Veränderung einer Variabilitätsstruktur in einer Population oder sagt sie etwas über irgendwelche Prozesse aus, die in jedem einzelnen ablaufen?

Solche und ähnliche Fragen liefern weitere Beispiele für Themen einer Allgemeinen Methodenlehre der Psychologie. Derartige Überlegungen sollen hier nicht weiter verfolgt werden; stattdessen zeigen wir am Beispiel der Themen dieses Enzyklopädie-Bandes Zusammenhänge auf und skizzieren eine erste Ordnung.

4. Themen Allgemeiner Methodenlehre

In den bisherigen Überlegungen und Ausführungen wurde versucht, Allgemeine Methodenlehre der Psychologie nicht durch eine - kaum sinnvoll mögliche - strenge Definition, sondern vielmehr durch Verweis auf möglichst typische Themen und Fragestellungen nach außen abzugrenzen, und zwar einerseits gegenüber einer Spezifischen Methodenlehre des Faches und andererseits gegenüber einer zwar Allgemeinen Methodenlehre, die aber keine solche der Psychologie ist. Es stellt sich nun die Frage nach möglichen Binnengliederungen einer Psychologie-Methodologie, die zumindest Arbeits- und Aufgabenbereiche aufzeigt.

4.1 Zugänge

Ein Blick auf die aktuelle Psychologie zeigt, daß es eine Vielzahl von Zugängen zum Gegenstands- und Phänomenbereich, von Forschungsmöglichkeiten und -Varianten gibt. Eine Suche nach der „richtigen Art, Psychologie zu be-

treiben“ (Grawe, Hänni, Semmer & Tschan, 1991) ist vergebliches Bemühen. Wir finden eine Vielzahl von Varianten psychologischer Forschung, die sich sowohl in den jeweils untersuchten Fragestellungen und Phänomenbereichen als auch in ihren Methodenarsenalen mehr oder weniger deutlich unterscheiden.

Allgemeine Methodenlehre kann solche Ansätze psychologischer Forschung darstellen und untersuchen, welche Themen und Fragestellungen jeweils mit welchen Methoden auf der Basis welcher Vorannahmen bearbeitet werden. Analog zur Abgrenzung einer Allgemeinen Methodenlehre der Psychologie nach außen erhält man so gewissermaßen eine Menge gegeneinander abgrenzbarer Allgemeiner Methodenlehren von XY-Psychologien. Hierauf können Arbeiten aufbauen, die etwa über die XY-Psychologien hinweg Korrespondenzen zwischen Fragestellungen, Rahmenkonzepten und Methodenvorräten analysieren. Teil B dieses Bandes mit den Kapiteln 2 bis 5 zeigt Spezifika einiger Zugänge der Psychologie zu ihren Gegenständen, die hier als Beispiele ohne jeden Anspruch auf Vollständigkeit dienen.

Erdfelder (Kap. 2) skizziert die Grundlagen eines Ansatzes, der **empirische Daten** generiert und nutzt. Dabei versteht er als ein „Datum“ nicht - wie bislang in diesem Kapitel - jede Aussage, von der allgemein unterstellt wird, sie gebe ein Beobachtungsergebnis wieder und für die keine weitere Begründung gefordert wird. Er unterscheidet vielmehr zwischen „Beobachtungsprotokollen“ und „Daten“ in einem engeren Wortsinne, die das Ergebnis der Abbildung von Beobachtungen in einen **Datenraum** sind. Dieser hinwiederum ist eine Menge möglicher Daten, die durch einen Satz von Variablen beschrieben werden kann, von denen jede bei jedem Einzeldatum mit einem Wert belegt ist. Welche Variablen dabei berücksichtigt werden, hängt von der jeweiligen Fragestellung ab, zu deren Beantwortung Daten erhoben werden. Daten sind bei Erdfelder stets soweit restringiert, daß sie möglichst genau die zur Beantwortung einer Fragestellung notwendige Information enthalten,

Von einem solchen „datenförmigen“ Vorgehen deutlich unterschieden sind **interpretierende Ansätze**, mit denen sich Soeffner und Hitzler im 3. Kapitel dieses Bandes beschäftigen. Das empirische Material besteht hier aus wenig restringierten Daten, die oft als (sprachliche wie auch nicht-sprachliche) „Dokumente“ bezeichnet werden. Der Vergleich zwischen dem 2. und dem 3. Kapitel macht deutlich, daß die Entscheidung für unterschiedliche Möglichkeiten der Datenrestriktion eng zusammenhängt mit verschiedenen Zielen wissenschaftlicher Forschung. Datenförmige Empirie, die Daten im Begründungszusammenhang verwendet, setzt „... ein psychologisches Gesetz, eine Hypothese, eine technologische Prognose oder eine psychologische Theorie ...“ (Erdfelder, Kap. 2 in diesem Band) voraus, die etwas über bestimmte Aspekte möglicher Daten aussagt. Damit kann der Datenraum auf eben diese Aspekte

restringiert werden. Bei Verwendung von Daten im Entdeckungszusammenhang gibt es „offene Fragestellungen“, in denen ebenfalls bestimmte Qualitäten beobachtbarer Sachverhalte vorkommen, auf deren Berücksichtigung der Datenraum restringiert werden kann. Sinnvolle Datenrestriktion wird also dadurch möglich, daß spezifizierte offene oder geschlossene Fragen Ausgangspunkt von Forschungen sind. Interpretation zielt dagegen auf Verstehen ab, auf die Entfaltung einer möglichst breiten Vielfalt von Sinn-Potentialen. Ausgangspunkt ist keine vorformulierbare Fragestellung, die einen Variablenrahmen vorgibt, sondern ein Dokument, für das Deutungsmöglichkeiten erschlossen werden sollen. Eine solche Forschungsaufgabe bietet nicht den geringsten Ansatzpunkt für eine wie auch immer geartete Datenrestriktion.

Zu den Spezifika der Psychologie gehört sicherlich, daß sie sich mit Beziehungen zwischen (äußeren) wahrnehmbaren Gegebenheiten und (inneren) Wahrnehmungsinhalten sowie zwischen (innerem) Erleben und (äußерem) Verhalten beschäftigt, also in der einen oder anderen Weise mit dem Leib-Seele-Problem befaßt ist. Weder datenförmiges noch qualitativ-interpretierendes Vorgehen sind Forschungsansätze, die spezifisch auf Fragen des Verhältnisses zwischen einem Innen und einem Außen abgestellt sind; dies gilt aber sehr wohl für die **Psychophysik**, die Mausfeld im 4. Kapitel skizziert. Zwei Fragenkomplexe charakterisieren Psychophysik: die Erforschung von Mechanismen der Wahrnehmung und die Frage nach der Meßbarkeit des Psychischen. Zum Methodenarsenal gehören Techniken der Darstellung und Analyse von Mechanismen der Wahrnehmung als Systeme, auf denen informationsverarbeitende Prozesse laufen. Eine solche Konzeption setzt voraus, daß eine zwischen einem Innen und einem Außen differenzierende Vorstellung gegeben ist, von der ausgehend die in physikalisch beschreibbaren (Reiz-)Gegebenheiten repräsentierte Information in ein kognitiv beschreibbares Format transformierbar wird. Die Frage nach der Meßbarkeit des Psychischen kann dann als Frage nach einer bestimmten Form der Repräsentierbarkeit mentaler Gegebenheiten im kognitiven Modell interpretiert werden.

Psychophysik liefert zum einen ein Beispiel für die Beschäftigung mit deutlich Psychologie-spezifischen Fragestellungen und daraus resultierende methodischen Konsequenzen, zum anderen schafft sie mit ihrem Anspruch nach Repräsentierbarkeit mentaler Gegebenheiten einen Übergang zum Kapitel 5 von Opwis und Spada über Möglichkeiten der Modellierung mit Hilfe **wissensbasierter Systeme**. Auch hier wird die Innen-Außen-Differenzierung thematisiert. Menschen verfügen über eine „interne, subjektive Repräsentation von Ausschnitten ihrer tatsächlichen oder einer vorgestellten Umgebung“ (Opwis und Spada, Kap. 5 dieses Bandes, Abschnitt 2). Die Frage ist, wie eine solche interne Repräsentation durch eine Symbolstruktur (als Konstruktion zweiter Ordnung) repräsentiert werden kann, auf der ein formales kognitives System operiert. Dabei soll zwischen dem Verhalten dieses artifiziellen Systems und

menschlichem Verhalten eine Analogiebeziehung bestehen, die es gestattet, das formale wissenschaftliche System als Modell des (natürlichen) mentalen Systems aufzufassen.

Diese vier Beispiele für Zugänge der Forschung zum Gegenstandsbereich der Psychologie mit unterschiedlichen Methodeninventaren lassen zumindest einige Zusammenhänge zwischen Forschungsansätzen und Methoden deutlich werden. Die Annahme, am Anfang eines Forschungsprozesses stehe eine bestimmte Art von Fragestellung (Kap. 2), legt eine angemessene Datenrestriktion nahe, während die Offenheit gegenüber einer Vielfalt von Deutungsmöglichkeiten (Kap. 3) sie geradezu ausschließt. Bestimmte Spezifika der Psychologie wie (in diesem Fall) ihre Beschäftigung mit einer Innen/Außen-Differenzierung spielen bei bestimmten Ansätzen keine Rolle (Kap.2), werden bei anderen teilweise relevant (Kap. 3) und rücken schließlich bei wieder anderen ins Zentrum der behandelten Fragestellungen (Kap.4 und 5). Die jeweils interessierenden Aspekte des (inneren) mentalen Systems werden als Variablen-Systeme (Kap. 4) oder als Symbolstrukturen in einem Wissensrepräsentations-Format (Kap. 5) erfaßt. Teil B dieses Bandes deckt so zumindest einen Teil der Variation ab, die Varianten des forschenden Zugangs zu psychologisch relevanten Phänomenen aufspannen.

4.2 Theorien und Programme

Die Darstellung und methodologische Analyse einzelner Forschungsansätze verstellt allzu leicht den Blick auf Aspekte, die den verschiedenen Zugängen - oder zumindest einer großen Mehrheit davon - gemeinsam sind. Allgemeine Methodenlehre erfordert mithin einen zweiten Themenbereich, der solche Gemeinsamkeiten behandelt und in diesem Band im Teil C exemplifiziert wird.

Hierher gehört zumindest das, womit sich wissenschaftliche Forschung in doppeltem Sinne beschäftigt: Einmal, indem sie es zum Ausgangspunkt ihrer Bemühungen nimmt, und zum anderen, indem sie es durch ihre Bemühungen bewertet, verändert oder gar neu konzipiert. Einer weit verbreiteten Vorstellung zufolge ist dieses „Etwas“ die Theorie. Forschung dient oft der Bewertung von Theorien; Forschung verändert häufig vorhandene Theorien oder führt gar zu einer Datenlage, die Neukonzeptionen nahelegt oder zumindest anregt. Liest man allerdings Artikel, in denen einzelne Untersuchungen dargestellt werden, so findet man oft einleitend statt eines Verweises auf eine spezifische zu prüfende Theorie Hinweise auf eine Art Thema und auf einige andere Untersuchungen, die ebenfalls diesem Thema zugeordnet sind und dabei ähnliche Konzeptionen benutzen wie der Verfasser der jeweils vorliegenden Arbeit. Untersuchungen werden damit eingeordnet in ein jeweils von mehreren Wissenschaftlern benutztes System von Annahmen über einen Ge-

genstandsbereich, in dem spezifizierbare und noch nicht beantwortete Fragen möglich sind. In Kapitel 6 dieses Bandes werden solche Systeme von Herrmann genauer abgegrenzt, als „**Forschungsprogramme**“ bezeichnet, vorgestellt und analysiert.

Bei der Identifikation konkreter spezifischer Forschungsprogramme geht es nicht ohne die Festlegung einer Korngröße. Bezeichnet man etwa die Forschungen zur Modularität des menschlichen Gedächtnisses (Engelkamp, 1990) als ein Forschungsprogramm, so ließe sich - bei größerer Granularität - die gesamte kognitiv orientierte Gedächtnispsychologie sicherlich auch als ein solches auffassen. Prinzipiell spricht nichts dagegen, jeden der im Teil B dieses Bandes dargestellten Ansätze oder sogar die gesamte Psychologie ein „Forschungsprogramm“ zu nennen. Von Bedeutung ist, daß eine solche Sicht bestimmte Darstellungs- und Analysemöglichkeiten eröffnet, die Herrmann (Kap.6 dieses Bandes) vorstellt und erläutert.

Theorien sind nun keine „kleinsten Forschungsprogramme“, sondern - wie Herrmann im 6. Kapitel dieses Bandes ausführt - „Mittel zur Lösung der in Forschungsprogrammen behandelten Probleme“. Im Detail gibt es eine Vielzahl von Versuchen, das zunächst recht unscharfe Konzept einer Theorie zu präzisieren. Gadenne zeigt im 7. Kapitel dieses Bandes solche Spezifikationsvarianten auf, problematisiert die Beziehungen zwischen Theorie und Empirie sowie zwischen Theorie und Realität, untersucht die Rolle der Theorie bei verschiedenen Arten von Erklärungen und weist darauf hin, daß es neben der Ansicht, Theorien seien stets als Aussagensysteme darstellbar, auch andere Theoriekonzeptionen - wie etwa den Strukturalismus (Sneed, 1971) - gibt.

Zu Theorien gibt es zwei methodologische Problemkreise: einen zur Frage, wie Theorien zustandekommen, und einen zweiten zu der Aufgabe, Theorien zu bewerten. Mit diesen beiden Komplexen befassen sich in diesem Band zwei Kapitel von Dörner (Kap. 8 zur Theoriebildung) und von Gadenne (Kap.9 zur Theoriebewertung). Dabei wird klar, daß keine Regeln die **Konstruktion von Theorien** ermöglichen und dabei die Brauchbarkeit ihrer Produkte garantieren. Es gibt aber sehr wohl Regeln zur Theoriekonstruktion, deren Produkte **möglicherweise** brauchbare Theorien sind. Regelsysteme, die angewandt auf eine Fragestellung ein konkretes Ergebnis produzieren, heißen auch „Produktionensysteme“; Regelsammlungen, deren Produkte lediglich möglichst oft (aber nicht immer) den Ausgangsanforderungen genügen, nennt man oft „Heuristiken“, wobei die Regeln selbst die „Heurismen“ sind. Dörner (Kap. 8 dieses Bandes) geht davon aus, daß wir lediglich über Heurismen der Theoriekonstruktion verfügen. Konsequenz ist, daß wir Methoden der Bewertung von Theorien (Gadenne in Kap.9) benötigen.

Zur **Theoriebewertung** gibt es eine Vielzahl von Kriterien. Dabei kann das Bewertungsergebnis je nach Kriterium von unterschiedlicher Art sein. Fragen

wir etwa nach der logischen Konsistenz einer Theorie, so ist diese gegeben oder nicht, kümmern wir uns um ihren Informationsgehalt, so ist dieser mehr oder weniger reichhaltig, und fragen wir schließlich gar danach, ob sie empirisch gültig ist oder nicht, so können wir unter günstigen Umständen vielleicht behaupten, sie sei mit höherer Wahrscheinlichkeit gültig als eine konkurrierende Theorie. Hinzu kommt das Problem, verschiedene Kriterien gegeneinander abzuwägen und eine multivariate Entscheidung zu treffen.

Da Psychologie - wie in Abschnitt 2 dieses Kapitels näher erläutert - in ihrem Selbstverständnis eine empirische Wissenschaft ist, spielt bei der Theoriebewertung die aufgrund von Daten mögliche empirisch basierte Bewertungskomponente eine entscheidende Rolle. Dies führt auf die Frage, ob und inwieweit überhaupt Schlüsse von endlichen Datensätzen auf Systeme von Allgemeinaussagen begründbar sind. Diesem Problem der **Induktion** ist das 10. Kapitel dieses Bandes von Westermann und Gerjets gewidmet. Alle Übergänge von Aussagen über einen limitierten Gegenstandsbereich zu Aussagen über einen größeren Bereich (die sogenannten „ampliativen Inferenzen“) sind insoweit unsicher, als sie die Richtigkeit der produzierten neuen Aussagen unter der Bedingung der Richtigkeit der Ausgangsbehauptungen nicht garantieren können. Damit gibt es keine (sicheren) induktiven Schlüsse, sondern lediglich (unsichere) induktive Argumente, für die allerdings Qualitätskriterien und rationale Begründungen existieren. Westermann und Gerjets untersuchen, welche Begründungen man zur Stützung induktiver Argumente heranziehen kann und welche Voraussetzungen dabei zu beachten sind.

Insgesamt ergibt sich aus den Ausführungen in Teil C dieses Bandes die Unsicherheit und Vorläufigkeit aller wissenschaftlichen Befunde und Ergebnisse. Nun ist es sicher trivial darauf hinzuweisen, daß Resultate wissenschaftlicher Forschung keine unumstößlichen Wahrheiten sind. Die Autoren der Kapitel aus Teil C untersuchen, aus welchen Bedingungen Unsicherheit und Vorläufigkeit resultieren, und nutzen die Ergebnisse dieser Analysen, um Möglichkeiten methodischer Kontrolle aufzuzeigen, die zu solchen Forschungsergebnissen führen, die mit vernünftigen Argumenten als akzeptabel und brauchbar (nach jeweils zu spezifizierenden Kriterien) anzusehen sind.

4.3 Forschungsschritte

Sowohl die Betrachtung verschiedener Zugänge psychologischer Forschung zu ihrem Gegenstandsbereich als auch die Analyse von Forschungsprogrammen, Theorien und damit zusammenhängenden Problemkreisen liefern relativ große Analyseeinheiten. Stärker mit methodischen Details befaßte Themen findet man, wenn einzelne Schritte des Forschungsprozesses und die dabei eingesetzten Methoden analysiert werden. Einen großen Teil aktueller psy-

chologischer Forschungen machen Arbeiten aus, in denen irgendwelche Hypothesen geprüft werden sollen. Es ist daher nicht verwunderlich, daß zur hypothesengeleiteten Forschung auch die meisten methodischen und methodologischen Arbeiten vorliegen. Dieser Band versucht dem gerecht zu werden, indem er einen Teil D enthält, in dem Forschungsschritte beim hypothesengeleiteten Vorgehen untersucht werden.

Die Orientierung der Aufeinanderfolge der einzelnen Kapitel von Teil D am Forschungsablauf ist offensichtlich. Es beginnt mit einer Hypothese (Hussy & Möller, Kap. 11). Bei der Planung von Untersuchungen zur Prüfung dieser Hypothese geht es um den Einsatz von Kontrolltechniken (Rehm & Strack, Kap. 12). Dabei wird häufig unterstellt, daß relevante Variablen meßbar sind und daß entsprechend sinnvoll interpretierbare Skalen zur Verfügung stehen (Mausfeld, Kap. 13). Auf der Basis erhobener Daten wird die jeweilige Ausgangshypothese geprüft (Erdfelder & Bredenkamp, Kap. 14). Dabei operiert man in den weitaus meisten Fällen mit Wahrscheinlichkeits- und Zufallskonzepten, benutzt also stochastische Modelle (Steyer, Kap. 15). Der Prozeß endet mit einer statistischen Entscheidung (Ostmann & Wutke, Kap. 16).

Da alle Beiträge aus Teil D dieses Bandes Teile eines Forschungsprozesses behandeln, die auf das gleiche Ziel ausgerichtet sind, gibt es notwendig zahlreiche Überschneidungen. Man kann nun einmal - um nur ein Beispiel zu erwähnen - Kontrolltechniken bei der Versuchsplanung nur einsetzen, wenn man die Prüflöge kennt und weiß, welche Kontrollen sie ermöglicht. Auf der anderen Seite macht es wenig Sinn, Verfahren der Hypothesenprüfung zu analysieren ohne Verweis auf Kontrollmöglichkeiten, die bereits bei der Versuchsplanung zu berücksichtigen waren.

Das in diesem Zusammenhang zentrale Konzept der **Hypothese** wird in den einzelnen Beiträgen von Teil D zunehmend differenziert. Hussy und Möller (Kap. 11) beginnen mit der Alltagssprachlichen Rede von „Vermutungen“ und zeichnen wissenschaftliche Hypothesen dadurch aus, daß ihren Inhalten ein Wahrheitswert zukommt und daß sie prinzipiell überprüfbar sind. Sie führen mehrere Differenzierungen ein, darunter jene zwischen einer theoretisch-inhaltlichen Hypothese, einer möglicherweise durch Operationalisierungen, operationale Reduktionen oder möglicherweise Vorgehensweisen abgeleiteten empirisch-inhaltlichen Hypothese, einer durch Präzisierung von Häufigkeits- oder Intensitätsangaben formulierbaren statistischen Hypothese und einer für statistische Entscheidungen benötigten Prüfhypothese. Während bei Hussy und Möller (Kap. 11) auch Vorhersagen über individuelle Ereignisse durchaus als „wissenschaftliche Hypothesen“ angesprochen werden, gehen Rehm und Strack (Kap. 12, Abschnitt 1) davon aus, daß es sich im allgemeinen um Vermutungen handelt, die sich auf raum-zeitlich begrenzte „offene Populationen“ beziehen. Sie betrachten bei ihren weiteren Überlegungen auch nur noch diese

Art von Hypothesen, und so tun es auch die Verfasser aller noch folgenden Kapitel. Bei Erdfelder und Bredenkamp (Kap. 14) spielt dann wiederum die Unterscheidung zwischen einer **psychologischen**, einer **empirischen** und einer **statistischen** Hypothese eine Rolle, wobei diese bereits bei Hussy und Möller (Kap. 11) vorgestellte Differenzierung hier ausführlicher spezifiziert wird. Dabei geht es insbesondere um die Frage nach der Relation zwischen psychologischen und jeweils zugeordneten statistischen Hypothesen. Wie kann man es begründen, statistische Entscheidungen über statistische Hypothesen als Grundlage für Bewährungsurteile über psychologische Hypothesen zu benutzen? Sowohl Steyer (Kap. 15) als auch Ostmann und Wutke (Kap. 16) gehen - der Themenstellung ihrer Beiträge entsprechend - vom Umgang mit probabilistisch formulierten Hypothesen aus, wobei Aussagen über Verteilungen und deren Parameter ein häufiger Spezialfall sind.

Zur Kennzeichnung der Qualität einer hypothesenorientierten Untersuchung wird oft der Ausdruck „**Validität**“ benutzt. Im Umgang mit diesem Ausdruck tauchen teils beträchtliche Bedeutungsunterschiede auf. Unter 2.3.3 wurde in diesem Kapitel bereits auf ein recht unscharfes Konzept der Validität einer Untersuchung bei Bortz (1984) hingewiesen, das dann in einer Weise präzisiert wurde - nämlich als Güte der Prognostizierbarkeit eines neuen Datums -, bei der Validität zum Merkmal eines Befundes wird. Hussy und Möller (Kap. 11) sprechen von verschiedenen Validitätskriterien, die teilweise unterschiedlichen Komponenten eines hypothesenorientierten Forschungsprozesses zugeordnet sind. So ist bei ihnen etwa die Rede von der „Ableitungsvalidität“ einer statistischen Hypothese (Kap. 11, Abschnitt 2.3.3), von einer „Populationsvalidität“, einer „ökologischen Validität“, einer „Situationsvalidität“ und einer „internen Validität“, die auch als „Ceteris-Paribus-Validität“ angesprochen wird (Kap. 11, Abschnitt 3). Erdfelder und Bredenkamp (Kap. 14) stellen ein auf Untersuchungen bezogenes komparatives Validitätskonzept vor. Danach ist Untersuchung **A** valider als Untersuchung **B**, wenn bei **A** sowohl die Wahrscheinlichkeit eines falschen Bewährungsurteils für die (in Wirklichkeit falsche) in Frage stehende psychologische Hypothese als auch die Wahrscheinlichkeit für ein falsches Nichtbewährungsurteil bei in Wirklichkeit richtiger Hypothese kleiner ist als bei **B**. Man sieht, daß derartige Vergleiche nur beschränkt möglich sind; in vielen Fällen kann man von zwei Untersuchungen nicht aussagen, welche von ihnen valider ist.

Selbstverständlich gibt es auch eine große Zahl von Konzepten, die in den Kapiteln zu Teil D dieses Bandes einheitlich behandelt werden. So wird von fast allen Autoren unterstellt, daß bei den zur statistischen Prüfung anstehenden Hypothesenformulierungen problemlos Verteilungsparameter benutzt werden können, die quantitative Variablen voraussetzen. Werte und Verteilungen quantitativer Variablen sind aber nur dann sinnvoll interpretierbar, wenn diese Variablen Skalen sind, die ausgewählte Aspekte bestimmter Ob-

jekte und Phänomene im jeweils betrachteten Realitätsbereich repräsentieren. Damit zusammenhängende Fragen analysiert Mausfeld im Kapitel 13. Einheitlichkeit herrscht auch bezüglich der Auffassung, Hypothesenprüfung müsse die Wahrscheinlichkeiten beider möglicher Arten von Fehlurteilen (falsches Bewährungsurteil bei falscher Hypothese, falsches Nichtbewährungsurteil bei richtiger Hypothese) berücksichtigen. Ostmann und Wutke (Kap. 16) stellen verschiedene Strategien der Hypothesentestung vergleichend nebeneinander. Sie analysieren sowohl die Vorteile einer Neyman-Pearson-Testung, die beide Fehlermöglichkeiten berücksichtigt, gegenüber der simpleren Fisher-Testung als auch die Möglichkeiten einer an Bayes orientierten Entscheidungsstrategie.

Insgesamt zeigt Teil D dieses Bandes beispielhaft auf, wie die Identifikation von Forschungsschritten zu Themen methodologischer Untersuchungen führen kann. Zusammen mit der Identifikation unterschiedlicher Zugänge psychologischer Forschung zu ihren Gegenständen und Phänomenen (Teil B) und der - orthogonal dazu konzipierten - Zuwendung zu gemeinsamen Gegebenheiten und Problemen (wie der Existenz von Forschungsprogrammen, der Arbeit mit und an Theorien und der Auseinandersetzung mit den Varianten des Induktionsproblems in Teil C) wird so in diesem Band ein erstes und sicher vorläufiges System von Themen aufgespannt, das die Reichhaltigkeit der Einzelfragen innerhalb einer Allgemeinen Methodenlehre der Psychologie erkennen läßt.

Literatur

- Agnew, N.M. & Pyke, S.W.(1969). **The science game. An introduction to research in the behavioral sciences.** Englewood Cliffs, N.J.: Prentice Hall.
- Anderson, J. R. (1991). Optimality and human memory. **Behavioral and Brain Sciences**, **14**, 215-216.
- Baron, J. (1991). Optimality as an evaluative Standard in the study of decision-making. **Behavioral and Brain Sciences**, **14**, 216.
- Bell, D. E., Raiffa, H. & Tversky, A. (1988). **Decision making: Descriptive, normative, and prescriptive interactions.** Cambridge: Cambridge University Press.
- Bortz, J. (1984). **Lehrbuch der empirischen Forschung für Sozialwissenschaftler.** Berlin: Springer.
- Bradley, R.A. & Terry, M.E. (1952). Rankanalysis of incomplete block designs. 1. The method of paired comparisons. **Biometrika**, **39**, 324-345.
- Bush, R. R. & Mosteller, F. (1951). A mathematical model for simple learning. **Psychological Review**, **58**, 313-323.
- Carnap, R. (1936/1937). Testability and meaning. **Philosophy of Science**, **314**.

- Chalmers, A.F. (1989). **Wege der Wissenschaft** (zweite, durchgesehene Auflage; Original erschienen 1976: What is this thing called science?). Berlin: Springer.
- Coombs, C. H. (1964). A theory of **data**. New York: Wiley.
- Coombs, C. H., Dawes, R. M. & Tversky, A. (1970). **Mathematical psychology: An introduction**. Englewood Cliffs, N.J.: Prentice Hall.
- Damerow, P., Freudenthal, G., McLaughlin, P. & Renn, J. (1992). **Exploring the limits of pre-classical mechanics. A study of conceptual development in early modern science: Descartes, Galileo, Beekman**. New York: Springer.
- Descartes, R. (1637/1984). **Abhandlung über die Methode des richtigen Vernunftgebrauchs** (französisches Original 1637). Stuttgart: Reclam.
- de Groot, A.D. (1969). **Methodology. Foundations of inference and research in the behavioral sciences**. Den Haag: Mouton.
- Dingler, H. (1928). **Das Experiment: Sein Wesen und seine Geschichte**. München: Reinhardt.
- Dingler, H. (1955). **Die Ergreifung des Wirklichen**. München: Eidos.
- Egart, J.P. (1958). **Recognition memory and the operating characteristic** (Technical Note AFCRC-TN-58-51). Indiana: University Hearing and Communication Laboratory.
- Engelkamp, J. (1990). **Das menschliche Gedächtnis**. Göttingen: Hogrefe.
- Feyerabend, P. (1976). **Wider den Methodenzwang. Skizzen einer anarchistischen Erkenntnistheorie**. Frankfurt: Suhrkamp.
- Feyerabend, P. (1980). **Erkenntnis für freie Menschen** (veränderte Ausgabe). Frankfurt: Suhrkamp.
- Gigerenzer, G. & Murray, D.J. (1987). **Cognition as intuitive statistics**. Hillsdale, N.J.: Erlbaum.
- Grawe, K., Hänni, R., Semmer, N. & Tschan, F. (Hrsg.) (1991). **über die richtige Art, Psychologie zu betreiben**. Göttingen: Hogrefe.
- Herrmann, T. (1990). Die Experimentiermethodik in der Defensive? **Spruche und Kognition, 9**, 1-11.
- Holzkamp, K. (1964). **Theorie und Experiment in der Psychologie**. Berlin: de Gruyter.
- James, W. (1890). **Principles of psychology** (Volume I and II). New York: Holt.
- Kächele, H. (1986). **Maschinelle Inhaltsanalyse in der psychoanalytischen Prozeßforschung**. Ulm: PSZ-Verlag.
- Lord, F. & Novick, M. R. (1968). **Statistical theories of mental test scores**. Reading, Mass.: Addison-Wesley.
- Luce, R.D. (1959). **Individual choice behavior**. New York: Wiley.
- Mergenthaler, E. (1985). **Textbank systems: Computer science applied in the field of psychoanalysis**. Berlin: Springer.
- Nagel, E. (1961). **The structure of science. Problems in the logic of scientific explanation**. New York: Harcourt, Brace & World.
- Nelson, R. J. (1977). Structure of complex systems. In F. Suppe & P.D. Asquith (Eds.), **PSA 1976, Vol. 2: Symposia**. East Lansing, Mich.: Philosophy of Science Association.

- Nelson, R. J. (1989). *The logic of mind* (2nd ed.). Dordrecht: Kluwer.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, Mass.: Harvard University Press.**
- Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, N.J.: Prentice Hall.
- Palmer, S. E. (1986). The information processing approach to cognition. In T. J. Knapp & L.C. Robertson (Eds.), *Approaches to cognition: Contrasts and Controversies* (S. 37-77). Hillsdale, N.J.: Erlbaum.
- Paul, W. J. (1978). *Komplexitätstheorie*. Stuttgart: Teubner.
- Peterson, W. W., Birdsall, T. G. & Fox, W.C. (1954). The theory of signal detectability. *Transactions of the Professional Group on Information Theory* [Institute of Radio Engineers], 4, 171-212.
- Popper, K.R. (1966). *Logik der Forschung* (zweite, erweiterte Auflage). Tübingen: Mohr.
- Posner, M. I. (Ed.) (1989). *Foundations of cognitive science*. Cambridge, Mass.: MIT Press.
- Roskam, E. E. (1983). Allgemeine Datentheorie. In H. Feger & J. Bredenkamp (Hrsg.), *Enzyklopädie der Psychologie, Themenbereich B: Methodologie und Methoden, Serie I: Forschungsmethoden der Psychologie, Band 3: Messen und Testen (S. 1-135)*. Göttingen: Hogrefe.
- Schoemaker, P. J. H. (1991). The quest for optimality: A positive heuristic of science? *Behavioral and Brain Sciences*, 14, 205-245.
- Selten, R. (1989). Evolution, *learning and economic behavior* (1989 Nancy L. Schwartz Memorial Lecture. Sonderforschungsbereich 303, discussion paper B-132). Bonn: Rheinische Friedrich-Wilhelms-Universität.
- Simon, H. (1977). How complex are complex systems? In F. Suppe & P.D. Asquith (Eds.), *PSA 1976, Vol. 2: Symposia*. East Lansing, Mich.: Philosophy of Science Association.
- Sneed, J. D. (1971). *The logical structure of mathematical physics*. Dordrecht: Reidel.
- Störig, H. J. (1970). *Kleine Weltgeschichte der Wissenschaft* (zwei Bände; Originalausgabe 1954, Stuttgart: Kohlhammer). Frankfurt: Fischer.
- Swijtink, Z.G. (1987). The objectification of observation: Measurement and statistical methods in the nineteenth century. In L. Krüger, L.J. Daston & M. Heidelberger (Eds.), *The probabilistic revolution. Vol. 1: Ideas in history* (Pp.261-285). Cambridge, Ma.: MIT Press.
- Tack, W. H. (1991). Rationales Handeln in sozialen Situationen. *Zeitschrift für Sozialpsychologie*, 22, 151-165.
- Tseëlon, E. (1991). The method is the message: On the meaning of methods as ideologies. *Theory and Psychology*, 1, 299-316.
- van Acker, P.O.F.C. (1977). *Models for intransitive choice*. Doctoral Dissertation, Katholieke Universiteit te Nijmegen, Nijmegen.
- van Meter, D. & Middleton, D. (1954). Modern statistical approaches to reception in communication theory. *Transactions of the Professional Group on Information Theory* [Institute of Radio Engineers], 4, 119-145.