

Lehr- und Forschungstexte Psychologie

- Band 1: I. Borg, Anwendungsorientierte Multidimensionale Skalierung. VI, 553 Seiten. 1981.
- Band 2: F. Rösler, Hirnelektrische Korrelate Kognitiver Prozesse. XI, 471 Seiten. 1982.
- Band 3: F. Rohrmeier, Langzeiterfolge Psychosomatischer Therapien. XII, 289 Seiten. 1982.
- Band 4: H. Rochel, Planung und Auswertung von Untersuchungen im Rahmen des allgemeinen linearen Modells. VI, 262 Seiten. 1983.
- Band 5: Fortschritte der Experimentalpsychologie. Herausgegeben von K. Pawlik. VII, 71 Seiten. 1984.
- Band 6: G. Strube, Assoziation. XII, 324 Seiten. 1984.
- Band 7: U. Schmidt-Denter, Die soziale Umwelt des Kindes. VII, 223 Seiten. 1984.
- Band 8: E.M. Steinmeyer, Depression und gelernte Hilflosigkeit. V, 198 Seiten. 1984.
- Band 9: H. Colonius, Stochastische Theorien individuellen Wahlverhaltens. XIV, 162 Seiten. 1984.
- Band 10: Psychologische Aspekte des Verstehens. Herausgegeben von J. Engelkamp. VIII, 254 Seiten. 1984.
- Band 11: J. Beckmann, Kognitive Dissonanz. VIII, 165 Seiten. 1984.
- Band 12: G. Haubensak, Absolutes und vergleichendes Urteil. XI, 198 Seiten. 1985.
- Band 13: W. W. Wittmann, Evaluationsforschung. XI, 547 Seiten. 1985.
- Band 14: G. Lehmann, Modell- und rekursionstheoretische Grundlagen psychologischer Theorienbildung. XXII, 297 Seiten. 1985.
- Band 15: Perspektiven der Kognitionspsychologie. Herausgegeben von O. Neumann. III, 276 Seiten. 1985.
- Band 16: G. Winneke, Blei in der Umwelt. IV, 192 Seiten. 1985.
- Band 17: K. Westhoff, Erwartungen und Entscheidungen. II, 197 Seiten. 1985.
- Band 18: J. Funke, Komplexes Problemlösen. VI, 145 Seiten. 1986.

Lehr- und Forschungstexte Psychologie 18

Herausgegeben von
D. Albert, K. Pawlik, K.-H. Stapf und W. Stroebe

Joachim Funke

Komplexes Problemlösen

Bestandsaufnahme und Perspektiven



Springer-Verlag

J. Funke: Komplexes Problemlösen

Das Buch gibt einen Überblick über aktuelle Arbeiten aus dem Bereich des „komplexen Problemlösens“. Darunter versteht man die Steuerungs- bzw. Regelungstätigkeiten von Individuen beim Umgang mit unbekanntem dynamischen, computersimulierten Systemen größeren Umfangs. Neben einer Darstellung von Forschungsproblemen, die am Beispiel von Simulationsprogrammen wie „Lohhausen“ oder „Schneiderwerkstatt“ erläutert werden, geht es auch um die Entwicklung einer Theorie über den Umgang mit komplexen Systemen.

Die Reihe *Lehr- und Forschungstexte Psychologie* zielt darauf ab, aktuelle Entwicklungen in allen Gebieten psychologischer Forschung auf hohem wissenschaftlichen Niveau Lehrenden, Forschenden und Studierenden zugänglich zu machen. Die einzelnen Bände sind thematisch in sich abgeschlossen und eigenständig. Dabei kann es sich handeln um: Berichte über eigene Forschungsergebnisse, Bände mit Arbeiten mehrerer Autoren, Monographien vom Typ 'Lehrbuch für Fortgeschrittene' oder 'Einführendes Lehrbuch für ein Spezialgebiet'. Darüber hinaus sind die Manuskripte so abgefaßt und illustriert, daß auch ein gut ausgebildeter Nichtspezialist dem Inhalt ohne Quellenstudium mit Verständnis folgen kann.

ISBN 3-540-16247-X

Lehr- und Forschungstexte Psychologie

Herausgegeben von
D. Albert, K. Pawlik, K.-H. Stapf und W. Stroebe

Joachim Funke

Komplexes Problemlösen

Bestandsaufnahme und Perspektiven



Springer-Verlag

Autor

Joachim Funke
Psychologisches Institut, Universität Bonn
Römerstr. 164, D-5300 Bonn 1

ISBN 3-540-16247-X Springer-Verlag Berlin Heidelberg New York Tokyo
ISBN 0-387-16247-X Springer-Verlag New York Heidelberg Berlin Tokyo

CIP-Kurztitelaufnahme der Deutschen Bibliothek. Funke, Joachim: Komplexes Problemlösen: Bestandsaufnahme u. Perspektiven / Joachim Funke. – Berlin; Heidelberg; New York; Tokyo: Springer, 1986.

(Lehr- und Forschungstexte Psychologie; 18)

ISBN 3-540-16247-X (Berlin ...)

ISBN 0-387-16247-X (New York ...)

NE: GT

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, der Entnahme von Abbildungen, der Funksendung, der Wiedergabe auf photomechanischem oder ähnlichem Wege und der Speicherung in Datenverarbeitungsanlagen bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten.

Die Vergütungsansprüche des § 54, Abs. 2 UrhG werden durch die 'Verwertungsgesellschaft Wort', München, wahrgenommen.

© Springer-Verlag Berlin Heidelberg 1986
Printed in Germany

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, daß solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten

Vorwort

Die vorliegende Arbeit besteht aus drei Teilen. Nach einer kurzen Einführung (Kapitel 1) erfolgt in einem ersten Teil eine Bestandsaufnahme gegenwärtiger Bemühungen im Bereich des komplexen Problemlösens (Kapitel 2 und 3). Unter komplexem Problemlösen werden dabei diejenigen Prozesse verstanden, die bei der Steuerung unbekannter, dynamischer, computersimulierter Systeme ablaufen. Neben einem kurzen historischen Rückblick wird versucht, die Vielfalt der Untersuchungsparadigmen aufzuzeigen und dabei Trennendes wie auch Gemeinsames deutlich werden zu lassen. Thesenartig werden Probleme dieser Arbeiten aufgezeigt und erste Ansätze zu deren Bewältigung genannt.

Der zweite Teil (Kapitel 4) vertieft verschiedene dieser Probleme an einem konkreten Beispiel, dem "TAILORSHOP". Neben der detaillierten Schilderung des computersimulierten Systems werden Fragen wie die nach der Optimierbarkeit des Gleichungssystems ebenso behandelt wie generelle Themen, etwa die Präzisierung des Transparenz-Konzepts oder das Problem möglicher Versuchsleitereffekte.

Im dritten Teil (Kapitel 5 und 6) geht es um das Aufzeigen eines konstruktiven Wegs aus dem diagnostizierten, unbefriedigenden Zustand gegenwärtiger Forschungsbemühungen. Eine formale Darstellung kognitionspsychologischer Konzepte mit Hilfe von Kausalmodellen soll zeigen, daß auch für die Erforschung des Problemlösens in komplexen Situationen ein Begriffsinventar zur Verfügung steht, das zum einen eine exaktere Beschreibung des Untersuchungsgegenstandes erlaubt, zum anderen eben dadurch die Möglichkeit gibt, präzisere Theorien über den Umgang des Menschen mit einfachen wie komplexen Systemen zu formulieren. Ein erster Ansatz zu einer derartigen Theorie wird vorgelegt.

Allen Kolleginnen und Kollegen, die mit ihren Ideen und Ratschlägen mein Vorhaben unterstützt haben, danke ich herzlich!

Bonn, im November 1985

Joachim Funke

INHALTSVERZEICHNIS

Seite

1	Einleitung: Ein Blick zurück	1
Teil I: Bestandsaufnahme		
2	Problemlösen in komplexen, computersimulierten Realitätsbereichen: Eine Übersicht	5
2.1	Der erste Impuls: "TANALAND"	5
2.2	Der Höhepunkt: "LOHHAUSEN"	7
2.3	Nebenschauplatz Testintelligenz: "TAILORSHOP"	9
2.4	Folgestudien	10
3	Thesen zur gegenwärtigen Lage der Forschung zum komplexen Problemlösen	20
3.1	Die These von der Theoriearmut	20
3.2	Die These von der mangelnden Integration anderer psychologischer Teildisziplinen	23
3.3	Die These von der fehlenden fachübergreifenden Kooperation	25
3.4	Die These von der unzureichenden Nutzung des systemtheoretischen Ansatzes	26
3.4.1	Komplexität	26
3.4.2	Vernetztheit	32
3.4.3	Stabilität und Katastrophe	37
3.5	Die These vom Fehlen einer Taxonomie von Problem- bzw. Systemtypen	41
3.6	Die These von der mangelhaften Beachtung des Meßfehlers	44
3.7	Die These von der suboptimalen Versuchsplanung und -auswertung	45
3.8	Die These von der unzureichenden Bestimmung der Lösungsgüte	47
3.9	Nachbemerkung zu den Thesen	51

Teil II: Das Beispiel "TAILORSHOP"		Seite
4	Das Simulationsprogramm "TAILORSHOP"	53
4.1	Beschreibung des Computerprogramms	53
4.2	Eine Simulationsstudie zum Simulationsprogramm	63
4.3	Bisherige Studien zum "TAILORSHOP"	69
4.4	Zur Kontroverse um die Interpretation der Ergebnisse	77
Teil III: Weiterführende Perspektiven für eine Theorie des Umgangs mit dynamischen Systemen		
5	Problemlösen als Konstruktion von Kausalmodellen	84
5.1	Ein Begriffsinventar für Kausalmodelle in der Denkpsychologie	85
5.2	Demonstration einiger typischer Formen von autoregressiven Prozessen	91
5.3	Konstruktion von Kausalmodellen: Eine Fallstudie ..	99
5.4	Anwendungsmöglichkeiten	103
5.5	Bausteine für eine Theorie der Konstruktion von Kausalmodellen über dynamische Prozesse	107
5.6	Abschließende Bemerkungen zum vorgestellten Rahmenkonzept	117
6	Abschluß: Ein Blick nach vorne	123
6.1	Rahmenmodell	123
6.2	Eine Heuristik zum Verständnis des Vp-Verhaltens .	124
6.3	Systematisierung der Forschungsaufgaben	125
Literatur		129
Register		141

1 Einleitung: Ein Blick zurück

Im deutschsprachigen Raum beschäftigt man sich seit gut zehn Jahren mit einem Thema, das unter dem Titel "komplexes Problemlösen" inzwischen Eingang in die Lehrbücher gefunden hat. Dieses Gebiet fasziniert nicht nur Fachwissenschaftler, sondern auch eine breitere Öffentlichkeit wegen der vermuteten Realitätsnähe des Forschungsgegenstandes. Im Unterschied zu klassischen Studien der Denkpsychologie, die sich beispielsweise mit Feinanalysen "künstlicher" Problemstellungen wie dem "Turm von Hanoi" (vgl. SYDOW 1980), dem "Kannibalen-und-Missionare-Problem" (vgl. JULISCH & KRAUSE 1976) oder "Schach" (vgl. SEIDEL 1977) beschäftigen, richtet sich die Erforschung komplexen Problemlösens auf die Bewältigung von Problemen in der alltäglichen Umwelt des Menschen. Insbesondere solche Probleme, die sich aus dem Umgang mit nicht näher bekannten Systemen ergeben, werden dabei fokussiert.

Das Aufgreifen dieser Thematik hat natürlich seinen zeitgeschichtlichen Hintergrund. In einer Epoche, in der das Überleben von Menschen auf unserem Planeten nicht nur durch wachsende Ressourcenknappheit, Hunger, Armut, ökologische Desaster und ökonomische Fehlregulationen gefährdet wird, sondern zusätzlich die angeblich friedenssichernden Pläne der Großmächte ein Bedrohungspotential allererster Ordnung schaffen, stellt sich tatsächlich die Frage nach den menschlichen Fähigkeiten des Umgangs mit derartigen Problemen. Ist es wirklich so, daß der angerichtete Schaden durch kurzfristige Eingriffe des Menschen in seine natürliche Umgebung nicht mehr zu korrigieren sein wird?

Sieht man einmal von diesen zeitgeschichtlichen Einflüssen ab, deren Kulmination durch so spektakuläre Berichte wie "Global 2000 - Bericht an den Präsidenten" oder die älteren Hochrechnungen des "Club of Rome" (MEADOWS, MEADOWS, ZAHN & MILLING 1972) genauso markiert werden wie durch den "Ölpreis-Schock" Mitte der 70er Jahre, gibt es aus der wissenschaftshistorischen Perspektive heraus eine mindestens ebenso tiefgreifende Wende: die Ablösung Newton'scher Denktradition durch die in den 30er Jahren entstandene systemtheoretische Sichtweise (vgl. CAPRA 1983²). Inwiefern es sich dabei tatsächlich um eine tiefgreifende Wende handelt oder ob nicht lediglich ein Austausch der Begrifflichkeit zur Beschreibung der immer gleichen Sachverhalte vorliegt, kann nicht ohne weiteres entschieden werden. Tatsache ist, daß durch systemtheoretische Betrachtungen Wissenschaften wie Medizin oder Biologie (vgl. PATTEN 1971, 1972) erheblich an Präzision gewonnen haben. BISCHOF (1981) diskutiert zwei komplementäre Denkweisen in den Naturwissenschaften: Während er das mit dem Namen "Galilei" verbundene Programm einer materiellen Reduktion des Untersuchungsgegenstandes für gescheitert erklärt, verspricht er sich eine mit dem Namen "Darwin" verbundene "teleonome Wende der Psychologie", bei der die Systemstruktur des untersuchten Objekts im Mittelpunkt steht. Inwiefern die hier wie dort heftig strapazierte Systemtheorie tatsächlich fruchtbar gemacht wurde, wird zu prüfen sein.

Problemlösen ist ein traditioneller Untersuchungsgegenstand deutscher Psychologen: Zu Beginn dieses Jahrhunderts haben sich Forscher wie ACH, DUNCKER, KULPE, SELZ und WERTHEIMER ausführlich mit Fragen des "produktiven Denkens" beschäftigt. Während in Frankreich und später in den USA die Psychometrie der höheren kognitiven Vorgänge in Form von Intelligenztests betrieben wurde, ging es deutschen Denkpsychologen weniger um die differentialdiagnostische Frage als um das Problem des Findens (griech.: heuriskein) von Problemlösungen. Typisch hierfür ist ein Zitat aus dem ersten Kapitel von DUNCKERS (1935, p.1) Monographie "Zur Psychologie des produktiven Denkens":

"Ein 'Problem' entsteht z.B. dann, wenn ein Lebewesen ein Ziel hat und nicht 'weiß', wie es dieses Ziel erreichen soll. Wo immer der gegebene Zustand sich nicht durch bloßes Handeln (Ausführen selbstverständlicher Operationen) in den erstrebten Zustand überführen läßt, wird das Denken auf den Plan gerufen. Ihm liegt es ob, ein vermittelndes Handeln allererst zu konzipieren. Die 'Lösung' eines solchen praktischen Problems hat somit zwei Forderungen zu genügen: ihre Verwirklichung (Umsetzung in die Praxis) muß erstens die Verwirklichung des erstrebten Zustandes zur Folge haben und zweitens vom gegebenen Zustand aus durch 'bloßes Handeln' erreichbar sein."

36 Jahre später finden wir in der Monographie "Information und Verhalten" von KLIX (1971) eine sehr ähnliche Bestimmung des Gegenstandes (p.639f): Neben einem gegebenen Ausgangszustand und einem angestrebten Zielzustand besteht das dritte Merkmal eines Problems in der nicht unmittelbar gelingenden Überführbarkeit des einen in den anderen Zustand. Auch andere aktuelle Themen der Allgemeinen Psychologie, etwa den Zusammenhang von Emotion und Kognition (vgl. KUHL 1983, MANDL & HUBER 1983) findet man in der Psychologie des Denkens und Wollens der Würzburger Schule bereits angelegt. Von SELZ (1913, p.1) stammt eine Bemerkung, die nach wie vor Gültigkeit besitzt:

"Eines der wichtigsten Probleme der Psychologie des Denkens ist daher die Bestimmung der richtungsgebenden Faktoren, die den geordneten Ablauf des Denkens herbeiführen, und die Auffindung der Gesetze ihrer Wirksamkeit."

Das Versuchsmaterial, das in den Studien der frühen Denkpsychologen Verwendung fand, muß nach heutigen Maßstäben als wenig komplex bezeichnet werden: Bei SELZ etwa ging es um Teil-Ganzes-Relationen bzw. Relationen der Über-, Neben- und Unterordnung (SELZ 1913, p.12; vgl. auch SELZ 1922, Zweiter Abschnitt, §§ 1-5), bei DUNCKER um die bekannte Bestrahlungsaufgabe:

"...gesucht ein Verfahren, um einen Menschen von einer inoperablen Magengeschwulst zu befreien mit Hilfe von Strahlen, die bei genügender Intensität organisches Gewebe zerstören - unter Vermeidung einer Mitzerstörung der umliegenden gesunden Körperpartien." (DUNCKER 1935, p.1).

Die ersten Verwendungen des Komplexitätsbegriffs - wenn wir einmal von Bezeichnungen wie "Komplexergänzung" (SELZ) oder ähnlichen Termini absehen - findet man in amerikanischen Arbeiten der 50er Jahre. RAY (1955) stellt etwa unter dem Titel "Complex tasks for use in human problem solving research" insgesamt 29 Aufgaben vor, die seiner Ansicht nach als komplex bezeichnet werden müssen. Kriterien dafür sind nach RAY (1) die Notwendigkeit mehrerer Antworten, (2) die Verfügbarkeit von mehr als einer Hypothese, (3) ein 'trial-and-error' mit mehr als ein oder zwei Schritten und (4) der gegenüber einem einfachen Akt erhöhte Zeitbedarf. Interessanterweise handelt es sich bei den genannten Problemen um Aufgaben gestaltpsychologischer Provenienz (also von DUNCKER, MAYER, SZEKELY, WERTHEIMER und anderen). Auch das Wasserumschütt-Problem, das LUCHINS zur Demonstration der Set-Bildung benutzt, gehört dazu. Einfache Aufgabentypen sind für RAY das Lernen von sinnfreien Silben oder das an Tieren untersuchte Labyrinthlernen.

Zusammen mit der im Nachkriegsaufschwung rapide fortschreitenden Entwicklung technischer Systeme (vor allem im militärischen Bereich) finden sich in den 50er Jahren ebenfalls erste Arbeiten von Psychologen zu Problemen, die im Umgang mit "der Technik" entstehen: Wie kann etwa schnellstmöglich der Fehler in einem defekten Radio erkannt werden ("trouble shooting"; vgl. GLASER, DAMRIN & GARDNER 1954 mit ihrer Methode des "tab item"). Die Beantwortung derartiger Fragestellungen gewinnt natürlich angesichts der enormen Veränderungen in immer weiteren Bereichen unserer Lebenswelt (Kernkraft; Waffentechnik; Raumfahrt; Elektronik in allen Varianten) an Bedeutung. In der Edition von RASMUSSEN & ROUSE (1981) über die Entdeckung von "system failures" wird dies ebenso anschaulich demonstriert wie in der Edition von MORAAL & KRAISS (1981) über "manned systems".

Deutsche Problemlöseforscher haben erst spät dieses Thema wieder aufgegriffen, vielleicht wegen einer eher grundwissenschaftlichen Orientierung, die nach den "Bausteinen" des Denkens suchte. Der "General Problem Solver" von NEWELL, SHAW & SIMON (1960) nahm mathematische Beweisführungen vor, ein Prozeß, den LÜER (1973) bei Probanden (im folgenden auch bezeichnet als Pbn oder Vpn) untersuchte, die logische Theoreme beweisen sollten (ähnlich DÖRNER 1974). In einer Übersicht über die Entwicklung der deutschen Problemlöseforschung nach dem zweiten Weltkrieg kommen LÜER & PUTZ-OSTERLOH (1978) zu dem Schluß, daß (1) eine Entwicklung der Denk- hin zur Problemlöseforschung stattfindet, (2) Problemlösen in Termini der Informationsverarbeitung beschrieben wird und (3) experimentelle Forschung zum Problemlöseprozeß durch den Einsatz komplexer statistischer Verfahren der Datenanalyse möglich wird. "Marksteine" dieser Entwicklung sind Arbeiten wie der Kongreßbeitrag von SÜLLWOLD (1960), der Handbuch-Beitrag von BERGIUS (1964) oder die Anthologie von GRAUMANN (1965). Deutschsprachige Lehrbücher gibt es wenige: OERTER (1971) und DÖRNER (1976) sowie AEBLI (1980, 1981) und jüngst HUSSY (1984a, 1986) stecken das Feld ab.

"Komplexes Problemlösen" im heutigen Verständnis wird im deutschsprachigen Raum meines Wissens erstmals von DÖRNER, DREWES & REITHER (1975) angesprochen und gleich mit einem spezifischen Inhalt gefüllt:

Es geht um den Umgang von Menschen mit unbekanntem Szenarien, in denen sie Entscheidungsbefugnis übertragen bekommen und für eine bestimmte Zahl von Simulationstakten eine Steuerungsfunktion übernehmen sollen. Die Etablierung eines Projekts "Systemdenken" durch die DFG garantierte einen finanziellen und personellen Rahmen für die Arbeit an diesem Thema. Zentrale Absicht ist die Ablösung der traditionellen Paradigmen der Problemlöseforschung durch computersimulierte Systeme unterschiedlichster Art. Doch noch Ende der 70er Jahre liegen kaum mehr als kurze Veröffentlichungen in Kongreßberichten vor (etwa DÖRNER 1979; DÖRNER & STAUDEL 1979; KREUZIG 1979a, b; REITHER 1979). Dies hat sich in den 80er Jahren verändert: Man kann durchaus von einer "Welle" sprechen, die von der Arbeit über das Kleinsystem "TANALAND" (DÖRNER & REITHER 1978) und dem darauf folgenden Großsystem "LOHHAUSEN" (DÖRNER, KREUZIG, REITHER & STAUDEL 1983) ausging. Der heutige Stand verlangt bereits nach einem orientierenden Überblick, den der Leser im folgenden Kapitel findet. Die Übersichtsartikel von LÜER & PUTZ-OSTERLOH (1978) sowie von KRAUSE (1982a, b) gehen auf diese Arbeiten noch nicht ein.

Teil I: Bestandsaufnahme

2 Problemlösen von komplexen, computersimulierten Realitätsbereichen: Eine Übersicht

Die Studien, über die nachfolgend berichtet wird, erfüllen drei Kriterien:

- (1) Es liegt ein Simulationssystem über einen bestimmten Realitätsbereich vor mit mindestens zwei voneinander abhängigen Variablen.
- (2) Probanden werden mit diesem Simulationssystem konfrontiert.
- (3) Das Verhalten der Probanden beim Umgang mit diesem Simulationssystem wird unter psychologischer Perspektive betrachtet.

Entgegen der ursprünglichen Absicht einer Darstellung in alphabetischer Folge wird der Versuch gemacht, die Entwicklungslinien dieses Forschungszweiges hervorzuheben. Am Anfang steht somit das System "TANALAND" und dessen Steigerung in Form von "LOHHAUSEN". Es folgt "TAILORSHOP" und die damit verbundene Frage nach dem Zusammenhang von Testintelligenz und Problemlösen. Schließlich sind die zahlreichen Neu- und Weiterentwicklungen zu betrachten, die ein Aufbrechen der Forschungs"tradition" in mehrere, nach Intention und Ansatz verschiedene Stoßrichtungen erkennen lassen.

2.1 Der erste Impuls: "TANALAND"

Das System "TANALAND" (vgl. DÖRNER 1975; DÖRNER, DREWES & REITHER 1975; DÖRNER & REITHER 1978) ist das früheste publizierte deutsche Simulationsprogramm. Simuliert wird ein ökologisches System einer fiktiven afrikanischen Landschaft mit verschiedenen Tier- und Pflanzenarten sowie zwei Menschengruppen, den ackerbauenden "Tupis" und den "Moros", die von Rinder- und Schafzucht leben. Die gut 50 Systemvariablen waren durch eine "Vielzahl positiver und negativer Rückkopplungen" verknüpft. Die zwölf Pbn dieser Erkundungsstudie - Studenten mit einem mittleren IQ von 122 - sollten sich in die Rolle eines landwirtschaftlich-technischen Beraters versetzen und die Lebensbedingungen der dort ansässigen Bevölkerung verbessern. Jedoch:

"Es zeigt sich in den Ergebnissen, daß die Probanden fast ausnahmslos das ursprünglich stabile Gefüge der Variablen des simulierten Landes zerstörten und dadurch häufig katastrophale Zustände schufen. Trotz großen persönlichen Engagements und einer ständig wachsenden Zahl von Maßnahmen und Entscheidungen, ist doch festzustellen, daß viele Probanden sich mit der Zeit auf wenige Aufgabenschwerpunkte konzentrieren und nicht mehr bereit sind, diese zu wechseln. Diese Tatsache, gepaart mit der offensichtlichen Unfähigkeit der Probanden, exponentielle Entwick-

lungsverläufe und die 'Vernetztheit' der Variablen angemessen zu berücksichtigen, stellen die Hauptdefizite des Verhaltens der Probanden in dem von uns konstruierten komplexen Realitätsbereich dar." (DÖRNER & REITHER 1978, p.527f).

Die beobachteten Mißerfolge werden nicht auf mangelndes Wissen zurückgeführt, vielmehr offenbaren sich darin Defizite allgemeiner Natur. So sei es klar, daß "die Vpn über die Denkmittel für den Umgang mit sehr komplexen Systemen nicht in hinreichendem Maße verfügten" - so DÖRNER, DREWES & REITHER (1975, p.340). Die Brisanz dieser Behauptung schlug bis in die Tagespresse durch, wo Titel wie "Überfordern Umwelt- und Wirtschaftspolitik das Denkvermögen von Wählern und Politikern?" ("SPIEGEL" Heft 50, 1981, p.66) das Scheitern des "linearen Denkens" verkündeten. In den realisierten Systemen, die durch die Merkmale der Dynamik, Komplexität, Vernetztheit und Intransparenz beschrieben werden, sei vielmehr ein Denken in Wirkungsnetzen angebracht.

Bereits 1978 haben DÖRNER & REITHER Kritik an ihrem "experimentellen Vorgehen" antizipiert. Folgende mögliche Einwände sahen sie (vgl. DÖRNER & REITHER 1978, p.534-535):

- (1) "Man denkt sich ein System aus, welches den Vpn z.T. unbekannt ist, überläßt es den Vpn zur Manipulation, um dann trivialerweise festzustellen, daß die Vpn mit ihrer Aufgabe überfordert sind." Gegenargument: Eine derartige Situation ist typisch z.B. für den Umgang mit politischen, wirtschaftlichen oder ökologischen Problemen.
- (2) "...gewöhnlich (verfügen) Entscheidungsbefugte in solchen Bereichen,..., über ein größeres einschlägiges Fachwissen ...und (sind) aus diesem Grunde für diese Aufgabe besser gerüstet...". Gegenargument: Oft genug werden Entscheidungen ohne Sachwissen getroffen bzw. wird das Sachwissen erst durch die Tätigkeit im jeweiligen Feld erworben.
- (3) Entscheidungsprozesse finden gewöhnlich in Gruppen statt. Gegenargument: Das Argument wird akzeptiert, Entscheidungsprozesse des Einzelnen interessieren jedoch insofern, um seine Verhaltenstendenzen zu erfahren, die er in die Gruppenkommunikation einbringt.
- (4) Planspiele sind ergebnis- und prozeßorientiert. Gegenargument: Hauptsächlich interessiert der Prozeß der Entscheidungsfindung.

Vor allem die beiden ersten antizipierten Einwände sind auch heute noch diskutabel. Sie laufen auf einen Haupteinwand hinaus, der bereits 1974 auf dem damaligen DGfPs-Kongreß in der Diskussion zu "TANALAND" geäußert wurde:

"PUBLIKUM: War der Probanden Verarbeitungskapazität simultaner Daten nicht überschritten?"

DÖRNER: Gewiß. Es gilt bessere Strategien der Abstraktion und der zweckmäßigen Datenreduktion den Probanden beizubringen." (DÖRNER, DREWES & REITHER 1975, p.341).

Die Gegenargumentation auf die beiden erstgenannten Einwände stützt sich ausschließlich auf die angenommene Realitätsnähe der Simulationssituation, eine Argumentation, die - wie wir sehen werden - angesichts der Validierungsprozedur Zweifel aufwirft.

Die Pilotstudie "TANALAND" blieb Vorläufer des eigentlichen "Paradepferdes", auf das nun gesetzt wurde: "LOHHAUSEN".

2.2 Der Höhepunkt: "LOHHAUSEN"

Hatte man mit "TANALAND" einen ersten Schritt in Richtung auf komplexes Untersuchungsmaterial hin gemacht, geht es nun weiter: "Simuliert wird das Leben" - so Dieter E. Zimmer in seinem Bericht über "das große Bürgermeisterspiel" ("DIE ZEIT" Nr. 37/38, 1981). Gut 2000 Variablen liegen dem simulierten Realitätsbereich - einer Kommune namens "LOHHAUSEN" - zugrunde. "LOHHAUSEN" ist das wohl prominenteste der hier erwähnten Simulationssysteme. Die umfangreiche Monographie von DÖRNER, KREUZIG, REITHER & STÄUDEL (1983) stellt den Abschlußbericht über die gut fünfjährige Projektarbeit an dieser einmaligen Studie dar. In drei Sätzen beschreiben die Autoren den Inhalt des 446 Seiten starken Buches:

"Der nachfolgende Bericht schildert die Ergebnisse eines relativ langwierigen psychologischen Experimentes. Wir haben uns bemüht, etwas über die Bedingungen und Formen des Handelns in Unbestimmtheit und Komplexität zu erfahren. Zu diesem Zweck haben wir 48 Versuchspersonen bei ihrem Handeln über eine ziemlich lange Zeitstrecke systematisch beobachtet und die vielfältigen Ergebnisse dieser Beobachtungen ausgewertet." (p.13)

Der Name "LOHHAUSEN" steht seitdem für ein Programm: ein Vertiefen kognitionspsychologischer Erkenntnisse durch Aufbruch zu neuen Paradigmen der Problemlöseforschung. Paradigmen, die im Unterschied zu den herkömmlichen Problemtypen wie etwa Denksportaufgaben ganz wesentlich das Merkmal der Komplexität und Unbestimmtheit aufweisen (vgl. DÖRNER 1981). Als ein Beispiel dafür betrachten die Autoren das von ihnen erstellte Simulationsprogramm der kleinen, fiktiven Gemeinde namens "LOHHAUSEN". Über einen simulierten Zeitraum von zehn Jahren - aufgeteilt auf acht zweistündige Sitzungen - sollten die Probanden in der Rolle eines Bürgermeisters von Lohhausen für das "Wohlergehen der Stadt in der näheren und ferneren Zukunft sorgen", eine Aufgabe, die angesichts der Programmkomplexität nicht gerade leicht fiel. Doch auch die Untersucher hatten Komplexes zu bewältigen: Pro Versuchsperson fielen ca. 100000 Daten an, aus denen man Wichtiges von Unwichtigem,

Zufälliges von Bedeutsamem herauszufiltern hoffte.

Die Analyse der Befunde erfolgte - mit wenigen Ausnahmen wie z.B. den Einzeldarstellungen von ausgewählten Probanden - gruppenstatistisch. Zunächst werden objektive und subjektive Maße der Problemlöse-güte zu einem einzigen "Generalgütekriterium" agglomeriert, anhand dessen die Leistungen der Vp ganz generell ablesbar sind und das eine Aufteilung der Stichprobe in zwei Extremgruppen besonders guter bzw. besonders schlechter Qualität (jeweils N=12) ermöglicht. Es stellt sich heraus, daß sich bei den schlechten Vpn verschiedenste Variablen im System "LOHHAUSEN" (so z.B. das Fabrik-, Stadt- und Bankkapital, Produktion und Verkauf, Einwohner- und Arbeitslosenzahl) ungünstiger entwickeln als bei den guten Vpn. Doch auch "die Guten" sind nicht das, was ihr Name verspricht: Systemkenner erreichen in einigen Variablen noch höhere (=bessere ?) Werte.

Interessanter als diese Verhaltenseffekte sind die damit verbundenen Denk-, Planungs- und Entscheidungsprozesse: Neben formalen Kennzeichen (z.B. Häufigkeit und Konsistenz von Entscheidungen) und inhaltlichen Schwerpunktbildungen (z.B. "Finanzsituation der Uhrenfabrik") des Grobprotokolls der Versuchsleiter gibt es interessante Hinweise in den "Denkprotokollen", also den Daten lauten Denkens, zu dem die Vpn angehalten wurden. Der Zusammenhang zwischen Problemlöse-güte und Testintelligenz wird überprüft, der von Konstruktseite her gesehen hoch ausfallen sollte: "(Wir erwarten) eine hohe Korrelation zwischen Generalgütekriterium und Intelligenzleistung" (p.304; kursiv). Die Daten entsprechen jedoch nicht dieser Hypothese: Weder RAVENS "Advanced Progressive Matrices" (APM) noch CATTELLs "Culture Fair Intelligence Test" (CFT) korrelieren substantiell mit der Lösungsgüte, wohl dagegen das spontane VI-Urteil "Vp macht intelligenten Eindruck"! Die Autoren diskutieren zu Recht einige Mängel klassischer IQ-Tests (z.B. fehlende Berücksichtigung der Informationssuche), sehen jedoch eine mögliche Interpretation ihres Befundes nicht: die möglicherweise mangelnde Güte des Problemlösemaßes. Von Intelligenztests weiß man, daß bei wiederholter Messung ein ähnliches Resultat erzielt wird; das Problem des Meßfehlers beim Problemlösemaß wird dagegen nicht angesprochen. Auch das Stichprobenproblem (Studenten mit durchschnittlich höherem IQ) sollte nicht übersehen werden. - Weitere Befunde betreffen Persönlichkeitsmerkmale in ihrer Beziehung zur Lösungsgüte. Hier ist vor allem das Konstrukt "Selbstsicherheit" zu nennen, das in positiver Beziehung zum komplexen Problemlösen steht und gegen das "völlige Versagen der Intelligenztests" ins Feld geführt wird (vgl. KREUZIG 1979a, b, 1981).

Das theoretische Kondensat der umfangreichen Arbeit besteht im wesentlichen in der Aufzählung elementarer Prozesse der Informationsverarbeitung beim Lösen komplexer Probleme wie etwa der Komponenten- und Dependenzanalyse sowie Sub- und Superordinationsprozessen. Die Darstellung der Teilzielbildung und -verfolgung mündet in das bekannte Modell der Absichtsbehandlung (vgl. DÖRNER 1982, 1983c). Ausgehend von einer emotionalen Einbettung kognitiver Prozesse (DÖRNER 1983a; DÖR-

NER, REITHER & STAUDEL 1983) kann die "intellektuelle Notfallreaktion" - eine schnelle und allgemeine Reaktion des kognitiven Systems auf unspezifische Gefahrensituationen - mit der aktuellen Kompetenz des Handelnden in Verbindung gebracht werden, wobei Selbstsicherheit als Indikator für heuristische Kompetenz steht, der Fähigkeit, "auch für unbekannte Situationen adäquate Handlungsmöglichkeiten erzeugen zu können" (p.436). Zentralen Stellenwert in der "Theorie" besitzt der Begriff der "Kontrolle": Kontrollkompetenz garantiert Handeln in Unbestimmtheit, Verlust der Kontrolle führt zu den negativen emotionalen Konsequenzen, von denen problemlösendes Denken überlagert ist.

"LOHHAUSEN" steht nicht nur für ein neues Untersuchungsfeld der Denkpsychologie, es ist zugleich ein Plädoyer gegen die im Wissenschaftsbetrieb vorherrschende "analytische Prozedur" (vgl. DÖRNER 1983b). Die Untersuchung des hochgradig vernetzten Systems "Mensch" kann nach DÖRNER nicht streng experimentell erfolgen, da die Isolation einiger ausgewählter Variablen im Labor wenig über den "normalen" Prozeßablauf sagen könne, der in Wechselwirkungen mit anderen Variablen eingebettet ist. Die Forderung nach einer Vertiefung des "Käfer- und Schmetterlingssammelns", also der genauen Beschreibung der untersuchten Sachverhalte, geht einher mit der Suche nach "übergreifenden Rahmenvorstellungen über den Ablauf des psychischen Gesamtgeschehens" (DÖRNER 1983b, p.24). Nicht zuletzt durch diese wissenschaftstheoretischen Randbemerkungen trägt "LOHHAUSEN" das Flair der Wende: Bruch mit der klassischen Vorgehensweise streng experimenteller Laborforschung, Vergessen der vielen "l'art pour l'art"-Untersuchungen, Aufgreifen der "alltäglichen" Probleme.

Daß diese Wende nicht von allen Mitgliedern der "scientific community" nachvollzogen wird, dürfte klar sein. Das übernächste Kapitel gibt einige Beispiele dafür. Doch zunächst ein paar Bemerkungen über einen Nebenschauplatz, bei dem es hauptsächlich um das "Versagen" klassischer Intelligenztests geht.

2.3 Nebenschauplatz Testintelligenz: "TAILORSHOP"

Der "TAILORSHOP" (vgl. PUTZ-OSTERLOH 1981, 1983b) ist ein Kleinsystem, in dem Pbn die Leitung einer Schneiderwerkstatt übernehmen müssen: Durch Einkauf von Rohmaterial und Bereitstellung von Produktionskapazität in Form von Arbeitern und Maschinen werden Hemden produziert und mit Gewinn verkauft. Das Ziel der Pilotstudie (PUTZ-OSTERLOH 1981) faßt die Autorin wie folgt zusammen:

"Es gilt..., die Anforderungen von komplexen Problemen und von Intelligenztestaufgaben zu beschreiben und zu untersuchen, welche Problemlöseprozesse zu ihrer Überwindung eingesetzt werden. Dabei sollen komplexe Probleme alltäglichen Problemsituationen stärker angenähert sein als es Intelligenztestaufgaben sind." (PUTZ-OSTERLOH 1981, p.80).

Im Unterschied zu IQ-Tests wird bei einem komplexen Problem das Aufstellen und Ableiten von Problemlösezielen, die Auswahl von Handlungen zum Erreichen der Ziele und die aktive Suche nach Informationen über relevante Systemvariablen gefordert. - Auf die Befunde dieser Pilotstudie soll hier nicht eingegangen werden. Die im gleichen Jahr erschienene und in gleicher Weise wie die Pilotstudie angelegte Folgestudie von PUTZ-OSTERLOH & LUER (1981) prüft die Hypothese fehlender Beziehungen zwischen Testintelligenz und Problemlösequalität durch Vergleich einer Transparenz- mit einer Intransparenzbedingung (N = 70 studentische Vpn), die durch das Vorliegen bzw. Nichtvorliegen einer Vernetzungsgrafik der Systemvariablen gekennzeichnet ist. Die Autoren finden unter der so operationalisierten Transparenzbedingung einen statistisch bedeutsamen Zusammenhang zwischen den beiden Variablen, unter Intransparenz dagegen nicht, was sie im Sinne einer Kritik an herkömmlichen Intelligenztests interpretieren (zur Frage der Transparenz vgl. Kap. 4.4). "Alltagserfahrungen machen plausibel, daß 'reale' Probleme eher intransparent und hochkomplex sind und somit ganz entscheidend jenes Verhalten erfordern, das durch Intelligenztestverfahren nicht abrufbar ist." (PUTZ-OSTERLOH & LUER 1981, p.332). Diese Argumentation -sie findet sich in vielen Studien dieser Arbeitsgruppe (vgl. KREUZIG 1979a, b, 1981; DÖRNER & KREUZIG 1983)- ist nicht unwidersprochen geblieben (vgl. FUNKE 1983b; HUSSY 1985; JÄGER 1984; TENT 1984). In Kap. 4.3 finden sich detailliertere Überlegungen hierzu, so daß sich an dieser Stelle eine ausführliche Darlegung der Kontroverse erübrigt.

In einer neueren Arbeit unterscheidet PUTZ-OSTERLOH (1985b) zwischen "Problemlöseintelligenz" und "Testintelligenz", wobei sie mit Überschneidungen und nicht Deckungsgleichheit der jeweiligen Leistungs determinanten rechnet. Damit dürfte dieser Nebenschauplatz an Bedeutung verlieren, ein Prozeß, der bei frühzeitiger Elaboration der theoretischen Konzepte schon eher hätte einsetzen können. Das Simulationssystem "TAILORSHOP" ist damit natürlich nicht vom Tisch. Kritische Anmerkungen zu Details dieses Programms wie auch zu den damit durchgeführten Untersuchungen finden sich in Teil II dieser Arbeit.

2.4 Folgestudien

Eine Systematisierung der Folgestudien fällt nicht leicht: Die inhaltlich untersuchten Fragestellungen divergieren vielfach, die gewählten Realitätsbereiche sind äußerst heterogen und die Simulationssysteme selbst lassen sich ebenfalls nur vordergründig vergleichen. Tab. 2.1 zeigt in einer Übersicht die im deutschsprachigen Raum verwendeten Simulationsprogramme.

Als Gruppierungsmerkmal wurde in Tab. 2.1 die Anzahl der beteiligten Variablen zugrundegelegt, ein Kriterium, das gelegentlich als wesentlicher Indikator von Komplexität angesehen wird (vgl. die Kritik hierzu in Kap. 3.4). Wie wenig aufschlußreich dieses Kriterium ist, wird deutlich bei der Frage der subjektiven Wirkung dieses Parameters: Ist etwa ein System mit 12 Variablen nur halb so schwer wie ein System

Tab. 2.1: Übersicht über die im deutschsprachigen Raum verwendeten Simulationsprogramme.

System	Variablenzahl	Publikationen (alphabetisch geordnet)
(a) Systeme mit bis zu 10 Variablen		
HAMURABI	8	GEDIGA 1983; GEDIGA, SCHÖTTKE & TUCKE 1982, 1983; GEDIGA, SCHÖTTKE & TUCKE-BRESSLER 1984; SCHÖTTKE & GEDIGA 1982
MINISEE	6	OPWIS & SPADA 1983, 1985; OPWIS, SPADA & SCHWIERSCH 1985; SPADA, MAY & OPWIS 1983
MONDLANDUNG	3	FUNKE 1981, 1983a; FUNKE & HUSSY 1984; HUSSY 1985; THALMAIER 1979
ÖKOSYSTEM	6	FUNKE 1985
PORAEU	8	PREUSSLER 1985; PREUSSLER & DÖRNER 1985
SIM002	10	KLUWE & REIMANN 1983; REIMANN & KLUWE 1983
WELT	4	EYFERTH et al. 1982
(b) Systeme mit bis zu 100 Variablen		
DAGU/DORI	12	HESSE 1982; HESSE & GERRARDS 1985; PUTZ-OSTERLOH 1985a, b; REITHER 1981; ROTH 1985
EPIDEMIE	13	HESSE, SPIES & LUER 1983; PUTZ-OSTERLOH 1985b; SPIES & HESSE 1983
MORO	49	PUTZ-OSTERLOH 1985a,b; STROHSCHNEIDER 1985
SIM003	15	KLUWE, MISIAK & REIMANN 1984; KLUWE, MISIAK & SCHMIDLE 1985
TAILORSHOP	24	FUNKE 1983b; LUER & HUBNER 1985; LUER, HUBNER & LASS 1984; PUTZ-OSTERLOH 1981, 1983b; PUTZ-OSTERLOH & LUER 1981; ROTH 1985
TANALAND	54	DÖRNER 1975; DÖRNER & REITHER 1978; SIMON & WEDEKIND 1980
(c) Systeme mit mehr als 100 Variablen		
ENERGIE	>2000	ROST & VENT 1985
LOHHAUSEN	>2000	DÖRNER et al. 1983 (stellvertretend für eine Vielzahl von Publikationen)

mit 24 Variablen? Ist ein System mit mehr als 2000 Variablen hundertmal schwerer als das letztgenannte? Besteht eine nichtlineare (logarithmische ?) Beziehung zwischen empfundener Schwierigkeit und Variablenzahl bzw. anderen, aussagekräftigen Systemmerkmalen? Auf diese Fragen wird Kap. 3 eingegangen; hier soll zunächst eine knappe Skizze der jeweiligen Arbeiten vorgelegt werden.

(a) Studien, in denen Personenmerkmale eine zentrale Rolle spielen

Unter diesem Titel fassen wir Arbeiten zusammen, bei denen Eigenschaften des Simulationsprogramms als gegeben akzeptiert werden und die Fragestellung dahingehend orientiert ist, welche personseitigen Merkmale für beobachtete Systemeffekte (in der Regel "erfolgreiches" vs. "erfolgloses" Manipulieren) verantwortlich zu machen sind. Neben dem Einfluß von Testintelligenz interessieren hier Merkmale wie Selbstreflexion, Motivation, Differenziertheit des semantischen Gedächtnisses und Selbstsicherheit. Sprachstatistische Analysen (ROTH 1985), Untersuchungen der Blickbewegungen (LUER, HÜBNER & LASS 1984) oder zum Einfluß politischer Werthaltungen (KÜHLE & BADKE 1985) markieren weitere Fragestellungen in diesem Bereich.

Mit der Simulation der klimatischen, ökologischen und ethnischen Bedingungen eines fiktiven Entwicklungshilfegebietes namens "DAGU" beschäftigt sich REITHER (1981). Zwölf Entwicklungshelfer mit 6- bis 8-jähriger Praxis in afrikanischen bzw. asiatischen Entwicklungsländern (=Experten) werden hinsichtlich ihres Umgangs mit der Problemstellung verglichen mit zwölf Personen vor Beginn ihres Einsatzes als Entwicklungshelfer (= Novizen). Aufgabe der Pbn war es, den Menschen in "DAGU" zu besseren Lebensbedingungen zu verhelfen und die Zahl der Einwohner zu erhöhen, gleichzeitig jedoch extreme Überbevölkerung zu vermeiden. Sieben Eingriffsbereiche (Nahrung, Tierfutter, Geburtenkontrolle, medizinische Versorgung, Aktionen gegen die Tse-Tse-Fliegen, Anlage von Bewässerungsprojekten, Verkauf von erzeugten Produkten) standen hierfür zur Verfügung. Die Ergebnisse der Studie, bei der die Pbn in Dreiergruppen zusammenarbeiteten, belegen Unterschiede in Denken und Handeln bei Experten und Novizen: Novizen denken eher in Kausalketten als in Kausalnetzen, zeigen mehr thematische Sprünge und äußern mehr "Metastatements". Kennzeichen von Experten ist deren blindes Handeln, d.h. unter allen Bedingungen werden Maßnahmen getroffen, um Handlungskontinuität zu demonstrieren. Die Befunde zeigen auch, daß selbst Experten nicht in der Lage sind, etwa die kritische Variable "Bevölkerungszahl" zu stabilisieren.

Das Programm "DORI" simuliert die Lebensbedingungen eines Noma-stammes in der Sahelzone, der sich von der Rinderzucht ernährt. In der Arbeit von HESSE (1982) wird die Bearbeitung dieser Fassung als semantische Version verglichen mit den Ergebnissen derjenigen Version, bei der die Variablen durch lateinische Buchstaben gekennzeichnet werden (=nichtsemantische, abstrakte Version). Der Faktor Semantik ist gekreuzt mit einem Kontrollfaktor Transparenz, dessen zwei Stufen das Vorliegen bzw. Fehlen einer Variablengrafik bezeichnen. Insgesamt 120 Studenten, je 30 pro Versuchsbedingung, dienten als Stichprobe. Die gefundenen Unterschiede zwischen guten und schlechten Problemlösern deuten auf eine durch die semantische Einkleidung bewirkte Differenz in der Bearbeitungsstrategie hin; während in der abstrakten Bedingung stärker von Notizen Gebrauch gemacht wird ("externes Gedächtnis"; vgl. MUTHIG & SCHÖNPFUG 1981), stellen Pbn der Semantik-Gruppe gezieltere Fragen und organisieren ihre Vorgehensweise besser. In der abstrakten Bedingung tritt ein positiver Zusammenhang zwischen Testintelligenz

und Problemlösegüte auf.

Beim "EPIDEMIE"-Programm, das "mit Ausnahme neu gewählter Semantik und einzelner Verknüpfungen" (sic!) dem "DORI"-System entspricht, muß die Vp in der Studie von HESSE, SPIES & LUER (1983; SPIES & HESSE 1983) als Leiter der Gesundheitsbehörde einer kleinen Stadt die Folgen des Ausbruchs einer Epidemie regulieren, d.h. die von ihm getroffenen Maßnahmen sollen den Krankenstand weitmöglichst reduzieren. Dabei stehen ihm sieben verschiedene Maßnahmevariablen zur Verfügung. Fragestellung der "EPIDEMIE"-Studie war die Wirkung unterschiedlicher persönlicher Betroffenheit, realisiert durch zwei Epidemie-Arten, die von je N=30 Studenten bearbeitet wurden: Während niedrige Betroffenheit durch eine simulierte Grippe-Epidemie hergestellt wird, soll durch die Variante einer gefährlichen Pocken-Epidemie höhere Betroffenheit erzeugt werden. Die experimentellen Befunde sprechen für eine Wirksamkeit dieser Variable auf die Problemlösegüte: Die stark betroffenen Pbn erzielten bessere Güterwerte, arbeiten ausdauernder, nehmen mehr sinnvolle Handlungen vor und erkennen eher effektive Maßnahmen.

"HAMURABI" heißt der absolutistische Herrscher des Agrarstaates "Summaria", der in den Studien von GEDIGA (1983) sowie von GEDIGA, SCHÖTTKE & TÜCKE (1982, 1983; GEDIGA, SCHÖTTKE & TÜCKE-BRESSLER 1984; SCHÖTTKE & GEDIGA 1982) simuliert wird. Die Pbn (N=28 Psychologie-Studenten) erhalten die Aufgabe, durch vier sequentielle Entscheidungen (Ackerland kaufen, Ackerland verkaufen, Größe der Kornanbaufläche bestimmen, Anzahl der Nahrungsmittel pro Einwohner festlegen) in zwei Durchgängen zu je 30 Spieljahren möglichst wenige Bürger von Summaria sterben zu lassen; geringe Zufallsschwankungen konnten dabei auftreten. Die Ergebnisse der Arbeiten belegen zum einen, daß Problemsituationen mit exponentiellen zeitlichen Veränderungen von einem Teil der Vpn beherrscht werden, zum anderen zeigt sich, daß die Hypothesenbildung vieler Vpn der komplexen Problemsituation angemessen war und auch zu einer besseren Leistung führte.

STROHSCHNEIDER (1985) verwendet das auch von PUTZ-OSTERLOH (1985a, b) eingesetzte Szenario "MORO", das die Situation eines kleinen Noma-stammes in der südlichen Sahara in "holzschnittartiger Weise" simuliert. Seine Studie geht der Frage nach, inwieweit mit diesem Untersuchungsinstrument stabile Daten erfaßt werden können und welche Hinweise auf die externe Validität dieser Daten zu finden sind. Bezüglich Test-Retest-Stabilität kommt er zu dem Schluß, daß Systemzustandsmaße geringe, Verhaltensindizes dagegen höhere Reliabilitäten aufweisen. Aus einer ausführlichen Nachbefragung der Vpn schließt STROHSCHNEIDER auf die Validität der durch MORO geschaffenen Untersuchungssituation.

EYFERTH et al. (1982) stellen sich die Aufgabe, den Aufbau komplexer Handlungsmöglichkeiten - die "Genese von Handlungskompetenz" also - in einer für die Vp bislang unerprobten Situation zu untersuchen. Ein computersimuliertes System namens "WELT" dient dabei als Untersuchungsinstrument. Diese "WELT"

"existiert als Bildfolge auf einem Fernsehschirm, auf dem sich wenige Objekte rechnergesteuert und nach festen Regeln bewegen und miteinander interagieren. Der Beobachter kann über Tasten in das Geschehen eingreifen und sich in ihm handelnd orientieren. Seine Aufgabe ist es, die Regeln zu begreifen und die Objekte zielgerecht zu manipulieren." (p.2f)

Die vier mit Ziffern versehenen Quadrate können vier verschiedene Bewegungsformen ("Turm", "Läufer", "Querläufer" und "Bandenschleicher") annehmen, die bei Kollision nach einer bestimmten, herauszufindenden Regel wechseln. Außerdem gibt es spontane Richtungswechsel. Der Beobachter kann (1) die Geschwindigkeit variieren, mit der ein Quadrat über den Bildschirm "läuft", (2) die Bewegungsrichtung eines Quadrates verändern sowie (3) das System anhalten. Die Autoren beschäftigen sich bei ihren theoretischen Erwägungen mit dem stark vernachlässigten Thema des allmählichen Aufbaus einer Systemrepräsentation und dessen Zusammenhang mit "mitgebrachten" Schemata.

(b) Studien, in denen Situationsmerkmale eine zentrale Rolle spielen

Hierunter fallen Arbeiten, in denen bestimmte Eigenschaften des Simulationsprogramms als potentielle Ursachen für Verhaltenseffekte systematisch untersucht wurden. Da eine Zuweisung zu dieser wie auch zur vorangehenden Kategorie nicht immer einfach ist, bitte ich den Leser um Nachsicht bei möglichen Fehlklassifikationen.

Die "ENERGIEVERSORGUNG" privater Haushalte in der Bundesrepublik wird von ROST & VENT (1985) in einem Großsystem simuliert, in dem individuelle Präferenzen bei der Energieauswahl zeitlich und räumlich hochgerechnet werden. Den Autoren geht es um die Effekte unterschiedlicher Präsentations- und Rückmeldungsformen, die bestimmte Denkweisen (z.B. analytische oder ganzheitliche) stimulieren sollten. Die Ergebnisse deuten auf eine Überlegenheit der "visuell-ganzheitlichen" Denkweise.

Mit dem Ökosystem "MINISEE" beschäftigten sich OPWIS & SPADA (1983, 1985; vgl. auch SPADA, MAY & OPWIS 1983; OPWIS, SPADA & SCHWIERSCH 1985). Verschiedene biologische Populationsmodelle (isolierte wie auch interagierende Populationen) mit bekannten systemtheoretischen Eigenschaften wurden zu diesem Zweck ausgewählt. Die Probanden sollen sich über Zustandsänderungen im Rahmen der verschiedenen Modelle informieren und dabei Wissen aufbauen, damit sie Prognosen zukünftiger Zustände machen können und dementsprechend Eingriffe mit genau definierten Zielvorgaben treffen können. Um den komplexen Prozeß der Hypothesenbildung und -modifikation, der Informationssammlung, -gewichtung und -bewertung zu erfassen, wird das sogenannte WEIV-Paradigma (vgl. SPADA, REIMANN & HAUSLER 1983) angewandt: Wahl einer Information - Erwartung eines bestimmten Variablenstandes - Informationsvermittlung durch einen Versuchsleiter - Verwertung dieser Information. Am Beispiel des Fischbestands könnte eine WEIV-Sequenz wie folgt aussehen (nach OPWIS & SPADA 1983):

- (W): Wahl eines Ausgangs-Fischbestands im Jahr k;
- (E): Abgabe einer begründeten Schätzung über den Fischbestand im Jahr k+1;
- (I): Information über den tatsächlichen Fischbestand im Jahr k+1;
- (V): Hypothesengenerierung bzw. -änderung bezüglich des angenommenen funktionalen Zusammenhangs.

Im Unterschied zur "klassischen" Vorgehensweise betonen OPWIS, SPADA & SCHWIERSCH (1985), daß (1) ein solides Modell der Untersuchungssituation vorliegen sollte (vgl. Kap. 6), (2) die Präsentation von Systemdaten nicht nur aus Anfangs- und Folgezustand bestehen soll, sondern den dynamischen Verlauf der Variablen erkennbar werden läßt, und (3) die Datenerhebungsmethoden sorgfältiger ausgewählt werden. Hierzu zählen die Autoren insbesondere die Problematik reliabler und valider "Güteindikatoren": Uneingeschränkte Eingreifbarkeit, unbekannte "Lösbarkeit" der Steuerungsaufgabe und Unkenntnis der gedächtnismäßigen Repräsentation des Systems durch die Vp machen eine experimentelle Untersuchung anfallender Denkprozesse unmöglich. OPWIS et al. (1985) verwenden daher einen Untersuchungsplan, der hinsichtlich der genannten Aspekte ein umfassendes Maß an Kontrolle erlaubt. Ein auf die Individualdaten gerichtetes Modell über die Wissensanwendung einer Vp läßt die Vorhersage von ca. 80 % der Vp-Antworten auf ihr gestellte Fragen zum System zu.

THALMAIER (1979) greift in seiner Arbeit zur "MONDLANDUNG" einen m.E. zentralen Kritikpunkt an den bisherigen Studien mit hochkomplexen Systemen auf, indem er darauf hinweist,

"daß ohne genaue Kenntnis der Struktur des betrachteten Systems ... im allgemeinen kaum Maßstäbe für die Bewertung der Güte vorliegender Steuerungen anzugeben sind" (p.390).

Oftmals sei es entsprechend schwierig zu entscheiden, was konkret falsch gemacht sei. Auch müßte es jeweils möglich sein festzustellen, ob nicht der bisherige Verlauf der Bearbeitung eine Lösung gar unmöglich gemacht habe. Aus diesen Argumenten heraus, die THALMAIER auf dynamische Problemstellungen deterministischer wie stochastischer Art bezieht, leitet er die Schlußfolgerung ab, daß die mathematische Durchdringung des Problemtyps (d.h. des simulierten Systems) notwendige Voraussetzung für das Verständnis von Vpn-Verhalten ist. Am Beispiel des Mondlandeproblems (vgl. auch FUNKE 1981, 1983a; FUNKE & HUSSY 1984) zeigt THALMAIER (1979, p.401 ff) mehrere vorteilhafte Aspekte dieser systemtheoretischen Analyse:

- (a) die Unterscheidung zwischen Gewinnlösungen (= open-loop-Steuerung: keine Möglichkeit zur Korrektur, von Zeitpunkt t-null an auf Gewinn ausgerichtet) und optimalen Strategien (= feed-back-Steuerung: abhängig vom augenblicklichen Zustand wird der optimale Pfad gesucht, ein erfolgter Fehler kann möglicherweise kompensiert werden). Ihre Bestimmung ist bei Problemen mit unendlich vielen Situationen oder kontinuierlichen Lösungsgraphen keines-

wegs trivial, jedoch läßt sich die Bestimmung optimaler Strategien auf mathematische Standardprobleme reduzieren wie z.B. auf das Potryaginische Maximumprinzip (vgl. p.399f). Dieses liefert zwar die notwendigen Bedingungen für die Existenz optimaler Lösungen, läßt jedoch offen, ob es solche überhaupt gibt.

- (b) Die analytische Lösung des Mondlandeproblems (p.406) erlaubt die Synthese eines optimalen Steuerprogramms in Form einer sogenannten "bang-bang-Steuerung" (zunächst freier Fall bis zu einem definierten Punkt im Lösungsraum, dann Vollbremsung bis zur weichen Landung).
- (c) Die in wiederholten Problemdarbietungen ermittelten Eingriffsstrategien erlauben die Berechnung der jeweiligen Distanz zum Optimum und damit eine "exakte Quantifizierung des Lernfortschritts" (THALMAIER 1979, p.411; Unterstreichung von mir).

Die elaborierte Formaldarstellung der Überlegungen bzw. Ableitungen, die dem mathematisch weniger geschulten Leser den Zugang zunächst erschwert, führt zu interessanten Einsichten. So zeigt sich etwa, daß es den 20 Mathematikstudenten, die als Vpn zur Verfügung standen, sehr wohl gelang, im Laufe der insgesamt 20 Landungsversuche die dynamischen Aspekte des Problems sowie dessen nichtlineare Entwicklungen zu erkennen. THALMAIER (1979) kommt zu dem Schluß, "daß Vpn bei nichtlinearen Extrapolationen wohl nicht von vornherein überfordert sind" (p.418). Zu berücksichtigen ist allerdings, daß bei Problemen der optimalen Steuerung dynamischer Systeme für die Vp das Identifikationsproblem des zu steuernden Systems im Vordergrund steht, also das sukzessive Erkennen der Systemstruktur durch eine Input-Output-Analyse (vgl. dazu FUNKE & STEYER 1985).

FUNKE (1981) geht es bei seinen Vorstellungen zum Paradigma der "MONDLANDUNG" in erster Linie um die experimentellen Ausbaumöglichkeiten dieses Mini-Systems; auch ihm geht es - ähnlich wie DÖRNER und Mitarbeitern - nicht um die möglichst realistische Simulation, sondern um die "Prüfung spezifischer Hypothesen und damit eine Präzisierung der theoretischen Annahmen zum hochkomplexen Problemlösen" (p.4). Hierzu eignen sich natürlich die Paradigmen am ehesten, bei denen ohne Mühe theoretisch relevante Parameter wie etwa Problemschwierigkeit variiert werden können. Ohne hier im Detail auf die Bestimmung der Lösungsgüte oder die Möglichkeiten einer quantitativen Strategie-Klassifikation eingehen zu wollen, sei noch einmal FUNKES Einschätzung der "MONDLANDUNG" in versuchs- wie auswertungstechnischer Hinsicht in Erinnerung gerufen:

"Einerseits bestehen zahlreiche Möglichkeiten zur Erzeugung eines hochkomplexen, dynamischen und schwer durchschaubaren Systems, zum anderen bleibt das Problem für den Versuchsleiter durchschaubar (zumindest prinzipiell), wodurch die Möglichkeit zur exakten Quantifizierung des Lösungsprozesses und seiner Güte besteht." (1981, p.19).

Empirische Befunde zu diesem Paradigma stammen von FUNKE & HUSSY (1984), die diesen Problemtyp sowie eine strukturgleiche Form ("Koch-

problem") präsentieren. Behauptet wurde dort, daß Vorerfahrung in dem geforderten Realitätsbereich zu besseren Leistungen führen sollte und daß bei gleichem Vorwissen in den beiden verwendeten Paradigmen keine Unterschiede auftreten. Als Vpn wurden 24 weibliche sowie 24 männliche Vpn (je als Experten des Bereichs "Mondlandung" bzw. "Kochen" aufgefaßt) herangezogen. Die Ergebnisse fallen allerdings insofern nicht hypothesenkonform aus, als zwar die Haupteffekte "Bereich" bzw. "Vorerfahrung" nur schwache Effektstärken erreichen, die erwartete Interaktion jedoch ausbleibt. Statistische Argumente verbieten jedoch eine ausgedehnte Interpretation dieses Befunds. - Im modifizierten "Zielannäherungsparadigma" kann HUSSY (1984b) zeigen, daß die Problemlösequalität bei steigender Variablenzahl ebenso wie bei Verwendung nichtlinearer Vernetzungsfunktionen oder bei Senkung der Transparenz der Problemstellung abnimmt. Eine signifikante Korrelation zwischen Testintelligenz und Problemlösequalität findet sich dort nur unter der transparenten Bedingung mit wenig Variablen, was die Ergebnisse von PUTZ-OSTERLOH & LÜER (1981) zu stützen scheint.

Beim "ÖKOSYSTEM" (FUNKE 1985) geht es für die Vp darum, in einem fiktiven Ökosystem durch Manipulation von "Gift", "Schädlingfressern" und "Dünger" Einfluß zu nehmen auf die "Käferzahl", "Blätterzahl" und "Wasserverschmutzung". Realisiert werden fünf Durchgänge mit je sieben Takten, wobei die ersten vier Durchgänge eine beliebige Systemmanipulation erlauben ("Wissenserwerb"), während der letzte Durchgang die Erreichung eines vorgegebenen Zielzustands erfordert ("Wissensanwendung"). Durch systematische Variation zweier ausgewählter Systemeigenschaften, "Vernetztheit der Variablen" und "Grad der Zeitverzögerung", ergibt sich, daß die beiden untersuchten Einflußgrößen eine erhebliche Auswirkung sowohl auf die Repräsentationsgüte (ein diagnostiziertes "mentales Modell" der Vp über das System) wie auch auf den Grad der Zielerreichung haben, wobei die Zeitverzögerungseffekte jedoch relativ zu den Vernetztheitseffekten gesehen schwächer ausfallen.

Mit einem kleinen Räuber-Beute-System namens "PORABU" beschäftigt sich PREUSSLER (1985; vgl. PREUSSLER & DÖRNER 1985). In einem Vorhersage-Experiment kreuzt die Autorin drei Semantikbedingungen (förderliche, hinderliche bzw. keine Einbettung), zwei Prognosebedingungen (nur Räuberwerte bzw. Räuber- und Beutewerte vorhersagen) und zwei Präsentationsformen (mit und ohne Kurvendarstellung). Angesichts der gut 20 abhängigen Variablen und gut 30 Hypothesen fällt es nicht leicht, in wenigen Worten das Ergebnis der Arbeit zusammenzufassen. Die Haupteffekte der drei Faktoren auf das Prognoseverhalten sind unbedeutend; deutlichere Effekte zeigen einzelne Interaktionen. Aus einer Zusatzuntersuchung zu Rückmelde-Effekten schließt die Autorin, "daß Individuen nicht in der Lage sind, exponentielle Entwicklungsverläufe zu prognostizieren" (p.84).

Aus der kritischen Betrachtung von bisherigen Studien zum komplexen Problemlösen leiten KLUWE & REIMANN (1983; vgl. auch REIMANN & KLUWE 1983) ihren Ansatz ab, der in der gezielten Konstruktion eines abstrakten Systems namens "SIM002" besteht, das keinen Anspruch auf Realitätsnähe erhebt.

"Mit unserer Systemerstellung verfolgen wir weniger das Ziel, Realität zu simulieren, sondern wir wollen Systeme entwickeln, die an vielfältige experimentelle Fragestellungen angepaßt werden können... Wir halten es für sinnvoll, spezielle Fragestellungen zunächst einmal an reduzierten Systemen zu überprüfen." (KLUWE & REIMANN 1983, p.6).

Das System "SIM002" besteht aus zehn Systemvariablen, deren Beziehungen in einer Parametermatrix erster Ordnung festgelegt sind. Die Zustandsdarstellung erfolgt in Histogramm-Form auf dem Monitor eines Kleincomputers, die Probanden können durch einfache Manipulation zu einem Eingriffszeitpunkt beliebig viele Variablen ändern. Aufgabe des Problemlösers ist es, auf dem Bildschirm angezeigte Sollzustände zu erreichen, wobei als Gütemaßstab die Differenz des erreichten Systemzustands zum Zielzustand herangezogen wird. - Bezüglich der Fragestellung ihrer Pilot-Studie, nämlich den Effekten des Verbalisierens auf die Problemlöseleistung, kommen KLUWE & REIMANN (1983, p.36) zu folgendem, für bestimmte Untersuchungsansätze der Denkpsychologie möglicherweise fatalen Schluß:

"Durch den Versuch haben wir nachgewiesen, daß in unserer Versuchsanordnung das Erheben von Verbaldaten einen deutlichen Effekt auf die Leistungen der Versuchspersonen hat. Es ist zu vermuten, daß unsere Versuchspersonen durch das Verbalisieren erheblich gestört wurden. Unsere Konsequenz aus den Ergebnissen ist es, so weit wie möglich auf die Erhebung von Verbaldaten zu verzichten."

Eine Weiterentwicklung des abstrakten Systems SIM002 besteht in der Systemvariante "SIM003" (vgl. KLUWE, MISIAK & REIMANN 1984; KLUWE, MISIAK & SCHMIDLE 1985), in der die Variablenzahl auf 15 erhöht sowie eine Gruppierung von Systemvariablen vorgenommen wurde. Neben beliebiger Eingreifbarkeit in das System erfolgt eine taktweise Beschränkung der Zustandsdarbietung, in unregelmäßigen Abständen soll der Pb den letzten Systemzustand reproduzieren bzw. den nächsten antizipieren. Eine zentrale Annahme dieser Arbeiten besteht im Postulat verschiedener Stadien der Ordnungsbildung, die in längerdauernden Einzelfallstudien identifiziert werden können. Die ablaufenden komplexen Lernprozesse betrachten die Autoren unter der Perspektive der "chunk"-Bildung. Da aufgrund der Systemkonstruktion zu jedem Zeitpunkt ein "Idealeingriff" hinsichtlich der Zielvorgabe bestimmbar ist, läßt sich der Prozeß des Lerngewinns deutlich beschreiben. Mit der Zunahme der Leistung geht ein Zeitgewinn einher, der allerdings (wie auch die chunk-Bildung) mit erheblichen individuellen Unterschieden verbunden ist. Am Ende einer langen Systemsteuerung (200 Simulationstakte) verfügen die Pbn über verbalisierbares Systemwissen sowohl in bezug auf die Variablenvernetzung als auch in bezug auf spezifische Eigenschaften einzelner Variablen.

In diesem Zusammenhang kann auch auf eine Studie von SCHOTT, NEEB & WIEBERG (1984) zur kognitiven Repräsentation von Kleinsystemen verwie-

sen werden: Die Autoren untersuchen die Wissensrepräsentation dynamischer Systeme mit Kausal-, Regelungs- bzw. Steuerungsstruktur, ohne allerdings den Pbn Gelegenheit zum konkreten Umgang mit dem System zu geben. Die Ergebnisse machen differentielle Effekte der drei Arten von Systemstruktur auf die Anzahl der von den Pbn korrekt rekonstruierten Variablenverknüpfungen deutlich, auf die hier jedoch nicht näher eingegangen wird.

3 Thesen zur gegenwärtigen Lage der Forschung zum komplexen Problemlösen

Im Sinne einer kritischen Bestandsaufnahme der aktuellen bundesrepublikanischen Forschung zum komplexen Problemlösen werden nachfolgend einige Behauptungen aufgestellt und begründet, die gewisse Schwachpunkte bisheriger Bemühungen betreffen. Dabei sollen keineswegs die Verdienste in Zweifel gezogen werden, die sich allein schon durch die Eröffnung dieses Sujets für die bundesdeutsche psychologische Forschung ergeben haben. Bereits die Wirkung in der Öffentlichkeit dürfte nicht gering zu schätzen sein (vgl. die Berichte über "LOHHAUSEN", "DAGU" und "TAILORSHOP" in Presse und Fernsehen). Die hier verfolgte fachinterne Kritik soll in erster Linie Anregungen für eine fruchtbare Weiterbehandlung dieses herausfordernden Themas geben und so den konstruktiven Aspekt dieses Vorgehens betonen.

An den Anfang jedes der acht folgenden Unter-Kapitel wollen wir zunächst die These stellen und daran anschließend die Begründung liefern. Die Thesen selbst sind natürlich nicht alle gleichwertig: Einige sind neben-, andere unter- bzw. übergeordneter Natur; auch die Reihenfolge sollte nicht unbedingt als Gewichtung angesehen werden. Eine ältere Fassung dieser Thesen liegt mit der Arbeit von FUNKE (1984b) vor, zu der es eine kurze Stellungnahme (DÖRNER 1984) gibt.

3.1 Die These von der Theoriearmut

Die bundesdeutsche Forschung im Bereich des komplexen Problemlösens wird momentan theoriearm betrieben. Die Phase des explorierenden Vorgehens muß abgelöst werden durch eine Phase der Modellbildung und Modelltestung.

Die "LOHHAUSEN"-Studie von DÖRNER, KREUZIG, REITHER & STÄUDEL (1983), die als hervorsteckende Arbeit dieses Forschungszweiges herangezogen werden kann, belegt unsere These in eindrucksvoller Weise: An den insgesamt 48 Probanden, die das computersimulierte System der fiktiven Stadt "LOHHAUSEN" als Bürgermeister lenken durften, fielen jeweils ca. 100000 Daten an (p.137). Bei den vielen zusätzlich erhobenen Persönlichkeitsmerkmalen wird explizit von einem "Schrotschuß" (p.119) gesprochen, durchgeführt mit der Absicht, "die Ermittlung von statistischen Zusammenhängen zwischen den situativen Merkmalen und den Testergebnissen einerseits und Verhaltensmerkmalen im Versuch andererseits" (p.119) zu ermöglichen. Als Schutz vor zufällig bedeutsamen Korrelationen beschreiben die Autoren nun eine "Methode der theoretischen Konsistenz" (p.137ff), die -verkürzt dargestellt- folgendermaßen vorgeht:

- (a) heuristisch geleitete Auswahl eines Datenausschnitts;
- (b) die dort ermittelten Resultate bilden die "Induktionsbasis" für ein zu erfindendes "theoretisches System" als mögliche Erklärung für die gefundenen Zusammenhänge; neben der Erklärung von Daten

der Induktionsbasis muß das theoretische System Prognosen über Variablen ermöglichen, die aus dem noch nicht ausgeschöpften Variablenpool stammen;

- (c) Überprüfung dieses theoretischen Systems an weiteren empirischen Daten der Stichprobe (der sogenannten "Prüfbasis");
- (d) bei Abweichungen: Veränderungen im theoretischen System vornehmen und an anderer Prüfbasis testen, die alte Prüfbasis wird dabei Bestandteil der neuen Induktionsbasis.

Dieser Wechsel zwischen Entdecken und Testen muß genauer angesehen werden, um mögliche Fehlerquellen zu erkennen. Eine wesentliche Voraussetzung derartigen Vorgehens, auf die auch DÖRNER et al. (1983) hinweisen, betrifft den Zusammenhang von Induktions- und Prüfbasis: Nur dann erfolgt ein fairer Test des entwickelten theoretischen Systems, wenn die zur Prüfung herangezogenen Daten logisch unabhängig von der Induktionsbasis sind. Diese Einschränkung bezieht sich auf die klassische Einteilung von apriorischen und aposteriorischen Erkenntnissen nach KANT sowie dessen Unterscheidung zwischen analytischen und synthetischen Urteilen. BRANDTSTÄDTER (1982, p.268) formuliert diesen Sachverhalt so:

"Theoretische Zusammenhangsbehauptungen qualifizieren sich ... nur dann als empirische Hypothesen, wenn kein Prüfergebnis von vornherein, etwa aufgrund logischer oder terminologischer Regeln, ausgeschlossen werden kann."

Nun ist es nicht immer einfach, solchermaßen empirische von analytischen Sätzen zu unterscheiden, so daß manche empirische Prüfung eher "pseudoempirisch" ausfällt. BRANDTSTÄDTER verweist darauf, daß erst die "Bezugnahme auf ein explizites System von Prädikatorenregeln bzw. Bedeutungspostulaten" Klarheit schaffen kann. Dies setzt ein so fundiertes und präzise formuliertes "theoretisches System" voraus, wie es im Bereich komplexen Problemlösens noch fehlt. Das oben dargestellte Verfahrensprinzip wird damit aber im Kern getroffen, da die logische Unabhängigkeit zwischen Induktions- und Prüfbasis erst auf der Basis präziser Definitionen und Begriffsklärungen zu überprüfen wäre. So ist es erfreulich, wenn am Ende der "LOHHAUSEN"-Studie ihre Autoren feststellen:

"Die Tatsache, daß Theorie und Daten so wunderschön zusammenpassen, ... kann ... nicht allzusehr verwundern; die Theorie folgte den Daten." (DÖRNER et al. 1983, p.445).

Die dort weiterhin gestellte Forderung, die "im Grunde doch einfache Theorie" prognostisch zu verwenden, halten wir zwar für richtig, sind jedoch von ihrer Prüfbarkeit nicht überzeugt. Die dort gemachten Zusammenhangsaussagen zwischen Problemlösefähigkeit und der Fähigkeit zu "Dependenz- und Komponentenanalyse", zu "Sub- und Superordinationsprozessen", zur "Differenziertheit und Tiefe der Gedächtnisbilder", der Fähigkeit zur "Modellbildung" usw. kann man lediglich als eine Art "theoretische Vorform" (DÖRNER) akzeptieren, aus der eine Phänomenologie der "guten" bzw. "schlechten" Problemlöser resultiert. Die an-

spruchsvolle Bezeichnung "Theorie" sollte dagegen reserviert bleiben für eine definierte "Menge von Gesetzen, die durch logische Ableitbarkeitsbeziehungen miteinander verbunden sind" (OPP 1970, p.50) und aus denen prüfbare Aussagen abgeleitet werden können. Erst auf einer solchen Ebene, auf der man Aussagen der Theorie in eine formale Sprache übersetzt, kann überhaupt entschieden werden, ob die Theorie zu widersprüchlichen Hypothesen führt oder ob eine bestimmte Untersuchung mit der Theorie und deren Prognosen in Einklang steht oder nicht. Hierzu benötigt man natürlich auch Angaben über den Geltungsbereich der theoretischen Aussagen (z.B. in bezug auf das Universum möglicher komplexer Problemstellungen oder in bezug auf das Universum möglicher "Löser" komplexer Probleme). Dies sollte in künftigen Arbeiten erfolgen, um Generalisierungsfehler zu vermindern. Tab. 3.1 zeigt im übrigen die "Deduktionen" der "LOHHAUSEN"-Studie (vgl. DÖRNER et al. 1983, p.438f) für die guten (=P-Vpn) und schlechten (=N-Vpn) Problemlöser.

Tab. 3.1: Vergleich guter (=P) und schlechter (=N) Problemlöser hinsichtlich verschiedener Merkmale (=Deduktionen der "LOHHAUSEN"-Studie).

Merkmal	P-Vp	N-Vp
ALLGEMEIN		
heuristische Kompetenz	hoch	niedrig
Tendenz zu diversiver Exploration	hoch	niedrig
Gedächtnisbilder	viele	wenige
-, Differenziertheit, Tiefe	ja	nein
Verhaltensorientierung	"Folge"	"Ergebnis"
TEILZIELBILDUNG		
Dependenz-, Komponentenanalyse	gut	schlecht
Sub-, Superordinationsprozesse	gut	schlecht
Analogieschluß-Fähigkeit	gut	schlecht
Beziehung Teil- zu Generalziel	klar	unklar
Problemsuche wegen Zielanalyse	möglich	notwendig
Wichtigkeitsprofil der Teilziele	inhomogen	homogen
ABSICHTSRANGIERUNG UND ZEITPLANUNG		
Profil der Absichtshierarchie	markant	flach
Stabilität der Rangierung	hoch	niedrig
Absichtswechsel	selten	oft
ABSICHTSBEARBEITUNG		
Fähigkeit zur Modellbildung	gut	schlecht
aktuelle Kompetenz	stabil	instabil
Wahrscheinlichkeit einer Notfallreaktion	gering	hoch

Nimmt man einmal das präziseste Theoriesegment heraus, nämlich die Theorie der Absichtsbehandlung (DÖRNER 1982), die bereits eine wesentliche Überarbeitung der Kap. 6.5 bis 6.7 aus DÖRNER et al. (1983) darstellt, stellen sich sofort einige Fragen, die eine nochmalige Revision dringend erforderlich machen. Der Auswahldruck a_i einer Absicht i wird dort (DÖRNER 1982) als Produkt aus Wichtigkeit w_i , Dringlichkeit d_i , Erfolgswahrscheinlichkeit e_i und Erledigungswert l_i sowie einer additiven Komponente, dem Aktualitätsgewicht ag_i , dargestellt:

$$a_i = w_i * d_i * e_i * l_i + ag_i \quad (3.1)$$

Eine Konsequenz von (3.1) wäre etwa, daß -sobald nur einer der vier Produktterme null wird- ausschließlich das Aktualitätsgewicht den Auswahldruck determiniert. Daß dies nicht sinnvoll ist, scheint mittlerweile sicher. Eine wesentlich verbesserte Variante dieser Theorie wurde von DÖRNER (1983c) vorgelegt: Ein Simulationsprogramm "VORG" produziert unter verschiedenen Parameterkonstellationen plausibel interpretierbare Handlungssequenzen. Unklar bleibt allerdings auch hierbei, wie dieser Entwurf empirisch geprüft werden kann, da die verwendeten Größen (z.B. der "Aktivations- bzw. Resignationswert" einer bestimmten Absicht zu einem bestimmten Zeitpunkt) nicht leicht zu erfassen sind und viele Hilfsmittel herangezogen werden (z.B. ein "Inhibitionsprozeß", der die Stabilität von Absichtsauswahlen garantieren soll), die sich ebenfalls nur aus Plausibilitätsaspekten heraus begründen lassen.

Bei allen kritischen Bemerkungen dazu sei festgehalten, daß wir die bisherigen Arbeiten sehr wohl als fruchtbare, empirische Heuristik betrachten. Diese, dem gegenwärtigen Forschungsstand wie auch dem Forschungsgegenstand vielleicht angemessene Situation muß jedoch fortentwickelt werden durch die Konstruktion theoretischer Aussagesysteme, aus denen prüfbare Hypothesen abgeleitet werden können.

3.2 Die These von der mangelnden Integration anderer psychologischer Teildisziplinen

Der Einbezug von Erkenntnissen verwandter Forschungsgebiete findet momentan nur unzureichend statt. Psychologische Teildisziplinen, die nach ihrem Selbstverständnis enge Verflechtungen zum Bereich des Problemlösens aufweisen, bleiben nach wie vor unverbunden. Die Integrationskraft des "information processing approach", auf die KLIK (1979, 1980) hinweist, kommt erst ansatzweise zum Tragen.

Ohne Anspruch auf Vollständigkeit werden im folgenden einige Forschungsgebiete aufgeführt, deren Zusammenhang mit komplexem Problemlösen stärker aufgearbeitet werden sollte. Arbeiten aus Bereichen wie "Emotion und Kognition" (vgl. KUHLE 1983, MANDL & HUBER 1983) oder "Gedächtnis" (vgl. DÖRNER 1983a), auf die sich das Interesse der Problemlöseforscher bereits gerichtet hat, zeigen exemplarisch den Vorteil solcher Bereichsüberschreitungen, deren Kehrseite in der Unklarheit darüber besteht, wie ein derartig weitreichender Anspruch einer "Totalpsychologie" einzulösen ist.

Metakognition. Neben dem Einfluß von "Selbstreflexion" auf das Problemlöseverhalten (HESSE 1979, 1982; PUTZ-OSTERLOH 1983b, 1985b; REITHER 1979) wurde auch in der "LOHHAUSEN"-Studie eine ganze Reihe der von Metakognitionsforschern für wichtig erachteten Eigenschaften ("predicting", "checking", "monitoring", "reality testing", "coordination and control"; nach BROWN 1978) untersucht, die von offenkundiger Bedeutung für die Steuerung komplexer Systeme sein dürften. Eine Rezeption der amerikanischen (vgl. FLAVELL 1979) bzw. deutschen Forschung (vgl. KLUWE 1979, WEINERT & KLUWE 1984) aus dem Bereich "Metakognition" findet jedoch nicht statt.

Streß. In Situationen mit hoher Anforderungsbelastung -und als solche müssen komplexe Problemstellungen interpretiert werden- tritt bei Erreichen der Kapazitätsgrenze Streß auf (vgl. BATTMANN 1983; HEUSER 1978; SCHULZ 1979, 1980). Vor allem im Zusammenhang mit der "intellektuellen Notfallreaktion" (DÖRNER 1981) liegt es nahe, derartige Zusammenhänge anzunehmen. Die mangelnde Aufbereitung diesbezüglicher Forschungsergebnisse bedeutet zugleich den Verzicht auf eine Vielzahl von zum Teil physiologischen Indikatoren, deren Meßgüte oberhalb der vieler Ratings liegen dürfte. - Während mit "Streß" die Spitzenwerte mentaler Belastungen charakterisiert werden, trifft die Bezeichnung "mental workload" (vgl. MORAY 1979) eher den allgemeinen Fall "normaler" Belastung. Auch dazu liegen interessante Befunde vor.

Sprachpsychologie. Angesichts der intensiven Analyse von Verbalprotokollen (vgl. HUBER & MANDL 1982) und der umfänglichen verbalen Interaktion zwischen Versuchsleiter und Proband verwundert das fast völlige Fehlen von Überlegungen zu diesem Punkt. Während die Intelligenzforschung aus eben diesen Erwägungen heraus ihr Testmaterial sprachfrei gestaltet hat, führt die Problemlöseforschung nunmehr extensiv sprachliche Äußerungen in den Forschungsrahmen ein, ohne die daraus resultierenden Überlagerungseffekte kritisch zu reflektieren (vgl. die Arbeiten zur Methode des lauten Denkens, z.B. von ERICSSON & SIMON 1980; NISBETT & WILSON 1977). Jüngst hat allerdings ROTH (1985) den Versuch unternommen, durch sprachstatistische Analysen Hinweise auf den "Denkstil" von Pbn zu gewinnen. Dieses Vorhaben könnte -sofern andere Kriterien als momentan verwendet für die Gruppenbildung zum Einsatz gelangen- interessante Aufschlüsse über die sprachlichen "Niederschläge" des Umgangs mit komplexen Situationen liefern.

Methodologie und Methodik. Möglicherweise aus einer skeptischen Grundhaltung zum klassischen Experiment heraus (exemplarisch KREUZIG 1983), die bei einem Teil der Forscher zu vermuten ist, findet eine Rezeption neuerer Vorstellungen über die Feststellbarkeit von Kausalbeziehungen auch in komplexeren experimentellen wie nicht-experimentellen Kontexten sichtbar nicht statt. Gerade die Diskussion der seit den 70er Jahren zunehmend elaborierten Strukturgleichungsmodelle (vgl. GOLDBERGER & DUNCAN 1973) und der damit verbundenen Möglichkeit zur Beschreibung von Kausalmodellen mit wechselseitiger Abhängigkeit sollte berücksichtigt werden. Auch in nicht-experimenteller Forschung kann unter bestimmten Prämissen (vgl. STEYER 1983a, b, c) ein kausales Mo-

dell auf seine Verträglichkeit mit den Daten hin geprüft werden. Auch andere Entwicklungen der letzten Jahre, insbesondere im Bereich diskreter multivariater Datenanalyse (z.B. "latent class analysis"; zur Einführung: LANGEHEINE 1980, 1982; zum Problem der Modellprüfung: ERDFELDER 1984), finden in den einschlägigen Arbeiten keinen Niederschlag; dort sieht man selten mehr als bivariate Korrelationen oder einfache Varianzanalysen, obwohl die Zusammenhangshypothesen multiple Ursachen und multiple Wirkungen postulieren. Dies ist kein Plädoyer für den ungehemmten Einsatz neuester statistischer Entwicklungen, sondern für eine hypothesengeleitete Datenauswertung: Wenn die wissenschaftliche Hypothese adäquat in die statistische Hypothese eines Mittelwertunterschieds überführt werden kann, ist die Verwendung eines t-Tests völlig ausreichend.

3.3 Die These von der fehlenden fachübergreifenden Kooperation

Trotz gegenteiliger Erwartungen findet keine fachübergreifende Kooperation statt, die vom Gegenstand her jedoch gefordert wäre. Gedacht ist dabei an die Fächer Mathematik, Informatik und Systemwissenschaft sowie -je nach Simulationsgegenstand- an die entsprechende Fachdisziplin (z.B. Betriebswirtschaft, Volkswirtschaft, Politikwissenschaft).

Im Unterschied zur vorangegangenen These, in der das Fehlen fachinterner Überschreitungen von Teildisziplinen der Psychologie bemängelt wurde, geht es hier um den Einbezug anderer Wissenschaftszweige. Zusammenarbeit mit Informatikern, Mathematikern, Systemwissenschaftlern und -je nach simuliertem System- mit Betriebs- oder Volkswirten oder Politikwissenschaftlern halten wir für notwendig und günstig, um nicht aus den nach Augenschein validen Simulationsprogrammen Fehlschlüsse zu ziehen. Derartige Kooperation mag auch schon vor jeder Empirie dazu beitragen, kritische Konzepte zu präzisieren und so die Modellbildung zu erleichtern, natürlich auch durch die Übernahme von Modellvorstellungen, die in anderen Disziplinen erfolgreich Verwendung finden. Eine derartige Kooperation läßt sich (zumindest aus der jeweils angegebenen Literatur) nicht erkennen.

Dabei könnte man vor allem von den Erfahrungen anderer Disziplinen beim Einsatz computersimulierter Systeme profitieren. Insbesondere Volks- und Betriebswirte verfügen schon seit den 60er Jahren über zahlreiche Simulationsprogramme. ZUCKERMAN & HORN (1973; zitiert nach WOLFE 1978) berichten bereits über mehr als 600 vorliegende "business games" (siehe auch ELGOOD 1982²), viele Untersuchungen mit derartigen Simulationsprogrammen sind in den einschlägigen Zeitschriften wie "Management Science" (z.B. HAND & SIMS 1975), "Simulation & Games" (z.B. NORRIS & SNYDER 1982; RUBEN & LEDERMAN 1982; THOMPSON & KEON 1982) usw. publiziert. Neuere Arbeiten wie etwa OPWIS, SPADA & SCHWIERSCH (1985) machen regen Gebrauch von den fachspezifischen Erkenntnissen über ihren Simulationsgegenstand "Bodensee", wie überhaupt im Rahmen der Ökologie eine Reihe interessanter Überlegungen über dynamische Systeme angestellt wurden (vgl. RICHTER 1985).

3.4 Die These von der unzureichenden Nutzung des systemtheoretischen Ansatzes

Der systemtheoretische Ansatz wird zwar propagiert (vgl. DÖRNER 1983a), aber nur unzureichend genutzt. Dies betrifft vor allem das Verständnis der Begriffe "Komplexität", "Vernetztheit" und "Stabilität". Ihrer elementaren Bedeutung wegen erfolgt an dieser Stelle eine ausführlichere Darstellung der These.

Faßt man menschliches Handeln (und damit auch Problemlösen) im Rahmen eines systemwissenschaftlichen Ansatzes, genügt es nicht festzustellen, daß "vieles mit vielem" interagiert (DÖRNER 1983a, p.354). Gerade wenn man es damit ernst meint, sollte zuallererst das entsprechende Rüstzeug herangezogen werden, das von der Systemwissenschaft bereitgestellt wird (vgl. einfürend CASTI 1979; LUENBERGER 1979; PICHLER 1975; SILJAK 1978; VEMURI 1978; ZWICKER 1981), auch wenn sich dessen Anwendung als nicht einfach herausstellen sollte. Drei Aspekte, die von CASTI (1979) in den Mittelpunkt seiner Monographie gestellt werden, möchten wir herausheben: Komplexität, Vernetztheit und Stabilität (im Englischen spricht man von den drei C's: "complexity", "connectivity" und "catastrophe").

3.4.1 Komplexität

Zunächst ein Zitat:

"Of all the adjectives in common use in the systems analysis literature, there can be little doubt that the most overworked and least precise is the descriptor 'complex'. In a vague intuitive sense, a complex system refers to one whose static structure or dynamical behavior is 'unpredictable', 'counterintuitive', 'complicate', or the like. In short, a complex system is something quite complex - one of the principal tautologies of systems analysis!" (CASTI 1979, p.40f).

Während der Begriff "Komplex" in der Psychologie lange Zeit von den Psychoanalytikern okkupiert wurde zur Bezeichnung von tabuisierten, stark affektiv besetzten Erlebnisgehalten, hat sich in neuerer Zeit tatsächlich der systemwissenschaftliche Bedeutungsgehalt in den Vordergrund geschoben - mit all der mangelnden Präzision, die CASTI in dem genannten Zitat anführt. Zum Beleg dafür wiederum ein paar Stellen aus "LOHHAUSEN" (DÖRNER et al. 1983), dem Buch, das unser Stichwort "Komplexität" im Untertitel führt:

"Das System Lohhausen war relativ komplex, d.h. es bestand aus vielen Variablen (ca. 2000), die in vielfältiger Weise miteinander verknüpft waren (...)." (p.136).

"Der Realitätsausschnitt ist komplex, d.h. der Akteur kann in der

ihm zur Verfügung stehenden Entscheidungszeit selbst diejenigen Merkmale des Realitätsausschnitts nicht feststellen, die an sich feststellbar sind, da deren Zahl zu groß ist." (p.26).

"Komplexität eines Realitätsausschnitts ist eine objektive Größe. Sie hängt ab von der Anzahl der Elemente eines Systems und ihren Verknüpfungen. Ein grobes Maß für die Komplexität eines Systems wäre das Produkt aus der Anzahl von Elementen und der Anzahl der zwischen ihnen auffindbaren Verknüpfungen." (p.44).

Vergleicht man diese Aussagen der Autoren mit dem Wissensstand sieben Jahre zuvor, scheinen sich in der psychologischen Komplexitätsforschung kaum Fortschritte bei der Klärung des Komplexitätskonzepts ergeben zu haben:

"Die Komplexität eines Sachverhalts hängt einmal von der Anzahl von Komponenten ab, die an ihm unterschieden werden können. Zum anderen wird sie bestimmt durch die Vielfalt der Verknüpfungen zwischen den Komponenten. Der Grad der Komplexität eines Sachverhalts beeinflusst den Problemlöseprozeß dadurch, daß von einem bestimmten Komplexitätsgrad an komplexitätsreduzierende Maßnahmen seitens des Problemlösers notwendig werden, da der Sachverhalt im Ganzen nicht mehr überschaubar ist." (DÖRNER 1976, p.18).

Natürlich finden sich ähnliche Aussagen in fast allen Arbeiten der Arbeitsgruppe "Systemdenken", nirgendwo jedoch findet man eine tiefergehende Präzisierung dessen, was unter "Komplexität" nun genau verstanden werden soll: "In short, a complex system is something quite complex" (CASTI 1979, p. 40).

Aber auch von anderen Forschern auf diesem Gebiet finden wir nichts dergleichen. Der von UECKERT & RHENIUS (1979) edierte Band über die Hamburger Tagung "Kognitive Psychologie" von 1978, bezeichnenderweise betitelt "Komplexe menschliche Informationsverarbeitung", enthält nirgendwo eine Definition von Komplexität oder auch nur den Versuch dazu. Sollte dies daran liegen, daß es tatsächlich nicht möglich wäre, eine präzisere Beschreibung dessen zu liefern, was man unter der Bezeichnung "komplexes System" versteht? Mit "präzise beschreiben" meine ich dabei: über die umgangssprachliche Charakterisierung hinauskommen, nach der "zwar nicht alles mit allem, aber vieles mit vielem" (DÖRNER 1983a, p. 354) zusammenhänge.

Ich möchte damit keine überzogenen Ansprüche wecken: Die gestellte Aufgabe ist nicht leicht zu bewältigen! Eher allgemeine Beiträge zum Thema "Komplexität" findet man z.B. bei BEER (1970), MORIN (1974), ROSEN (1975, 1977) und WALLER (1982). Mit der Identifikation von Systemen beschäftigen sich die Arbeiten von ASTRÖM & EYKHOFF (1971), GAINES (1977) sowie KLIR (1976). Eher grundlegend sind die Arbeiten von FERDINAND (1974), LÖFGREN (1977), SAHAL (1976), TCHON (1983) sowie TCHON & WOJCIECHOWSKA (1981). Tests für Systemkomplexität beschreibt GEORGE (1977), ZIMMERMANN (1976) geht auf unscharfe ("fuzzy") Systeme ein. Ich will an dieser Stelle nicht mehr versuchen, als drei Diffe-

renzierungen des Komplexitätsbegriffs vorzustellen, die in der Systemwissenschaft -übrigens schon länger- üblich sind (vgl. die Bibliographie von CORNACCHIO 1977b zu diesem Thema). Dabei handelt es sich um die Unterscheidung von statischer, dynamischer und Kontrollkomplexität.

3.4.1.1 Statische Komplexität und Kontrollkomplexität

Unter statischer Komplexität versteht man Eigenschaften wie die hierarchische Struktur des Systems, die Verbindungsmuster der Variablen, die Verschiedenheit der Komponenten sowie die Stärke der Interaktionen. Es geht dabei also lediglich um strukturelle Eigenschaften des Systems und der Subsysteme.

Beginnen wir mit dem Aspekt der hierarchischen Struktur. Ein hoher Grad an statischer Komplexität bedingt hohe Raten des Informationsaustauschs zwischen verschiedenen Entscheidungsinstanzen; umgekehrt ergibt sich die Notwendigkeit einer hierarchischen Struktur gerade aus dem Bedürfnis nach einem angemessenen Informationsfluß sowie der Ausführung von Entscheidungen. Als Beispiel für diesen Aspekt der Komplexität verweist CASTI (1979, p. 99) auf ein klassisches Beispiel von SIMON (1962).

Zwei Uhrmacher, Chronos und Tempus, erzeugen mit gleicher Präzision Uhren aus 1000 Teilen in zweierlei Weise: Während Tempus seine Uhren sequentiell produziert und dadurch bei einer Störung jedesmal von vorne beginnen muß, teilt Chronos den Konstruktionsprozeß in Teilschritte auf. In jedem Teilschritt werden 10 Einheiten zusammengebaut. Auf der ersten Ebene entsteht so aus 10 Einzelteilen ein Bauteil. Auf der zweiten Ebene ergeben 10 Bauteile eine Komponente. Auf der dritten und letzten Ebene schließlich werden 10 Komponenten zur fertigen Uhr zusammengesetzt. Wenn Chronos gestört wird, verliert er also nur jeweils das gerade in Bearbeitung befindliche Teil der jeweiligen Ebene. Stört man ihn gerade, wenn er aus den 10 Komponenten die Uhr bildet, muß er diesen Schritt nochmals wiederholen. Die Komponenten, die aus je 100 Einzelteilen bestehen, bleiben ihm jedoch erhalten. Die Wahrscheinlichkeit für die beiden Uhrmacher, eine Uhr fertigzustellen bei gegebener Störwahrscheinlichkeit p beträgt für Tempus $(1-p)^{1000}$, für Chronos $(1-p)^{10}$. Nimmt man eine Störwahrscheinlichkeit von $p=0.01$ an, braucht auf lange Sicht Tempus 20000 mal mehr Zeit als Chronos (nach CASTI 1979, p.99).

Dieses Beispiel macht den Stellenwert der Hierarchie im Zusammenhang mit komplexen Systemen deutlich: Sie ist eine natürliche Konsequenz aus der Notwendigkeit, mit Fehlern und Mißverständnissen umgehen zu können. Den Systemanalytikern stellt sich die Aufgabe, eine derartige Struktur zu dekomponieren, also die Teilsysteme auf den verschiedenen Ebenen zu identifizieren (vgl. PICHLER 1981, SYDOW 1981).

Der zweite Aspekt statischer Komplexität, die Verbindungsmuster der Variablen, wird unter dem separaten Punkt "Konnektivität" aufgeführt,

so daß wir hier gleich zum dritten Merkmal übergehen können, der Variation der Komponenten. Das von ASHBY (1958) formulierte Prinzip, das sogenannte "law of requisite variety", besagt, daß die Vielfalt des Systemoutputs nur durch genügende Vielfalt auf seiten der Inputmöglichkeiten gesteuert werden kann. Das Verhältnis zwischen den Variationsmöglichkeiten im Output und meinen Manipulationsmöglichkeiten der Input-Variablen bestimmt maßgeblich meine Möglichkeiten der Kontrolle über das System. Kontrollkomplexität ist ein Maß für die Fähigkeit des Systems, Variation auf der Input-Seite in Variation auf der Output-Seite umzusetzen. Wir wollen dies an einem kleinen Beispiel verdeutlichen.

Tab. 3.2: Unterschiedliche Formen der Kontrollkomplexität. Einem Satz von Störvariablen wird ein Satz von Kontrollvariablen zugeordnet. Die Zellen der Matrizen enthalten die aus den jeweiligen Kombinationen resultierenden Zustände der Systeme S_1 (links) und S_2 (rechts).

Störvariable	Kontrollvariable			Kontrollvariable		
	α	β	γ	α	β	γ
1	b	a	c	a	a	a
2	a	c	b	a	a	a
3	c	b	a	b	b	c

Der linke und rechte Teil der Tab. 3.2 veranschaulicht zwei verschiedene Systeme S_1 und S_2 mit unterschiedlich hoher Kontrollkomplexität. Während in S_1 der Kontrolleur durch geeignete Wahl einer Kontrollvariable k e K bei Eintritt eines beliebigen Störungstyps s e S jeden gewünschten Zustand z e Z herbeiführen kann (=maximale Kontrolle), bietet S_2 nur noch beschränkte Manipulations- bzw. Steuerungsmöglichkeiten: Lediglich bei Störungstyp 3 kann der Kontrolleur zwischen Zuständen b und c wählen, in den beiden restlichen Störfällen bleiben alle Kontrollvariablen wirkungslos (=geringe Kontrolle). Bezeichnet man mit KV das Ausmaß der Kontrollvariation und mit SV die Störvariation, ist die totale Variation (tV) im Systemverhalten mindestens so groß wie das Verhältnis von Störvariation zu Kontrollvariation, also:

$$tV \geq SV/KV \quad (3.2).$$

Für System 1 stehen den drei Störmöglichkeiten drei Kontrollmöglichkeiten gegenüber, das Verhältnis ist also 3:3. Im System 2 finden wir nur eine Störvariable (nämlich die 3), auf die überhaupt reagiert werden kann; Kontrollmöglichkeiten bieten uns nur zwei der drei Kontrollvariablen, nämlich Gamma in Verbindung mit einem der beiden anderen. Das Verhältnis SV zu KV beträgt hier also 1:2.

Die Kenntnis von KV und SV dürfte bei der Beurteilung von Leistungen in computersimulierten Systemen sehr wichtig sein, da im Falle hoher SV und geringer KV schon einfache Systeme kaum mehr beherrschbar

sind, menschliches "Versagen" somit keine rein psychologisch relevante Ursache mehr besitzt, sondern systembedingt auftritt.

Der letzte Aspekt statischer Komplexität betrifft das Ausmaß an Interaktion zwischen den verschiedenen Systemkomponenten. Darauf geht das nächste Kapitel unter dem Stichwort "Vernetztheit" ein, so daß damit die Betrachtung der statischen Komplexität abgeschlossen werden kann.

3.4.1.2 Dynamische Komplexität

Einer anderen Art von Komplexität begegnet man unter Verweis auf die Dynamik des zu betrachtenden Systems. Wie CASTI (1979, p.102) ausführt, ist ein intuitives Hauptmerkmal komplexer Vorgänge die Tatsache, daß der Output keine einfache Funktion des Inputs darstellt. Ein System kann strukturell durchaus einfach sein, also über eine niedrige statische Komplexität verfügen, und dennoch mag die Systemdynamik hoch komplex sein. Es kann sogar Fälle geben, in denen der erzeugte Systemoutput trotz strenger Determination nicht von einem Bernoulli-Prozeß unterschieden werden kann (vgl. das Beispiel bei CASTI 1979, p.103).

Zunächst kurz einige Ausführungen zur Terminologie komplexer Systeme. Als Beispiel sei das Multiplikator-Akzelerator-Modell (SAMUELSON 1939) aus der Volkswirtschaft genannt, das die Beziehung zwischen Volkseinkommen, Konsum und Investition beschreibt (vgl. ZWICKER 1981, p.31ff).

$$C(t) = ax * Y(t-1) \quad (3.3)$$

$$I_i(t) = bx * (C(t) - C(t-1)) \quad (3.4)$$

$$Y(t) = C(t) + I_i(t) + I_a(t) \quad (3.5)$$

wobei $C(t)$ = Konsum in Periode t ,
 $Y(t)$ = Volkseinkommen in Periode t ,
 $I_i(t)$ = durch Konsum induzierte Investitionen in Periode t ,
 $I_a(t)$ = autonome Investitionen in Periode t ,
 ax, bx = Gewichtungsfaktoren.

Zunächst einmal erkennt man an den drei Formeln (3.3), (3.4) und (3.5) eine Eigenschaft dynamischer Modelle, nämlich die Indizierung der Variablen durch einen Zeitindex t . Diese Indexierung kann zwei verschiedene Strukturen aufweisen, einmal eine zeitkontinuierliche und zum anderen eine zeitdiskrete. Diese Unterscheidung ist jedoch nicht prinzipieller Natur, da jedes zeitkontinuierliche Modell auch zeitdiskret dargestellt werden kann. Für das Verständnis der Zeitindexierung ebenfalls wichtig ist die Annahme der Äquidistanz zwischen benachbarten Zeitpunkten. Am Beispiel der Investitionsfunktion (3.4) sieht man eine für dynamische Modelle typische Schreibweise in Form einer Diffe-

renzungsgleichung. Diese Form besitzt gegenüber der mathematisch äquivalenten Darstellung in Form von Differentialgleichungen den Vorteil erhöhter Anschaulichkeit, der einfachen Verwendung von Parameterschätztechniken sowie einfacheren und flexibleren Simulation auf einem Computer (vgl. ZWICKER 1981, p.30f).

Einige Kriterien zur Unterscheidung verschiedener Variablentypen sowie zur Charakterisierung von dynamischen Modellen seien nur kurz erwähnt. Neben der Unterscheidung endogener und exogener Variablen (letztere erhalten ihre Werte von außen vorgegeben) differenziert man zwischen verzögerten und unverzögerten Variablen, letztere erkennbar am Zeitindex t , erstere durch einen von t verschiedenen Index markiert. Die Gruppe der exogenen sowie der endogen verzögerten Variablen nennt man auch vorherbestimmte Variablen. Mit dem Grad einer Differenzengleichung bezeichnet man die größte Zeitverzögerung, die eine der verzögerten Variablen gegenüber t aufweist. Eine reduzierte Gleichung liegt vor, wenn eine endogene Variable durch eine Gleichung beschrieben wird, die nur vorherbestimmte Variablen enthält. Als Endgleichung schließlich bezeichnet man den eventuell durch Umformung eintretenden Fall der Abkoppelung einer endogenen Variablen vom übrigen System, wenn diese endogene Variable also allein von ihren eigenen verzögerten Wirkungen sowie von den verzögerten und unverzögerten Wirkungen exogener Variablen abhängig ist.

Ein Hinweis sei erlaubt auf die verschiedenen Darstellungsformen dynamischer Systeme. Wie Abb. 3.1 zeigt, kann ein und dasselbe System in Form eines TINBERGENSchen Pfeildiagramms, eines Blockdiagramms, eines Kausaldiagramms oder eines Signalflußdiagramms dargestellt werden. Inwiefern die dadurch erzeugten subjektiven Repräsentationen des Systems identisch sind, ist ungeklärt und untersuchenswert.

Natürlich gibt es in jeder Gruppe eine Reihe von Variationsmöglichkeiten, die hier nicht im Detail aufgezeigt werden sollen. Wichtig scheint mir der Hinweis, daß das in bisherigen Studien zum Thema "komplexes Problemlösen" verwendete komparative Kausaldiagramm (siehe Abb. 3.1) nur eine von mehreren Möglichkeiten nutzt, wobei sich sehr wohl die Frage stellen läßt, welche unterschiedlichen subjektiven System-Repräsentationen aus den jeweils unterschiedlichen Darstellungsmöglichkeiten dynamischer Systeme resultieren. Ganz offenkundig macht etwa das Pfeildiagramm den Aspekt der verzögerten Wirkung von Maßnahmen erheblich deutlicher als das Kausaldiagramm ohne Zeitindex. Gerade das bisher ausschließlich verwendete Kausaldiagramm ist ausgerechnet die einzige Systemdarstellung ohne erkennbare Dynamik; sollte dies eine Erklärung dafür sein, warum die Probanden die Dynamik zeitverzögerter Variablen trotz "transparenter" Bedingung nicht erkennen?

Nach diesen (unvollständigen) Bemerkungen zu verschiedenen Aspekten von Komplexität möchte ich übergehen zum Stichwort "Vernetztheit".

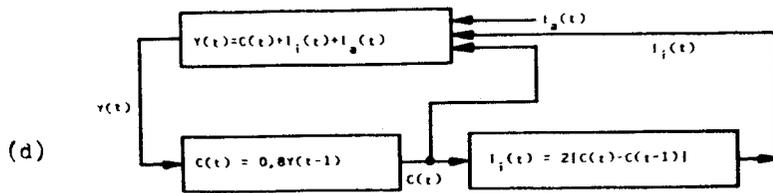
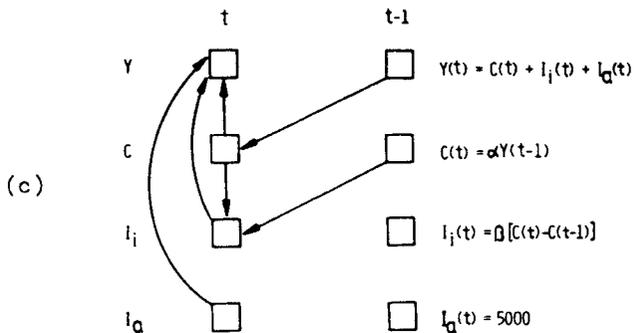
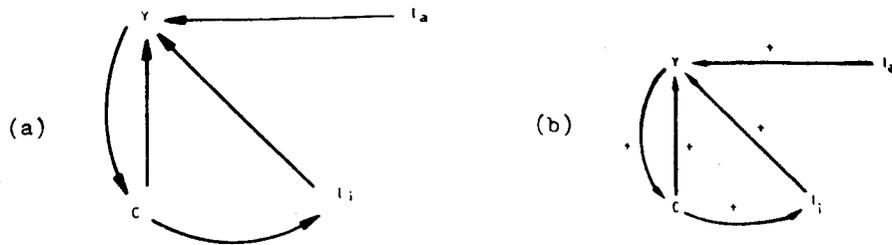


Abb. 3.1: Darstellungsformen eines dynamischen Systems (nach ZWICKER 1981).

3.4.2 Vernetztheit

Vernetztheit ("connectivity") ist oftmals implizit definierendes Merkmal des Begriffs "System": KLIR (1972, p.1) nennt z.B. ein System "... an arrangement of certain components so arranged as to form a whole.". Man kann es auch andersherum formulieren: Stehen Komponenten nicht in einem Zusammenhang miteinander, wird man nicht von einem System sprechen. Vernetztheit soll also die Verflechtungen und deren Einfluß auf den Verhaltensablauf des Systems angeben. CASTI (1979, p.35f) schreibt dazu:

"The essence of the connectivity issue is to understand the mathematical structures describing how the components of a system S are related to each other."

Eine Veranschaulichung der Konnektivität erlaubt beispielsweise die graphentheoretische Darstellung oder die - äquivalente - "incidence matrix" (Inzidenzmatrix; siehe Abb. 3.2).



"Räuber"	"Beute"					
	V	F	I	G	A	
Vogel	V	0	0	1	1	0
Fuchs	F	1	0	1	0	0
Insekt	I	0	0	0	1	0
Gras	G	0	0	0	0	0
Antilope	A	0	0	0	1	0

Abb. 3.2: Beispiel einer einfachen Nahrungskette: (a) grafische Darstellung, (b) "incidence matrix" (nach CASTI 1979).

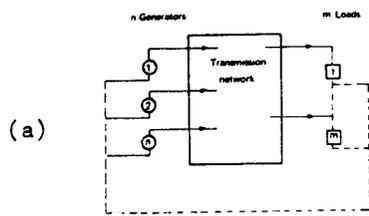
Bei den in Abb. 3.2 gezeigten Darstellungsformen einer Nahrungskette handelt es sich um einfache Verbindungen binärer Art (vorhanden / nicht vorhanden); selbstverständlich können hier auch numerische Werte größer Eins eingetragen werden, um der Verbindungsstärke Rechnung zu tragen. Die gezeigten Darstellungsmöglichkeiten geben übrigens auch DÖRNER et al. (1983, p.42 und 46) für ein fiktives Beispiel an, ohne es im Verlauf der weiteren Arbeit nochmals aufzugreifen. Dabei bestehen z.B. Möglichkeiten, einen derartigen Komplex in sogenannte "Simplexe" aufzulösen und deren Dichte untereinander zu bestimmen (Ermittlung der Homotopie von Wegen; REINHARDT & SOEDER 1982^s, p.237).

Im gezeigten Beispiel gibt es zwei Simplexe: den Vogel-Simplex, der aus den Scheitelpunkten von Insekt und Gras gebildet wird, sowie den Fuchs-Simplex, bestehend aus den Beutetieren Vogel und Insekt. Natürlich kann man auf jeder derartigen Inzidenzmatrix von Input- und Output-Variablen zwei Relationen definieren, nämlich einmal aus der Sicht des Zeilen- und einmal aus der Sicht des Spaltenmerkmals. Diese beiden einander ergänzenden Aspekte sind Gegenstand der von CASTI (1979, 1982) präferierten q-Analyse, die auf ATKIN (1974, 1977) zurückgeht. Kurz gesagt geht es dabei um folgendes: Betrachtet man ein komplexes System als eine Art mehrdimensionalen Schweizer Käse, liefert die q-Analyse dieses Komplexes Aussagen über die Verbindungen der Substanz. Die Analyse des Systems kann sich gleichermaßen auf die

"Löcher" wie auf die Substanz erstrecken.

An einem Beispiel sollen die Möglichkeiten der q-Analyse einmal verdeutlicht werden. Man stelle sich ein Stromversorgungsunternehmen vor, das eine bestimmte Anzahl von Generatoren betreibt und eine bestimmte Anzahl an Abnehmern zu versorgen hat. Aus technischen wie aus Sicherheitsgründen ist der Verbrauch der einzelnen Abnehmer nun in unterschiedlichem Ausmaß auf die stromerzeugenden Generatoren verteilt. In der Abb. 3.3 findet man neben der Skizze des Systems die gewichtete Inzidenzmatrix, die den prozentualen Anteil jedes Abnehmers an den Generatoren anzeigt, sowie binäre Inzidenzmatrizen auf verschiedenem Auflösungs-niveau ("slicing").

Wie kann man nun diese (statische) Systemstruktur auf ihre Vernetztheit hin untersuchen? Zunächst einmal kann man die Gewichtungsmatrix aus Abb. 3.3b durch eine binäre Relation m in einer veränderten Form darstellen, in der eine 1 in der Matrix des kartesischen Produkts aus x und y bedeutet, daß zwischen Generator y_i und Abnehmer x_i eine Beziehung bestimmten Ausmaßes besteht. Im Beispiel finden sich zwei verschiedene "Schnitte": Einmal bedeutet eine 1 das Vorliegen einer Generator-Abnehmer-Beziehung mit einem Gewicht von mindestens 1% (= $U(1\%)$; siehe Abb. 3.3c), ein andermal das Vorliegen einer derartigen Beziehung mit dem Mindestgewicht von 20% Stromabnahme von Generator y_i (= $U(20\%)$; siehe Abb. 3.3d). Bei einem "slice"-Wert von 61% und größer würde eine Nullmatrix resultieren: Es gäbe dann keinen Abnehmer, der 61% oder mehr seines Bedarfs aus einem einzelnen Generator y_i befriedigen würde.



(b) Gewichtete Inzidenzmatrix U(x,y) mit x (loads) und y (generators).

i	X (loads)					
	x_1	x_2	x_3	x_4	x_5	x_6
y_1	30	20	10	10	30	20
y_2	10	30	0	20	0	20
y_3	40	50	40	20	0	10
y_4	0	0	20	10	60	0
y_5	20	0	30	40	10	50

(c) Binäre Inzidenzmatrix $U(1\%) = y_3$.

μ	x_1	x_2	x_3	x_4	x_5	x_6
y_1	1	1	1	1	1	1
y_2	1	1	0	1	0	1
y_3	1	1	1	1	0	1
y_4	0	0	1	1	1	0
y_5	1	0	1	1	1	1

$Q(1\%) = \begin{pmatrix} 5 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$

(d) Binäre Inzidenzmatrix $U(20\%) = y_3$.

μ	x_1	x_2	x_3	x_4	x_5	x_6
y_1	1	1	0	0	1	1
y_2	0	1	0	1	0	1
y_3	1	1	1	1	0	0
y_4	0	0	1	0	1	0
y_5	1	0	1	1	0	1

$Q(20\%) = \begin{pmatrix} 3 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$

Abb. 3.3: Beispiel einer q-Analyse: (a) Skizze des Systems, (b) gewichtete Inzidenzmatrix, (c) und (d) Verbindungen auf unterschiedlichem Niveau mit zugehörigen Strukturvektoren (aus CASTI 1982, p.202f).

Den so gebildeten Komplex $k_y(x;m)$, der die Vernetztheit der Zeilen zu analysieren erlaubt, und den Komplex $k_x(y;m)$, der die Vernetztheit der Spalten zu analysieren erlaubt, untersucht man nun auf seine Simplexe und deren Verbindungen hin. Simplexe werden in der algebraischen Topologie definiert: Sie sind Bausteine von Polyedern, die man sich als aus einer endlichen Anzahl von "Ecken", "Kanten", "Dreiecksflächen" und -im R^3 - "Tetraedern" aufgebaut denken kann. Je nach Dimension unterscheidet man 0-, 1-, 2- oder allgemein k-dimensionale Simplexe. "Ecken" sind z.B. 0-dimensionale Simplexe. Höherdimensionale Simplexe sind aus Simplexen niedrigerer Dimensionen aufgebaut: der 1-dimensionale Simplex "Kante" enthält als sog. Seitensimplexe seine zwei "Ecken". Für die Betrachtung komplexer Systeme ist nun wichtig, daß eine endliche Menge von Simplexen als Simplizialkomplex bezeichnet wird (vgl. REINHARDT & SOEDER 1982⁵, p.243), der aus einer Reihe von Teilkomplexen besteht, die wiederum auf Simplexe rückführbar sind. Zwei wichtige Eigenschaften sind an die Definition des Simplizialkomplexes geknüpft: (1) der Durchschnitt je zweier Simplexe ist entweder leer oder ein Seitensimplex beider Simplexe, und (2) mit jedem Simplex ist auch jeder seiner Seitensimplexe aufgeführt. Die Exzentrizität eines Simplex ist eine Kennzahl seiner relativen Bedeutsamkeit für den gesamten Komplex sowie seine relative Bedeutung als Agent der Verknüpfung.

Eine weitere wichtige Angabe liefert die q-Analyse in Form des sogenannten Strukturvektors (siehe Abb. 3.3c und d, unterer Teil: $Q(1\%)$ und $Q(20\%)$). Für jede der möglichen Dimensionen gibt dieser Vektor an, wieviel verschiedene Simplexe unverbunden vorliegen. Im vorliegenden Fall findet sich auf der Analyseebene $U(1\%)$ -d.h. Generator-Abnehmer-Verhältnisse von mindestens 1%- sowohl für k_y wie auch für k_x maximale Konnektivität für die Dimensionen 5 bis 0 bzw. 4 bis 0. Auf der Analyseebene $U(20\%)$ -hier fallen alle x-y-Beziehungen kleiner als 20 heraus- zeigt sich dagegen Diskonnektivität für höhere Dimensionen. Dies besagt nichts anderes als daß es bei $U(20\%)$ nicht mehr möglich ist, z.B. den Ausfall eines Generators durch einen anderen zu ersetzen, da keine Verbindungen mehr bestehen.

Die Strukturvektoren werden nach folgendem Verfahren bestimmt. Zunächst wird entsprechend dem gewünschten Auflösungs-niveau x die Inzidenzmatrix $U(x)_{m \times n}$ bestehend aus den Werten 0 und 1 gebildet. Dann bildet man das Produkt aus der Matrix U und ihrer Transponierten. Man erhält eine Matrix der Form $m \times m$, in deren Zellen $U_{i,j}$ die Zahl der in Zeilen i und j gemeinsam auftretenden Einsen enthalten sind. Die Diagonalelemente der neuen Matrix geben die jeweilige Dimension des Zeilensimplexes plus Eins an. Der Wert des größten Diagonalelements minus Eins gibt die Dimension des gesamten Komplexes (= $\dim K$) an. Nun beginnt man, auf dem höchstmöglichen Konnektivitäts-niveau $q = \dim K$ nach dem Vorliegen von Vernetzungen zu suchen. Dafür kommen alle Elemente in Frage, deren Diagonalwert größer oder gleich dem aktuellen Wert von q ist. Besteht unter allen fraglichen Elementen eine Querverbindung, so liegt der erste Wert des Strukturvektors fest: Er beträgt dann nämlich Eins. Lägen auf der untersuchten q-Ebene alle x Elemente diskonnektiv vor, nähme der entsprechende Wert des Strukturvektors den

Wert x an. Diese Prozedur wiederholt man solange, bis man auf der niedrigst möglichen Konnektivitätsebene angelangt ist. Aus den jeweils erhaltenen Angaben über die Zahl der diskonnektiven Elemente bildet man dann den Strukturvektor für den simplizialen Komplex.

In seinem Buch "Mathematical Structure in Human Affairs" beschreibt ATKIN (1974) verschiedene Anwendungen der "set theory" (=Mengenlehre). Unter einem set S wird eine Zusammenfassung bestimmter wohldefinierter Objekte, der Elemente dieser Menge, verstanden. Für jedes Element muß klar entscheidbar sein, ob es zur Menge S gehört oder nicht. Als Elemente können Zahlen, geometrische Abbildungen usw. in Frage kommen, jedoch auch die Bestandteile meines Schreibtischs, die Zahl der Alkoholleichen usw. Unterschieden werden finite von infiniten Mengen. Zwei grundlegende Ideen spielen in der Mengenlehre eine entscheidende Rolle: (a) die Idee des Elements (=die Dinge, die Mitglieder einer Menge sind), (b) die Idee der Mengenzugehörigkeit.

Die Zugehörigkeit zu einer Menge gibt zum einen an, daß ein Element bestimmte Eigenschaften besitzt, zum anderen, daß es eine Menge solcher Elemente gibt, die spezifische Eigenschaften untereinander teilen.

"Thus the 'membership' acts like a kind of cement which creates out of the individual objects some new thing which is different in quality, in kind (...). This idea constitutes our interpretation of the ancient Aristotelian view that the set is more than (different from) the sum of its parts." (ATKIN 1974, p.3).

Diese Idee der Mengenzugehörigkeit ist für das Konzept der Vernetztheit grundlegend: Dort werden die Relationen zwischen verschiedenen Mengen näher untersucht. An mehreren Beispielen demonstriert ATKIN (1974) den Einsatz verschiedener Auswertungsmöglichkeiten, so etwa am Beispiel des Schachspiels, der sozialen Beziehungen zwischen Personen (Motto: "everyman is a polyhedron"), der Struktur einer Stadt, den Beziehungen innerhalb einer Universität, der Gestalt von Kunstwerken usw. Die Struktur, die analysiert wird, besteht in der Relation zwischen zwei oder mehr Mengen; diese Struktur wird Simplizial-Komplex genannt. Indikatoren, die aus einer derartigen Struktur abgeleitet werden können, sind (zur genauen Bestimmung der nachfolgend genannten Maße siehe ATKIN 1974, Anhänge A bis D):

- (a) der bereits bekannte Strukturvektor, der die q -Vernetztheit für alle Stufen von q ($0 < q < \dim K$; letzteres bezeichnet die höchste Dimension des Simplizialkomplexes, den Rang der Matrix) angibt;
- (b) die Exzentrizität eines Simplexes als Differenz von "top- q " und "bottom- q ", den Angaben über die maximale Dimension des Simplexes bzw. über das höchste q -Niveau, auf dem der Simplex mit einem anderen Simplex verbunden ist. Diese Differenz wird am bottom- q relativiert. Ein hoher Exzentrizitätswert bedeutet demnach, daß der Simplex im Vergleich zu seinen "üblichen" Attributen über eine Reihe "außergewöhnlicher" Attribute verfügt; im Extremfall

kann dieses Maß den Wert unendlich annehmen, was totaler Diskonnektivität gleichkommt.

- (c) Ein Obstruktionsvektor für die verschiedenen Stufen der q -Vernetztheit gibt an, wie groß die Wahlmöglichkeiten beim Übergang von einem zum nächsten Simplex sind. Ein hoher Obstruktionswert läßt wenig Freiheitsgrade, ein niedriger viele.

Eine Vielzahl weiterer, hier nicht zu beschreibender Indikatoren findet sich auch in anderen Arbeiten von ATKIN (1977, 1981) sowie in den entsprechenden Passagen von CASTI (1979, 1982). Eine umfangreiche Anwendungsstudie über die politische, räumliche und personelle Struktur der Universität Essex (GB) findet sich in Sektion B des Buches von ATKIN (1977, p.70-199).

Soviel zu den Möglichkeiten der Vernetzungsanalyse. Aus den wenigen Worten dürfte deutlich werden, daß sehr wohl formale Beschreibungsmittel für die Zusammenhangsstruktur in einem komplexen System existieren und daß man seine Systemschilderung nicht auf der umgangssprachlichen Ebene enden lassen muß. Nicht eingegangen bin ich auf die vielfältigen Ansätze, die auf dem Konzept der (Trans-)Information aufbauen (vgl. ASHBY 1972; BROEKSTRA 1976; CORNACCHIO 1977a; DUSSAUCHOY 1982; HIGASHI & KLIR 1982), ebenso nicht auf die Möglichkeiten, ein komplexes System in Subsysteme zu dekomponieren. Die einschlägige systemanalytische bzw. systemtheoretische Literatur enthält dazu viele Hinweise (vgl. etwa DASTYCH 1973; PICHLER 1981; SYDOW 1981).

3.4.3 Stabilität und Katastrophe

Stabilität im umgangssprachlichen Sinne bezeichnet etwas nur schwer Veränderbares, Unbewegliches (z.B. die stabile Konstruktion eines Hauses). In der Allgemeinen Systemlehre wird dieser Begriff jedoch erheblich differenzierter verwendet, um Aussagen über Eigenschaften eines Systems zu machen (vgl. WALKER 1971; WALKER & AADRYAN 1971; WALKER & ASHBY 1966).

Stabilitätsmessungen erfolgen in drei Schritten. Zunächst einmal wird der "ungestörte" Zustand des Systems erfaßt. Dann wird eine definierte Störung auf das System gesetzt (z.B. Einheitsimpuls, Einheitsprung) und erneut der Zustand erfaßt. Überschreitet die Abweichung des gestörten Systems vom ungestörten Zustand eine bestimmte Norm, nennt man das System instabil, ansonsten wird es als stabil deklariert. So kann es etwa möglich sein, daß im obersten Stock eines Hochhauses bei starkem Wind eine Auslenkung um einen Meter auftritt, ohne daß man daher das Haus als instabil bezeichnen würde. Würde dieses Phänomen jedoch im Erdgeschoß beobachtet, wäre ein sofortiger Auszug zu empfehlen. Wie man daran erkennen kann, handelt es sich bei Angaben zur Stabilität eines Systems also stets um relative Aussagen, relativ in bezug auf die tolerierte Abweichung des Systems vom Grundzustand. Nachfolgend sollen kurz verschiedene Formen der Stabilität unterschieden werden (Abb. 3.4; vgl. ZWICKER 1981, p.75ff).

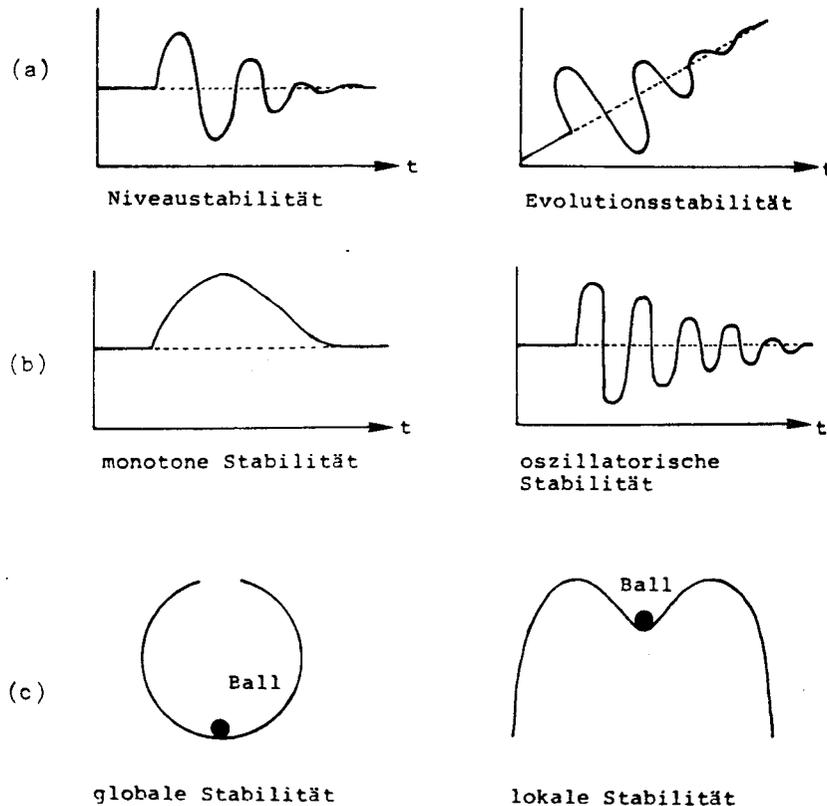


Abb. 3.4: Unterschiedliche Stabilitätsformen (nach ZWICKER 1981); auf der Ordinate in (a) und (b) ist der Wert der Variablen abgetragen, deren Stabilität untersucht wird.

Niveau- versus Evolutionsstabilität: Erstere kennzeichnet unser intuitives Vorverständnis von Stabilität als Rückkehr zum Ausgangspunkt, letztere dagegen bezieht sich auf Systeme, deren Gleichgewichtspfad wachsende oder fallende Tendenz besitzt, wo Stabilität also durch Erreichen des alten "Trends" angezeigt wird.

Monotone versus oszillatorische Stabilität: Hierbei geht es um das Rückkehrverhalten nach der Systemstörung. Schlägt das System nach der Störung in eine Richtung aus und kehrt dann langsam oder schnell wieder in den Grundzustand zurück, nennt man es monoton stabil. Oszillatorisch stabil ist das System, wenn es -wie das Pendel einer Uhr- nach der Störung abwechselnd und mit abnehmender Stärke in beide Richtungen ausschlägt.

Lokale versus globale Stabilität: Man bezeichnet ein System dann als global stabil, wenn es auch nach starken Störungen wieder in

seinen Gleichgewichtszustand zurückfindet (Beispiel: Pendel). Lokale Stabilität dagegen ist an einen Grenzwert der Störung gebunden, der nicht überschritten werden darf (Beispiel: Hochhaus nur bis Windstärke 10 stabil, danach Einsturz möglich).

Wie WEINBERG (1972, p.124) ausführt, darf man auch keinesfalls den Fehler begehen, die Stabilität eines Systems als etwas unbedingt Gutes zu deuten. Er führt als Beispiel den Schmelbrand in einem Flöz bei Ohio an, der seit mehr als 40 Jahren andauernd brennt - Beleg für stabiles Verhalten, aber nicht für einen wünschenswerten Zustand.

Genausowenig wie man den Fehler machen darf, Stabilität als etwas Positives zu bewerten, darf man ebenfalls den Begriff der Störung nicht als etwas Negatives bewerten. ZWICKER (1981, p.78) warnt vor einer derartig restriktiven Interpretation des Störungsbegriffs, der zum ersten eine implizite Zielsetzung enthält ("Sollwert") und der zweitens mit den Systemen in Widerspruch gerät, die nicht auf die Einhaltung bestimmter Sollwerte ausgerichtet sind (z.B. bei intendiertem Gewinn-Wachstum).

Das Gegenteil stabilen Verhaltens findet man bei katastrophalen Systementwicklungen, bei denen ein System also "aus dem Ruder läuft". Die Kenntnis des Stabilitätsverhaltens von Systemen kann deutlich machen, unter welchen Bedingungen katastrophale Entwicklungen auftreten. Wiederum liefert die Topologie Mittel zur Beschreibung katastrophaler Verläufe. Die topologische Katastrophentheorie, entwickelt von THOM (1975), eher allgemeinverständlich vorgetragen von ZEEMAN (1976, 1977), versucht die Variation einer oder mehrerer Verhaltensvariablen durch die Variation einer oder mehrerer Kontrollvariablen zu beschreiben, wobei unter bestimmten Bedingungen kleine Veränderungen der Kontrollvariablen große Veränderungen der Verhaltensvariablen nach sich ziehen können ("catastrophic jumps"). Diese Vorstellung ist für den Systemforscher wichtig, da möglicherweise festgestellte Schwächen menschlicher Individuen bei der Handhabung von Systemen schon durch geringfügige Änderungen der Kontrollvariablen bedingt sein können. Auch konservative Strategien können hierdurch scheitern. Ohne allzusehr ins Detail zu gehen, sollen ganz kurz wichtige Implikationen dieser Modellierung genannt sein (vgl. die Übersicht bei STEWART & PEREGOY 1983).

Katastrophentheorie ist eine allgemeine Theorie über diskontinuierliche Zusammenhänge, von CASTI (1982) weiterentwickelt zu einer mathematischen "theory of surprise". Innerhalb der Katastrophentheorie werden verschiedene Katastrophentypen unterschieden (vgl. Abb. 3.5), je nach der Anzahl der beteiligten Variablen. Wir wollen uns den einfachsten Fall, eine Faltung, näher anschauen.

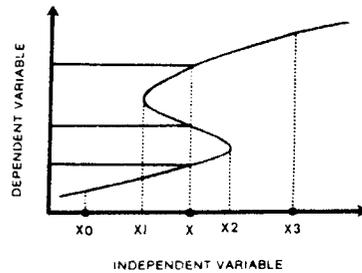


Abb. 3.5: Beispiel einer Faltung: ein Graph, für den keine einwertige funktionale Beziehung gilt (aus STEWART & PEREGOY 1983, p.337).

Erkennbar ist das Vorkommen von Verhaltenssprüngen ("catastrophic jumps"), wenn man die möglichen Werte der UV einmal durchgeht und sich die zugehörigen AV-Werte betrachtet. Auch die Eigenschaft der Hysterese (in der Physik eine Bezeichnung für die Fortdauer einer Wirkung nach Aufhören der Ursache) ist gut erkennbar: Je nachdem, ob man sich von links nach rechts oder von rechts nach links auf der Abszisse bewegt, gibt es einen unterschiedlichen Sprungpunkt. Aus diesem Phänomen resultiert (logisch paradox), daß innerhalb der "delay"-Phase zu einem UV-Wert zwei AV-Ausprägungen resultieren (Bimodalität). Zugleich wird klar, daß dadurch ein "unzugänglicher" Bereich entsteht, d.h. bestimmte AV-Ausprägungen können innerhalb der delay-Phase nicht auftreten.

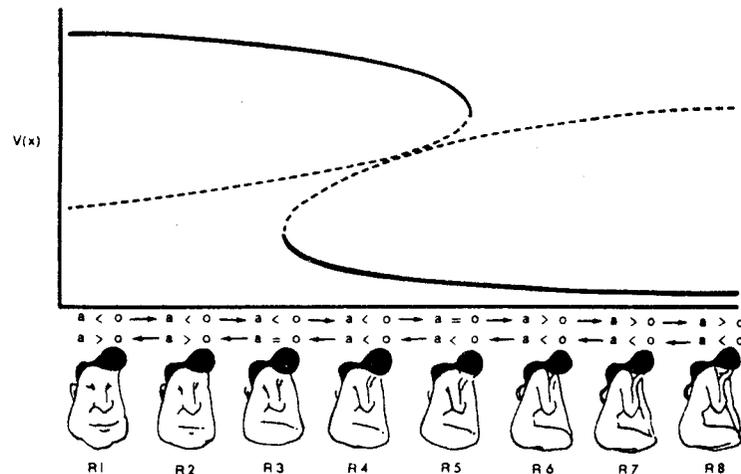


Abb. 3.6: Zwei Faltungskatastrophen durch Beginn des Experiments an verschiedenen Enden der Stimulus-Sequenz (Ordinate: Wert der Funktion; aus STEWART & PEREGOY 1983, p.343).

Ein psychologisches Beispiel aus dem Bereich der Wahrnehmungsforschung soll das eben Gesagte veranschaulichen (vgl. Abb. 3.6). STEWART & PEREGOY (1983) demonstrieren eine Faltungskatastrophe bei mehrdeutigen Figuren von FISHER, einem männlichen Gesicht (links), das sich zunehmend als weibliche Figur erweist (rechts). Fragt man bei sukzessiver Präsentation dieser Reihe Probanden nach dem empfundenen Eindruck der Männlich- bzw. Fraulichkeit der Abbildung, erhält man für auf- und absteigende Präsentationsfolge zwei unterschiedliche Antwortverläufe: Von links nach rechts vorgehend, ist R_2 die "kritische" Figur, im umgekehrten Fall R_3 . Damit ist die Hysterese-Eigenschaft des wahrnehmenden Systems deutlich demonstriert.

Mit diesem Beispiel seien die Ausführungen zum Thema "Stabilität" bzw. "Katastrophen" im Systemverhalten abgeschlossen. Aus psychologischer Perspektive heraus muß man festhalten, daß ohne tiefere Kenntnis des Stabilitätsverhaltens keine validen Schlußfolgerungen über Problemlöseverhalten beim Umgang mit komplexen Systemen gezogen werden können. Die Gefahr von Fehlinterpretationen aufgrund mangelnder Systemkenntnis -dieses Individuum ist "fähig", jenes dagegen "unfähig"- wird schnell auftreten, obwohl die "wahren" Ursachen andere als die angenommenen sein können. Inwiefern Katastrophentheorie im Kontext komplexer Probleme fruchtbar gemacht werden kann, ist noch nicht abzusehen. Kritische Stimmen (z.B. SUSSMANN & ZAHLER 1978) greifen bereits wesentliche Grundannahmen an, die für Nicht-Mathematiker nur schwer nachprüfbar sind. Andererseits liegen inzwischen psychologische Anwendungen vor (FREEDLE 1977, LEWIS 1982, SAARI 1977), die die Brauchbarkeit dieser Theorie vermuten lassen.

Insgesamt stimmen wir bei Betrachtung der These 4 mit DÖRNER (1984) darin überein, daß man zukünftig die Merkmale "komplexer" Systeme genauer beschreiben muß und zu untersuchen hat, welche Wirkung von ihnen ausgeht, um damit letztlich dem interessierenden Phänomen mit mehr Präzision begegnen zu können. Dies wird nur gelingen, wenn Rezeption und Diskussion der einschlägigen systemtheoretischen Literatur erfolgt.

3.5 Die These vom Fehlen einer Taxonomie von Problem- bzw. Systemtypen

Es fehlt eine differenzierte Taxonomie von Problem- bzw. Systemtypen, um die an unterschiedlichen Problemen (Systemen) gewonnenen Ergebnisse systematisch vergleichen zu können. Prognosen, die aus bestimmten Modellvorstellungen resultieren, können dann direkt zur Konstruktion eines entsprechenden Problemtyps mit bestimmten Eigenschaften führen.

Im Laufe der letzten 20 Jahre wurden einige Klassifikationsschemata für verschiedene Arten von Problemen vorgelegt, die in ihrer überwiegenden Mehrzahl als unbefriedigend, weil zu undifferenziert, bezeichnet werden dürfen.

DAVIS (1966) etwa verwendet nur eine Dimension, die er "Ausmaß an Versuchs- und Irrtums-Lernen" nennt. BOURNE, EKSTRAND & DOMINOWSKI (1971) verwenden immerhin drei Dimensionen: (1) Klarheit der Zieldefinition (well-defined, ill-defined), (2) Vorhandensein einer oder mehrerer Lösungen sowie (3) Abruf oder Erzeugung der Lösung. JOHNSON (1972) nennt eine Reihe wichtiger Aufgabencharakteristika wie z.B. Komplexität, Bekanntheit, Abstraktheit und Einbettung, die er als unterscheidende Kriterien bezeichnet.

Aus der Kombination extremer Ausprägungsgrade der beiden Dimensionen "Bekanntheitsgrad der Mittel" und "Klarheit der Zielkriterien" leitet DÖRNER (1976) drei verschiedene "Barriertypen" ab; implizite Bestandteile dieser Taxonomie sind aber auch Merkmale der Operation, die für die spezifischen Anforderungen des Problems von Bedeutung sind.

Drei große Gruppen von Problemen unterscheidet AEBLI (1981): (1) Probleme mit Lücke, (2) Probleme mit Widerspruch sowie (3) Probleme mit unnötiger Komplikation; alle drei Formen können sowohl abstrakter als auch konkreter Natur sein.

In einer neueren Arbeit unterscheidet HESSE (1983) drei verschiedene Arten von Problemen. Ausgehend von der Wichtigkeit der Wissensrepräsentation für die Problemlösung nennt er (1) Probleme ohne spezielle Anforderungen an die Wissensbasis, (2) Probleme mit speziellen Anforderungen an die Wissensbasis und (3) Probleme mit speziellen Anforderungen an die Informationsaufnahme, -selektion und -speicherung.

"Der erste Problemtyp ist entscheidend an der Entwicklung des Informationsverarbeitungsansatzes beteiligt, und der zweite hat erst in den letzten Jahren verstärktes Interesse gefunden. Unter Abstrahierung von vielen anderen relevanten Beschreibungsdimensionen bezeichnen wir den Typus I als 'knowledge-poor' Problembereich und den Typus II als 'knowledge-rich' Problembereich. Der uns interessierende Problembereich hat Elemente von beiden und stellt einen eigenständigen Typus III dar. ... Typus I liegt chronologisch am weitesten zurück, und das Interesse an Typus II und III hat sich danach in etwa parallel entwickelt, mit dem Schwergewicht für II in den USA und für III in Deutschland." (HESSE 1983, p.3).

Ein Prototyp für Typ I ist der "Turm von Hanoi", für Typ II das Schachspiel und für Typ III die Entwicklungslandsimulation "Dori". Während Typ I das Kurzzeitgedächtnis und Typ II das Langzeitgedächtnis (man schätzt -so HESSE-, daß Schachmeister innerhalb von 10 Jahren etwa 50000 Konfigurationen speichern!) fordert, verlangen Typ III-Probleme von beiden etwas.

Eine differenzierte Taxonomie stammt von SPEEDIE, TREFFINGER & HOUTZ (1976): Drei grobe Ebenen - Merkmale der Aufgabenwelt (Ambiguität, Anzahl möglicher Lösungen, Komplexität, Erfahrung), an der

Problemlösung beteiligte Prozesse (Präparation, Produktion, Evaluation) sowie Art der ableitbaren Maße (Zeit, Menge, Qualität, Sequenzen) - werden in jeweils mehrere Aspekte unterteilt (in Klammern angegeben), so daß insgesamt elf Kriterien resultieren. Von allen vorgestellten Taxonomien scheint es sich bislang um die differenziertere zu handeln, wenngleich gerade unter dem Aspekt der Verwendung komplexer Problemstellungen zusätzliche Ebenen hinzugenommen werden sollten (z.B. Angaben über Systemeigenschaften).

Es gibt allerdings auch Stimmen, die diesen Weg -Erstellung einer Problemklassifikation- nicht für günstig halten. OSTERLOH (1983, p.265) etwa fragt, "ob Aufgaben- beziehungsweise Problemklassifikationen sowie Kriterien für die Lösung eines offenen Problems überhaupt intersubjektive Gültigkeit haben können.". Sie begründet diese Zweifel mit dem Hinweis auf intersubjektive Unterschiede hinsichtlich Ausgangslage, Richtung und Beendigung des Suchprozesses. Der erste Aspekt betrifft also die Vorerfahrung des Problemlösers ebenso wie seine Wahrnehmung der Situation, ein unbestreitbarer Faktor interindividueller Variation, der bislang wenig beachtet wurde. Mit dem zweiten Aspekt, der Richtung des Suchprozesses, geht sie auf die Frage ein, welche Beschränkungen ("constraints") Probanden als vorgegeben ansehen und welche nicht (auch dies ist strenggenommen ein Aspekt der Vorerfahrung). Gerade in offenen Problemen sei es nicht möglich, "jeweils alle sichtbaren constraints zu problematisieren, das heißt in allen nur möglichen Richtungen die Probleme zu öffnen" (OSTERLOH 1983, p.266). In Anlehnung an LUHMANNs (1968) Konzept der "mechanisierten Entlastung" argumentiert sie, daß eine Übernahme fremder Informationsselektion unvermeidlich sei, ohne daß man bei einer Person vorhersagen könne, an welcher Stelle ein Proband diese Entlastung wählt. Mit dem dritten Aspekt, der Beendigung des Suchprozesses, sind interindividuelle Differenzen angelegt, die die Optimierungskriterien betreffen. Individuelle Wertordnungen bestimmen die Zielkriterien; wieviel Zeit und Kosten in eine vertiefte Problemlösung investiert werden, hängt damit auch von Interessen und Motiven des Problemlösers ab (vgl. das Konzept der "evaluativen Struktur" von HUSSY 1983).

Nimmt man OSTERLOHs (1983) Kritik ernst -und dies sollte man bei einer differenzierten Beurteilung tun-, fällt damit noch nicht das Bedürfnis nach einer objektiven Taxonomie von Problemen, sondern man muß im Gegenteil eine Erweiterung derselben um einen Personen- und Situationsteil fordern. Unter dieser Perspektive stellt sich auch die Problematik der Lösungsgüte-Bestimmung erneut: Während in ersten Arbeiten mit komplexen Systemen subjektive Leistungskriterien des Experimentators an alle Probanden angelegt wurden (vgl. PUTZ-OSTERLOH 1981; DÖRNER et al. 1983), was uns dazu herausgefordert hat, "objektive" Kriterien zu verlangen (FUNKE 1983b), kann man angesichts dieser Kritik versuchen, subjektive Kriterien anzuwenden, die dann allerdings mit den individuellen Zielsetzungen jedes Probanden genau abgestimmt sein müssen. Jeder Proband kann dann an seinen eigenen Kriterien gemessen werden (ipsative Diagnostik), eine Forderung nach Einzelfallanalysen also, bei deren Umsetzung man nur noch auf globaler Ebene Aussagen über Gruppen von Individuen machen könnte.

Was die Taxonomie von Systemen betrifft, liegen auch keine überwältigenden Konzepte vor. KLIR (1974) unterscheidet (1) nach dem Typ von Variablen Systeme mit Input-Output-Variablen (=kontrollierte Systeme) und solche ohne derartige Variablen (=neutrale Systeme), (2) offene versus geschlossene Systeme, (3) Systeme mit begrenzter bzw. unbegrenzter Variablenzahl, (4) materielle und abstrakte Systeme, (5) Systeme mit diskreten bzw. kontinuierlichen Variablen, (6) wohldefinierte Systeme bzw. "fuzzy systems", (7) Systeme mit und ohne Gedächtnis, (8) teleologische und nonteleologische Systeme, (9) deterministische und stochastische Systeme sowie (10) lineare und nichtlineare Systeme.

Ähnlich geht VEMURI (1978, p.53-65) vor: Er beschreibt eine Reihe von Eigenschaften, nach denen sich die Variablen eines Systems, das System selbst sowie der Modellierungsprozeß unterscheiden lassen. Aber auch dieser Ansatz liefert keine Typologie im strengen Sinn. Möglicherweise muß man auch den Anspruch an eine derartige Typologie reduzieren; die Kenntnis der psychologisch relevanten Eigenschaften von Systemen scheint mir jedoch ebenso unabdingbar zu sein wie deren Kenntnis im Bereich denkpsychologischer Untersuchungsinstrumente schlechthin.

3.6 Die These von der mangelhaften Beachtung des Meßfehlers

Das Problem der Meßfehler muß diskutiert werden. Man kann nicht einerseits den klassischen Intelligenztests mangelnde Validität vorwerfen, wenn die Reliabilität der Maße komplexen Problemlösens andererseits ungeklärt bleibt.

Zweifel an der Validität klassischer Intelligenztests sind sicherlich nicht unberechtigt. Die für bestimmte Subskalen wie auch für gesamte Tests vermuteten Invaliditäten (z.B. bei Skalen zur Erfassung des Raumvorstellungsvermögens: PUTZ-OSTERLOH & LÜER 1979; z.B. die Konstruktvalidität des CFT: LANGFELDT-NAGEL 1982) liefern den Kritikern das Material für diese Argumentation. Will man Intelligenztests durch Ergebnisse zum Problemlöseverhalten invalidieren, wie es in den Arbeiten von PUTZ-OSTERLOH (1981, 1983b, 1985b; vgl. auch PUTZ-OSTERLOH & LÜER 1981) und DÖRNER et al. (1983) versucht wird, muß man trivialerweise davon ausgehen können, daß die teilweise arbiträr abgeleiteten Maße der Problemlösefähigkeit auch tatsächlich Eigenschaften eines Maßes aufweisen. Überlegungen zum Meßfehler (vgl. ZIMMERMAN & WILLIAMS 1977) etwa, den man natürlich auch bei Maßen der Problemlösefähigkeit erwartet, finden sich in den einschlägigen Publikationen nicht, obwohl selbst das Problem der Situationsabhängigkeit von Meßwerten im Rahmen der klassischen Testtheorie formal dargestellt werden kann (vgl. TACK 1980).

Meß- und maßtheoretische Erwägungen müssen daher mehr Gewicht erhalten als ihnen momentan zugewiesen wird. Eine noch zu entwickelnde "Diagnostik des Problemlösens", die neben eine "Diagnostik der Intel-

ligenz" zu stellen wäre, hat genau wie die letztgenannte die Aufgabe, ihre Gültigkeit für verschiedene Populationen nachzuweisen (Problem der differentiellen Validität). Inwiefern hierbei die traditionellen Konzepte von Reliabilität und Validität überhaupt verwendbar sind, muß ebenfalls geklärt werden. Am Beispiel "TAILORSHOP" (Teil II dieser Arbeit) wird die Problematik reliabler und valider Indikatoren der Lösungsgüte weiter vertieft (vgl. auch Kap. 3.8).

3.7 Die These von der suboptimalen Versuchsplanung und -auswertung

Will man im Bereich komplexen Problemlösens experimentell arbeiten, sollten versuchsplanerische und auswertungstechnische Überlegungen verstärkt Gewicht erhalten. Auf postexperimentelle Aufteilungsverfahren kann oftmals verzichtet werden, das Mitteilen auch nichtsignifikanter Ergebnisse sollte -bei kontrolliertem Beta-Fehler- selbstverständlich sein.

Vorliegende (quasi-)experimentelle Studien zum komplexen Problemlösen (FUNKE 1983b; FUNKE & HUSSY 1984; HESSE 1982; HESSE, SPIES & LÜER 1983; PUTZ-OSTERLOH 1981, 1983b) gehen kaum über ein klassisches 2x2-Design hinaus. Es gibt sogar Studien mit einem einzigen dreistufigen Faktor (z.B. "Gruppenzugehörigkeit" bei PUTZ-OSTERLOH 1983b). Ob diese Designs der Komplexität des Untersuchungsgegenstandes und der theoretisch postulierten vielfältigen Einflußgrößen gerecht werden, scheint fraglich. Oftmals ist einer der Faktoren (z.B. "Testintelligenz") quantitativ und wird künstlich dichotomiert, ein Vorgehen, das auch auf die abhängige Variable "Lösungsgüte" angewandt wird, um die Stichprobe post-hoc in Gruppen "guter" oder "schlechter" Problemlöser einzuteilen. Einmal abgesehen von der Problematik, ein aus der Sicht des Systems und seiner Eigenschaften (vgl. Kap. 3.4) begründetes Maß der Lösungsgüte definieren zu können, erscheint es in derartigen Fällen gewinnbringender, die metrische Information regressionsanalytisch auszuschöpfen (vgl. Kap. 8.2.2 in COHEN & COHEN 1975: "When not to use the 2x2-design"), ein Argument übrigens, das auch für eine eigene Arbeit (FUNKE 1983b) Gültigkeit besitzt.

Die Anzahl der abhängigen Variablen scheint in den meisten Arbeiten schier grenzenlos, so daß sich mancherorts die Berichterstattung nach dem (öfter wechselnden) Signifikanzkriterium richtet (so z.B. PUTZ-OSTERLOH 1983b, p.112). Dennoch finden multivariate Analyseverfahren nur spärlich Verwendung, obwohl eine konsequente Verfolgung des systemtheoretischen Grundansatzes nicht nur die Kontrolle vieler UVn, sondern auch vieler AVn erfordert; auch auf seiten der abhängigen Variablen interessieren nicht isolierte Auswirkungen, sondern Auswirkungsgefüge.

Ein weiteres Argument bezüglich Versuchsplanung betrifft die sozialpsychologische Seite des Experimentierens (vgl. BREDEKAMP 1980, p.41f): Um Fehler zu Lasten des Forschers zu reduzieren, sollte der Versuchsleitereffekt in künftigen Studien dadurch ausgeschlossen bzw. verringert werden, daß dem Versuchsleiter keine derart intensive Ver-

der Bestimmbarkeit eines oder mehrerer als objektiv optimal zu bezeichnender Lösungswege konstruiert wurden.

Angesichts der geschilderten Schwierigkeiten bei der Festlegung eines Erfolgsmaßes wird in mehreren Studien (z.B. DÖRNER et al. 1983; PUTZ-OSTERLOH 1981) der Weg über Beurteilungen von grafisch aufbereiteten Verlaufsdaten beschriftet. Folgende Überlegung steht dabei im Hintergrund: Kommen verschiedene Beurteiler "in etwa zu gleichen Urteilen über die Leistungen der Vpn aufgrund der Graphiken ...", so wäre dies ein Beleg für einen gleichartigen Bewertungsmaßstab der Leistungsgüte" (DÖRNER et al. 1983, p.159) durch die verschiedenen Schätzer. Korreliert darüberhinaus die durchschnittliche Beurteilung durch die Rater hoch mit dem gewählten Erfolgsindikator, wird dies als indirekter Beleg für die Reliabilität bzw. Validität des Erfolgsindikators gewertet. Zunächst wird an zwei Beispielen gezeigt, wie dieses Vorgehen in der Praxis aussieht. Dann wird eine kleine eigene Studie vorgelegt, die diesen zweifelhaften Weg der Reliabilitäts- und Validitätsbestimmung als höchst fragwürdig erscheinen läßt.

Beispiel 1: PUTZ-OSTERLOH (1981, p.88) läßt sieben Beurteiler (3 Versuchsleiter und 4 "naive" Personen) den grafisch präsentierten Verlauf von 12 wichtigen Systemvariablen über die 24 Bearbeitungsmonate auf einer fünfstufigen Skala einschätzen. Neben diesen subjektiven Erfolgskriterien wird von der Autorin ein objektives Gütemaß herangezogen, nämlich die Anzahl der Simulationsmonate, in denen das Flüssigkapital des Probanden einen Anstieg verzeichnet. Dieses Trendmaß schwankt zwischen Null (=kein Monat mit Aufwärtstrend im Flüssigkapital) und der Anzahl der simulierten Monate minus Eins (=alle Monate Aufwärtstrend). Die mittlere Einschätzung des Lösungserfolgs korreliert mit ihrem Trendmaß zu 0.83, die Beurteilerübereinstimmung nach HORST (1949) beträgt $r_{cc} = 0.94$. Daraus wird auf die "Gültigkeit des Kriteriums zur Beschreibung des Problemlöseerfolgs" (PUTZ-OSTERLOH 1981, p.88) geschlossen.

Beispiel 2: Bei DÖRNER et al. (1983) sind es fünf Versuchsleiter und vier bis fünf naive Beurteiler, die jeweils 10 bis 12 anonyme Grafiken, bestehend aus dem Verlauf von 17 kritischen Variablen, auf einer (wohl mehrstufigen) Skala einschätzen. Die Interkorrelationen innerhalb und zwischen den beiden Gruppen fallen "ausnahmslos hoch signifikant" aus (mit gewissen Abstrichen bei den naiven Ratern), ebenfalls signifikante Korrelationen bestehen zwischen den "Grafik"-Urteilen und den zwei "objektiven" Systemindikatoren "Fabrikkapital" sowie "Einnahmen-/Ausgaben-Quotient" der Stadt.

"Tatsächlich finden sich ... so hohe Übereinstimmungen zwischen den verschiedenen Gütekriterien, sei es zwischen subjektiven oder objektiven, zwischen wissenschaftlichen oder 'naiven', oder sei es zwischen globalen und mehr detaillierten, daß sie zu einem gemeinsamen Kriterium, dem Generalgütekriterium zusammengefaßt werden können." (DÖRNER et al. 1983, p.163).

Daraus wird gefolgert, daß ein "brauchbares Maß für die individuelle

Problemlöseleistung" vorliegt, das sich u.a. auch für eine Extremgruppenbildung eigne (p.164). Abb. 3.7 veranschaulicht noch einmal das in beiden Beispielen zugrundegelegte Prinzip der Gütebestimmung.

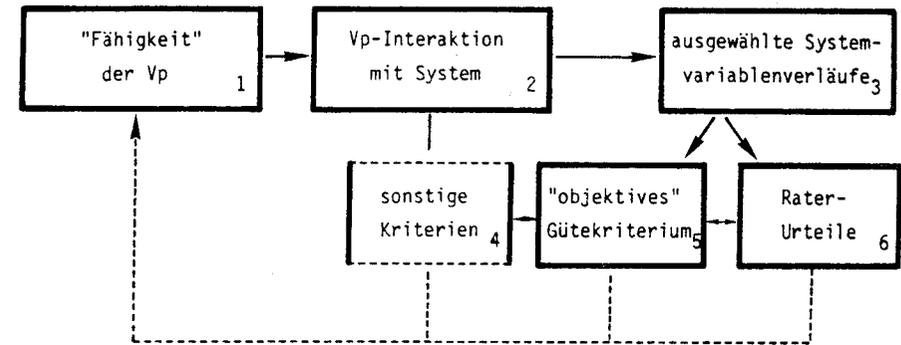


Abb. 3.7: Veranschaulichung des typischen Vorgehens bei der Bewertung von Vpn-Verhalten im Umgang mit komplexen Systemen.

Die Fähigkeit des Probanden (1) schlägt sich in der Interaktion mit dem System (2) nieder. Daraus werden Systemvariablenverläufe ausgewählt (3) und einerseits nach "objektiven" Kriterien (5) ausgewertet, andererseits nach Rater-Urteilen (6) bewertet. Die hohe Korrelation von (5) mit (6) sowie von (5) mit sonstigen aus (2) abgeleiteten Variablen (4) wird als Validitätshinweis interpretiert, womit der Bezug zur latenten Variable (1) hergestellt zu sein scheint.

Gegenbeispiel: Um zu veranschaulichen, daß hohe Rater-Reliabilitäten und hohe Korrelationen mit einem "objektiven" Außenkriterium keinen Rückschluß auf die Validität der Maße, noch nicht einmal auf deren Reliabilität erlauben, wird eine kleine Untersuchung vorgestellt.

Zu diesem Zweck wurde künstlich ein Datensatz mit bekannter Struktur generiert. Insgesamt 48 mal wurden Sequenzen der Länge 10 von einem Zufallszahlengenerator gebildet nach folgender Regel:

1. Ausgangswert jeder Sequenz ist der Wert $x=10$;
2. verändere den letzten Wert x per Zufall im Intervall $\pm 2.5 * x$ (gleichverteilt) und nimm den ganzzahligen Anteil des resultierenden Wertes;
3. wiederhole 2., bis 10 Daten vorliegen.

Das entsprechende BASIC-Programm für den COMMODORE-Kleinrechner zeigt Abb. 3.8.

```

1 OPEN1,3:X=RND(-1)
5 FOR Z=1 TO 50:PRINT#1,"*** FALL";Z
10 X=10:T=5:SA=0:PRINT#1,0,X,SA
20 FOR L=1 TO 10:XA=X
30 X=X+RND(1)*T-T/2:X=INT(X)
35 IF X>XA THEN SA=SA+1
40 PRINT#1,L,X,SA
50 NEXT
55 PRINT#1
60 NEXT

```

Abb. 3.8: Liste des BASIC-Programms zum Erzeugen "zufälliger" Trenddaten.

Neben der erzeugten Sequenz selbst (=X) wird auch angegeben, in wievielen Fällen ein "Aufwärtstrend" (=SA) vorliegt. Da das Programm mit dem COMMODEORE-eigenen seed-Wert des Pseudozufallszahlengenerators arbeitet, erübrigt sich ein Abdruck der Daten, da sie jederzeit wieder rekonstruiert werden können. Auf diese Art und Weise erzeugt, liegt ein Datensatz vor, der dem Leser möglicherweise bekannt vorkommt: Nehmen wir an, bei der erzeugten Sequenz handele es sich um Kapitalveränderungen eines simulierten Systems, dann können wir nunmehr fragen, ob Rater die hier per Zufall erzeugten "Kurven" treffsicher (d.h. übereinstimmend) beurteilen und wie die Korrelation zum "objektiven" Gütekriterium Trendmaß (die Zahl der Einheiten mit "Aufwärtstrend") ausfällt.

Vorab noch einige Angaben zu den statistischen Eigenschaften der 48 erzeugten Sequenzen. Wie auf der Basis des vorgestellten Erzeugungsalgorithmus nicht anders zu erwarten, beträgt die odd-even-Korrelation der Rohwerte 0.91 (jeder Meßwert wird ja vom vorausgehenden nur um maximal 25% verändert), während die odd-even-Korrelation des Trendmaßes -0.09 beträgt (der Trend aufeinanderfolgender Meßwerte ist ja vollkommen unabhängig voneinander).

Die 48 Kurven wurden grafisch veranschaulicht und zusammen mit einer kurzen Instruktion insgesamt sechs Ratern vorgelegt. Davon waren vier "Experten" (diese hatten die "Schneiderwerkstatt" mindestens einmal selbst gespielt) und zwei "Naive". In der Instruktion hieß es, die Kapitalkurven von 48 Vpn seien hinsichtlich ihrer Güte auf einer fünfstufigen Skala von "++" bis "--" zu beurteilen. Die Auswertung der erhaltenen Daten erfolgte mittels des Übereinstimmungskoeffizienten r_{cc} von HORST (1949; zit. nach SIXTL 1982², p.149f.) ganz analog zu der oben beschriebenen Vorgehensweise von PUTZ-OSTERLOH (1981).

Der Leser soll nicht länger auf die Folter gespannt werden: das r_{cc} der sechs Beurteiler beträgt 0.941, die Korrelation der durchschnittlichen Beurteilung mit dem Trendmaß beträgt dagegen -0.685 (die Skalierung der Raterurteile entsprach Schulnoten: "++" = 1, "--" = 5; ein hoher Trendwert deutet starken Anstieg und damit "Güte" an)! M.E. kann aus den erhaltenen hohen Werten (sie entsprechen den in den genannten

Studien vorgelegten Zahlen übrigens recht gut) kein Schluß auf die Reliabilität des Systemverhaltens gezogen werden.

Da es sich hierbei um einen Validierungs"versuch" handelte, soll der möglicherweise vorliegende Fehlschluß nicht zu hoch gewichtet werden. Er macht im übrigen eines deutlich: Wenn die Korrelation zwischen Rater-Urteil und Trendmaß wie berichtet hoch ausfällt, ist anzunehmen, daß das Trendmaß das Beurteilungskriterium für die Rater gewesen sein muß. Wenn die Verläufe von 12 (!) oder gar 17 (!!) Systemvariablen dem Rater vorgelegt werden, wird ein derartiger "information overload" auftreten, der nur durch massive Reduktionsmechanismen bewältigt werden dürfte. Eine naheliegende Strategie wäre, sich auf die (vermutete) zentrale Variable "Kapital" zu konzentrieren und aus deren Verlauf Rückschlüsse zu ziehen. Nur so ist ja wohl auch die Tatsache substantieller Korrelationen zwischen naivem und Expertenrating zu erklären: die Rater sind -verständlicherweise- der vermuteten Bedeutung einzelner Variablen und deren Auf und Ab erlegen. Im Normalfall würde man ja gerade eine Diskrepanz zwischen beiden Beurteilergruppen erwarten, da unterschiedliche Bewertungskriterien (nämlich differenziertere auf Expertenseite) unterstellt werden müssen. Die fehlende Diskrepanz als Hinweis auf ein "von möglichen Interaktionseffekten zwischen Vpn und V1 unbeeinflusstes Gütekriterium" (DÖRNER et al. 1983, p.162) anzusehen, halten wir daher für nicht schlüssig, selbst wenn sich bezüglich der Urteilsbegründungen von Naiven und Experten Unterschiede ergeben (vgl. hierzu die Unterstellung von PUTZ-OSTERLOH 1983a, ich hätte das Verfahren der Beurteilung in meiner Arbeit von 1983b verkürzt und damit falsch dargestellt).

Alle diese -hier ausführlich geschilderten- Probleme lassen sich reduzieren, wenn Probanden mit nach mathematischen Gesichtspunkten optimierbaren Systemen umgehen: In diesen Fällen kann man auf den Umweg bzw. Irrweg über die Rater-Urteile verzichten. Daß dennoch -auch in nicht-optimierbaren Systemen- unterschiedlich gute Kriterien der Systemmanipulation aus dem System selbst herleitbar sind, wird am Beispiel des "TAILORSHOP" in Teil II aufgezeigt werden.

3.9 Nachbemerkung zu den Thesen

Nach soviel Kritik erscheint eine selbstkritische Nachbemerkung angebracht. Vom Lehnstuhl des Kritikers aus die Schwächen und Lücken in den Arbeiten Anderer zu entdecken, mag ein gutes Stück weit nützlich und für die Kritisierten vielleicht auch hilfreich sein. Zentrale Frage jedoch bleibt: wie kann ein konstruktiver Weg aus den aufgeführten Problemen der Problemlöseforschung aussehen?

Diese Frage stellen heißt, dem Kritiker eine ziemliche Bürde aufzulasten: An vielen Stellen habe ich teilweise nur in Form der Angabe weiterführender Literatur versucht, erste Hinweise zu geben. Am zentralen Punkt, nämlich der Theoriefrage (vgl. Kap. 3.1), klafft einstweilen eine Lücke; in Teil III dieser Arbeit unternehme ich erste Ansätze mit dem Modell der Identifikation von Kausalstrukturen (vgl.

Kap. 5), die aber zum gegenwärtigen Zeitpunkt noch unvollständig entwickelt sind. Hier bedarf es größerer Anstrengungen, um tatsächlich falsifizierbare Prognosen über das Problemlöseverhalten machen zu können.

Es gibt Wissenschaftler -so etwa DÖRNER-, die von der (keineswegs unplausiblen) Annahme ausgehen, der Mensch sei ein "chaotisches System" im Sinne der Katastrophentheorie, d.h. bereits geringfügige Änderungen in den Bedingungsvariablen riefen nicht nur quantitative, sondern qualitative Änderungen hervor. Bei den momentan gegebenen Unschärfen der Meßinstrumente wird damit eine Prognose von konkretem Problemlöseverhalten nur stark eingeschränkt möglich sein. Die neueren Entwicklungen von DÖRNERs allgemeinem Modell der Handlungsregulation gehen daher konsequent in Richtung Computersimulation: Streng formalisierte Modellbausteine werden in ihrem Zusammenspiel auf dem Rechner beobachtet, Vergleiche mit Protokollen menschlicher Probanden zur Kontrolle herangezogen. So interessant diese Studien sind - den Verdacht, ein ausgeklügeltes Glasperlenspiel vor sich zu haben, wird man angesichts der eingestandenen Nichtprüfbarkeit des Modells nicht los.

Ein Wort zur Methodik scheint ebenfalls als Nachbemerkung notwendig. Angesichts fehlender ausformulierter Theorien kann der Einsatz inferenzstatistischer Prüfverfahren vielfach nicht oder nur schlecht begründet werden. Dennoch möchte man sich natürlich gegen "zufällige" Ereignisse absichern. Hier wäre daran zu denken, zunächst einmal die Replizierbarkeit der erhaltenen Daten zu sichern (vgl. These 6 und die Ergebnisse von STROHSCHNEIDER 1985) und die Präzision der Daten (etwa mittels Konfidenzintervall) abzuschätzen. An die Stelle von Inferenzstatistik könnte die Entscheidungstheorie treten, um Auskunft über wahrscheinliche oder unwahrscheinliche Ereignisse zu geben. Eines steht jedoch fest: Daten "an sich" sind vollkommen nichtssagend, erst im Lichte einer -wie auch immer gearteten- Hypothese liefern sie Erkenntnisgewinn!

Teil II: Das Beispiel "TAILORSHOP"

4 Das Simulationsprogramm "TAILORSHOP"

Nachdem im ersten Teil eine kritische Bestandsaufnahme der gegenwärtigen Forschung zum komplexen Problemlösen versucht wurde, soll im zweiten Teil anhand des Beispiels "TAILORSHOP" eine Vertiefung ausgewählter Kritikpunkte vorgenommen werden. Zunächst einmal erfolgt in Kap. 4.1 eine Beschreibung des Computersimulationsprogramms, die bestimmte Eigenheiten des Systems verdeutlicht. Eine "Simulation der Simulation" (Kap. 4.2) liefert interessante Hinweise über verschiedene Maße der Lösungsgüte. Dann folgt in Kap. 4.3 die Darstellung bisheriger Arbeiten zum "TAILORSHOP", d.h. die Arbeiten von PUTZ-OSTERLOH (1981, 1983b) und PUTZ-OSTERLOH & LÜER (1981) sowie die von diesen Arbeiten angeregte Replikationsstudie (FUNKE 1983b) werden noch einmal kurz skizziert. Auch die Kontroverse um die Interpretation der dort ermittelten Befunde (vgl. PUTZ-OSTERLOH 1983a, FUNKE 1984) kommt zur Sprache (Kap.4.4).

4.1 Beschreibung des Computerprogramms

Das Computerprogramm "TAILORSHOP" von Dietrich DÖRNER (damals Gießen) in einer "Version für den TI-59 mit Drucker PC 100" lag in der von Norbert STREITZ (Aachen) modifizierten und kommentierten Fassung vom März 1979 vor. STREITZ hat das Programm in einigen Punkten verändert: Neben der '50-Hemden'-Maschine gibt es einen '100-Hemden'-Maschinen-Typ. Zwischen beiden Typen kann umgerüstet werden. Weiterhin wurde die Möglichkeit geschaffen, Beiträge in einen Sozialfond einzuzahlen. Diese Variable wirkt sich auf die Zufriedenheit der Arbeiter aus, die eine interne Variable für die Berechnung der Produktion darstellt. Außerdem wurden eine Reihe von Parametern und Prozeduren verändert. Ein Problem stellte die begrenzte Anzahl verfügbarer Speicher auf dem TI-59 dar, was dazu führte, daß in dem System eine Reihe von vereinfachenden Annahmen gemacht wurden, vor allem bei dem Zusammenhang zwischen Maschinenkapazität und Reparaturinvestitionen beim Neukauf von Maschinen.

Das mir vorliegende Programmlisting von 559 Programmzeilen wurde zunächst auf Magnetkarte übertragen. Parallel dazu lief die Entschlüsselung der TI-59-spezifischen Maschinsprache und deren Umsetzung in die höhere Programmiersprache BASIC, über die der vorhandene Kleincomputer (Typ CBM-3032) verfügte.

Zur Prüfung der Äquivalenz des umgeschriebenen Programms wurde folgendes Vorgehen gewählt: In beiden Programmversionen, also der Taschen- und der Tisch-Rechnerversion, legte ich die Zufallsgrößen, die auf verschiedene Funktionen aufgesetzt waren (siehe unten), als Konstanten fest und fixierte folgende Eingriffsabfolge:

- (1) In den ersten zwei Simulationstakten keine Eingriffe.
- (2) Im dritten Monat: 1000 Einheiten Rohmaterial, 4000.- Werbung, 1300.- Lohn, 100.- Sozialkosten, 45.- Hemdenpreis, 10 Arbeiter an 50er-Maschinen.
- (3) Im vierten Monat: Lage = 1 (City), 600 Einheiten Rohmaterial.
- (4) Im fünften Monat: 400 Einheiten Rohmaterial, 3000.- Reparaturkosten.
- (5) Im sechsten Monat: 200 Einheiten Rohmaterial, Verkauf von 2 50er-Maschinen, Entlassung von 2 Arbeitern an 50er-Maschinen.

Unter den gegebenen Startwerten (siehe unten, Tab. 4.1) zeigten sich hinsichtlich aller Variablen in den TI-59-Speichern 12 bis 39 mit meinem Programm übereinstimmende Daten für die sechs simulierten Zeitpunkte. Dies wurde als ausreichender Test auf Äquivalenz des neuen BASIC-Programms mit dem Quell-Programm für TI-59 angesehen.

Da von den Autoren des "TAILORSHOP" keine dokumentierten Angaben zur verwendeten Zufallszahlenreihe vorliegen, verzichtete ich auf das Auslesen dieser Sequenz aus dem Standard-Software-Modul des TI-59 und verwende in meinem BASIC-Programm eine eigene, für jeden Probanden gleiche Pseudo-Zufallszahlen-Sequenz mit dem "seed" RND(-1).

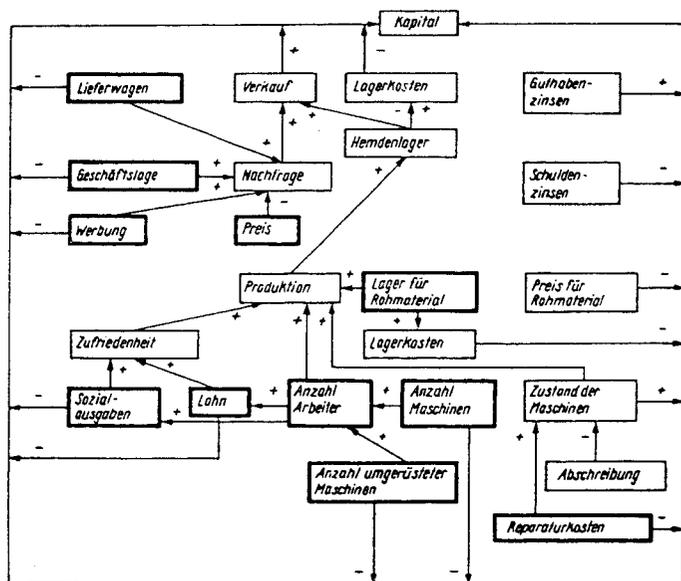


Abb. 4.1: Die Variablen des "TAILORSHOP" und ihre Verknüpfungen. Die dick umrandeten Variablen sind direkt beeinflussbar (N = 11); + = positive Korrelation, - = negative Korrelation (aus PUTZ-OSTERLOH 1981b, p. 83).

Nun zu den eigentlichen Details des Systems. In der vorliegenden Literatur zum "TAILORSHOP" (PUTZ-OSTERLOH 1981, 1983b; PUTZ-OSTERLOH &

LUER 1981) finden sich eher spärlich zu nennende und weitgehend textgleiche Beschreibungen. Neben der Darstellung der Variablenvernetzung (vgl. Abb. 4.1) durch Angabe von groben Wirkungscharakteristiken (Maßnahme wirkt positiv oder negativ auf andere Variable) wird gezeigt, welche Größen vom Probanden beeinflussbar sind und welche nicht.

Um die Autoren zu Wort kommen zu lassen, hier deren Darstellung des "TAILORSHOP":

"Das System besteht aus 24 netzwerkartig miteinander verknüpften Variablen, von denen 11 durch Maßnahmen der Pbn direkt beeinflussbar sind. Das Kapital stellt die 'Kernvariable' des Systems dar, da es mit den meisten, nämlich 15 Variablen, in direkter Verbindung steht. - Einige Variablen unterliegen in ihrem Zustand Zufallsschwankungen und werden somit nicht allein durch den Zustand der Beeinflussungsvariablen bestimmt. Der Effekt von Veränderungsmaßnahmen wirkt sich nicht immer sofort, sondern z.T. erst mit einer Zeitverzögerung auf den Zustand direkt benachbarter, d.h. verknüpfter Variablen aus. - Das System ist so programmiert, daß die Werkstatt ohne Eingriffe in kurzer Zeit Konkurs anmelden müßte. Denn Rohmaterial muß eingekauft werden, Reparaturkosten sind z.B. zu erhöhen usw., wenn der Betrieb bei bestehender Produktionskapazität Gewinn erwirtschaften soll. - Aufgabe des Problemlösers ist es, die Schneiderwerkstatt über einen vorgegebenen Zeitraum von simulierten Monaten hinweg mit Hilfe vorgegebener Maßnahmen so zu leiten, daß der Betrieb maximalen Gewinn auf die Dauer erwirtschaftet." (PUTZ-OSTERLOH & LUER 1981, p.313-315).

In Tab. 4.1 findet der Leser eine Liste wichtiger beteiligter Variablen des "TAILORSHOP" sowie ihrer Kürzel, in Tab. 4.2 die Gleichungen, aus denen die Vernetzung der Variablen hervorgeht.

Tab. 4.1: Liste der Variablen im Programm "TAILORSHOP".

Kürzel	Bezeichnung	Anfangswert
ZA	aktuelle Zufriedenheit	0.981
PA	aktuelle Produktion	403.932
PM	mögliche Produktion	403.932
KL	Hemden im Lager	80.716
RL	Rohmaterial im Lager	16.068
NA	Nachfrage aktuell	766.636
VH	verkaufte Hemden	407.216
PH	Preis pro Hemd	52.000
RS	Reparatur- und Servicekosten	1200.000
MA	aktuelle Maschinenkapazität	47.040
WE	Werbe-Kosten	2800.000
LO	Lohnkosten pro Arbeiter	1080.000
SM	Sozialausgaben pro Arbeiter	50.000
N1	Anzahl Arbeiter an 50er	8
N2	Anzahl Arbeiter an 100er	0
A1	Anzahl 50er-Maschinen	10
A2	Anzahl 100er-Maschinen	0
LW	Anzahl Lieferwagen	1
GL	Geschäftslage (0 = Vorort, 0.5 = Cityrand, 1 = City)	0.5
KA	Kapital	15774.659
RP	Rohmaterial-Preis	3.994
ND	Anzahl registrierter Monate	0

Tab. 4.2: Vernetzung der Variablen im "TAILORSHOP"*

Nr. der Gleichung	Variable	Beziehung
(1)	ZA	$ZA = \text{Min}(ZM, (0.5 + ((LO - 850)/550) + SM/800))$
(2)	PM	$PM = (\text{Min}(N1, A1) \cdot (MA + \xi \cdot 4 - 2) + \text{Min}(N2, A2) \cdot (MA \cdot 2 + \xi \cdot 6 - 3)) \cdot ZA$
(3)	PA	$PA = \text{Min}(PM, RL)$
(4)	HL	$HL = HL + PA - VH$
(5)	RL	$RL = RL - PA$
(6a)	NA ₁	$NA_1 = NA_2/2 + 280 \cdot 1.25 \cdot e^{-(PH^2/4250)}$
(6b)	NA ₂	$NA_2 = \text{Min}(WE/5, NM) + LW \cdot 100$ $NA_2 = NA_2 + NA_2 \cdot g + \xi \cdot 100 - 50$ wobei $g = 0$ für GL = Vorort $= 0.1$ für GL = Cityrand $= 0.2$ für GL = City
(7)	VH	$VH = \text{Min}(HL, NA_1)$
(8)	RP	$RP = 2 + \xi \cdot 6.5$
(9)	MA	$MA = \text{Min}((MA - 0.1 \cdot MA + (RS/(A1+A2)) \cdot 0.017), MM)$
(10a)	KA	$KA = KA - SM \cdot (N1+N2) - PA \cdot 1 - RL \cdot 5$ $- HL \cdot 1 + VH \cdot PH - WE - LW \cdot 500$ $- GL \cdot 2000 - RS - (N1+N2) \cdot LG$
(10b)	KA	$KA = KA + KA \cdot z$ wobei $z = 0.0025$, wenn $KA \geq 0$ $= 0.0066$, wenn $KA < 0$

* Das Zeichen ξ steht für eine Zufallszahl im Bereich zwischen Null und Eins, erzeugt von einem rechnerinternen Pseudozufallszahlengenerator.

Aus Tab. 4.2 geht hervor, daß in den Gleichungen (2), (6b) und (8) die Werte für die mögliche Produktion PM, Nachfrage für den nächsten Monat NA₂ sowie Preis des Rohmaterials RP innerhalb gewisser Grenzen "zufällig" schwanken können. Nur der Preis des Rohmaterials ist dem Probanden bekannt, er kann sich entsprechend darauf einstellen. Auf die beiden anderen Variablen dagegen kann er nicht reagieren, da sie "verdeckt" sind.

Zur Vervollständigung der Angaben über das Simulationsprogramm sind in Tab. 4.3 und 4.4 eine Liste der Konstanten sowie wichtige Nebenbedingungen und Zusatzinformationen enthalten.

Tab. 4.3: Liste der Konstanten im Programm "TAILORSHOP".

Kürzel	Bezeichnung	Wert
MM	maximale Maschinenkapazität	50
NM	maximale Nachfrage	900
ZM	maximale Zufriedenheit	1.7
GZ	Guthabenzinsen monatlich	0.0025
SZ	Schuldzinsen monatlich	0.0066

Tab. 4.4: Nebenbedingungen und Zusatzinformationen für den Betrieb des "TAILORSHOP".

Nr.	Bezeichnung	Wert
(1)	minimale Lohnkosten	850
(2)	a) Lagerkosten pro Hemd Rohmaterial	0.50
	b) Lagerkosten pro fertiges Hemd	1.00
(3)	Mietkosten	
	a) Vorstadt	500
	b) Cityrand	1000
	c) City	2000
(4)	Lieferwagenkosten	
	a) Anschaffungspreis pro LKW	10000
	b) Laufende Kosten pro LKW	500
	c) Verkaufserlös: 8000 - (Anzahl Monate * 100)	
(5)	Maschinenkosten	
	a) Anschaffung einer 50-Hemden-Maschine	10000
	b) Anschaffung einer 100-Hemden-Maschine	20000
	(nur möglich, wenn MA \geq 35)	
	c) Verkaufserlös 50er: MA/MM * 8000	
	d) Verkaufserlös 100er: MA/MM * 16000	

Es ist klar, daß -sofern stochastische Modelle verwendet werden- die Zufallszahlen für jeden Probanden in derselben Abfolge wirken müssen, um keine unerwünschten Effekte zu erzielen. In diesem Sinn handelt es sich auch nicht mehr um Zufallsschwankungen, sondern um zeitabhängige Konstanten des Systems. Überlegungen dazu sind auch deswegen wichtig, weil sie die Vergleichbarkeit verschiedener Studien mit ein- und demselben System betreffen. Es ist nicht geklärt, in welcher Weise die Ergebnisse mit veränderten Zufallszahlen variieren. Deswegen scheint die Forderung nach Publikation der verwendeten Zufallszahlenreihe und des Verhältnisses von Signal und Rauschen sinnvoll und notwendig. Lediglich eine Arbeit (SCHÖTTKE & GEDIGA 1982, p.15-18) beschäftigt sich mit dem Verhältnis von Signal und Rauschen bei dem von der Osnabrücker Arbeitsgruppe eingesetzten Simulationsmodell "Hamurabi, König von Summaria" (vgl. auch GEDIGA, SCHÖTTKE & TÜCKE 1982; GEDIGA 1983; GEDIGA, SCHÖTTKE & TÜCKE-BRESSLER 1984). Dort wurde für die "empfindlichste" Systemvariable ("Population") der "schlechteste mögliche Fall" des Verhältnisses von Signal und Rauschen über alle t Zeittakte untersucht, um sicherzustellen, daß der gewählte Fehlerwert die Variable nicht "verrauscht". Eine derartige Abschätzung sollte für alle nicht-deterministischen Systeme gefordert werden.

Wie sehr die Zufallsgrößen den Güteindikator "Anzahl Monate mit Aufwärtstrend im Flüssigkapital" beeinflussen können, sei an einer kleinen Simulation verdeutlicht. Um zu prüfen, welche Effekte durch die "zufälligen" Bedingungen eintreten können, wurde eine fixierte Interventionsstrategie über zwölf Monate unter zwei extremen Bedingungen durchgeführt: Unter der Bedingung "günstig" (G) wurde der Wert für

PM und NA₂ auf Eins und für RP auf Null festgelegt. Dadurch liegen die Werte für Produktion und Nachfrage maximal hoch, der Einkaufspreis für Rohmaterial maximal niedrig. Unter der Bedingung "ungünstig" (U) lagen die gegenteiligen Verhältnisse vor, also niedrige Produktion und Nachfrage bei teurem Einkaufspreis. Folgende Interventionsstrategie fand bei diesem Test Verwendung: In allen Monaten wird Rohmaterial für 500 Hemden eingekauft, zusätzlich wird im ersten Monat der Hemdenpreis auf 45 gesenkt, Werbung auf 3500 erhöht, zwei 50er-Maschinen verkauft, Reparaturkosten auf 3000 und Lohnkosten auf 1300 erhöht. Mit dieser einfachen Intervention wird der Laden im ersten Monat stabilisiert und anschließend nur durch Ankauf von Rohmaterial am Laufen gehalten. Abb. 4.2 zeigt den Verlauf der "zentralen" Variable Kapital unter den genannten Voraussetzungen.

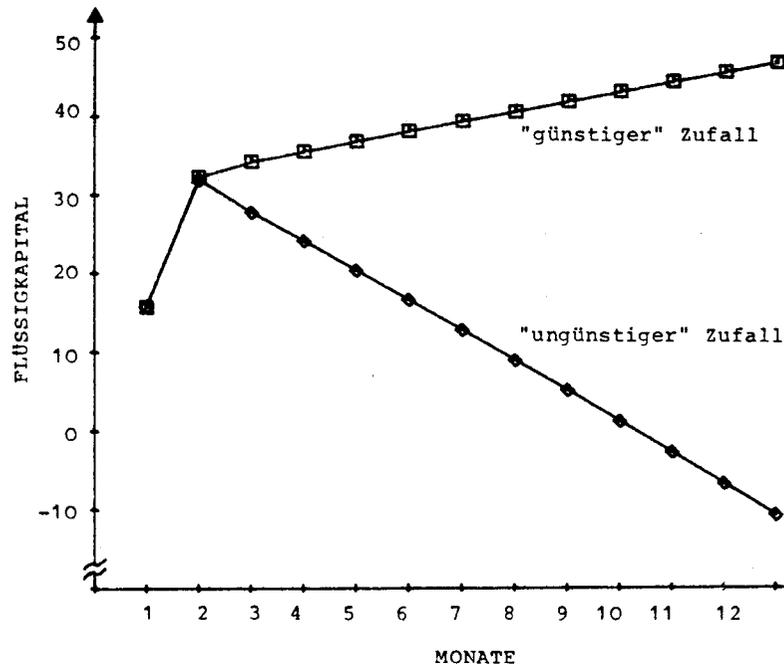


Abb. 4.2: Verlauf des "Erfolgs"-Maßstabs Flüssigkapital über zwölf Monate hinweg bei konstantem Vpn-Verhalten unter günstigen bzw. ungünstigen Zufallsbedingungen.

Abb. 4.2 verdeutlicht den Unterschied zwischen G und U nach zwölf Monaten (57547 Einheiten), den der Proband nicht zu verantworten hat, da er ihn nicht beeinflussen kann. Bezüglich des angeblichen Gütemaßstabes "Aufwärtstrend in Flüssigkapital" führt dies zu einem katastrophalen Resultat: Unter Bedingung G erzielt TRENDPO (so nennen wir den von PUTZ-OSTERLOH vorgeschlagenen Trendindikator) einen (maximalen) Wert von 12, unter U dagegen einen Wert von 1! Der von FUNKE (1983b,

p.292) vorgeschlagene Trendindikator TRENDFU -dort wird die Anzahl der Simulationsmonate ausgezählt, in denen ein Anstieg nicht des Barkapitals, sondern des Gesamtkapitals, also einer Art monatlicher Bilanz, vorlag- reagiert allerdings in gleicher Weise wie TRENDPO, da bei konstanten Investitionen (hier: konstant keine) natürlich das Flüssigkapital, das Bestandteil von TRENDFU ist, allein zur Varianz beiträgt; läßt man dagegen variable Investitionstätigkeiten zu -und dies tritt im Experiment natürlich auf-, spricht nur TRENDPO auf dieses Verhalten an, während TRENDFU davon nicht beeinflusst wird. In TRENDFU wird der durch Investitionen entstehende Verlust an Flüssigkapital durch Bilanzierung des Maschinen- bzw. Materialwertes wieder ausgeglichen; Investitionen führen bei TRENDFU also nicht zu einem Abfall des Güteindicators, wie dies bei TRENDPO der Fall ist. Bei gleicher Interventionsstrategie kommen die Trendindikatoren somit zu gegenläufigen Resultaten, die lediglich zu Lasten der zufällig schwankenden Werte gehen. Nun kann man argumentieren, in den tatsächlich eingesetzten Programmen seien solche Extrem-Bedingungen nicht möglich; dies stimmt natürlich, läßt aber zumindest die Notwendigkeit einer quantitativen Abschätzung solcher Random-Effekte nicht überflüssig erscheinen, die in einem konkreten Fall realisiert wurden.

Ein paar Bemerkungen zur Realitätsnähe des Simulationsprogramms sind notwendig. Bei der Variablen "Maschinenschäden" (MA; vgl. (9) in Tab. 4.2) wurde eine dynamische "Verschlechterung" vorgesehen, die durch einen der Maschinenzahl entsprechenden Betrag für Reparatur/Service (RS) aufgefangen werden kann (pro Maschine wird ein Betrag von 297.- benötigt, wobei übrigens beide Maschinentypen gleichgewichtet werden). Allerdings -und dies ist das Merkwürdige daran- spielt das Alter der Maschinen keine Rolle; werden also zu einer Zahl "schlechter" Maschinen neue hinzugekauft, nehmen diese sofort denselben schlechten Zustand an. Probanden, die durch einen kompletten Austausch alter Maschinen gegen neue die Variable MA beeinflussen wollen, machen nur finanzielle Verluste (Verkaufswert der Maschinen sinkt mit zunehmenden Schäden), erzielen aber keinen Effekt bei MA. Dies führt auch zu der Beschränkung, daß bei mehr als 30 % Maschinenschäden (definiert als $100 - (MA/MM) * 100$) keine neuen 100-Hemden-Maschinen erworben werden können (wohl aber 50-Hemden-Maschinen; in der Variante von PUTZ-OSTERLOH können unter der genannten Bedingung keinerlei Maschinen gekauft werden).

Eine weitere wichtige Merkwürdigkeit betrifft die Variable "Lieferwagen" (LW) und deren Einfluß auf die Nachfrage (NA₂; vgl. (6b) in Tab. 4.2). Unabhängig von sonstigen Bedingungen erhöht jeder LW die NA₂ um einen Betrag von 100. Da nun insbesondere das Erzeugen von Nachfrage nach den produzierten Hemden für den Manager von Bedeutung ist, kann er auf diesem (nicht alltäglichen) Weg seine Nachfrage sehr hoch treiben. Nimmt andererseits jemand an, in der City benötige man keine Lieferwagen, schädigt dieser (vernünftig denkende) Manager seine zentrale Nachfrage-Verkaufs-Relation; die im zweiten Teil von (6b) vorgenommene "Kompensation" durch eine standortabhängige Gewichtung der Zufallsvariable (vgl. Tab. 4.2) mindert diesen möglichen Schaden nur unbedeutend.

Eine Relation, die der Realität -und damit dem Vorwissen der Pbn-nahekommen könnte, betrifft die Verbindung von Hemdenpreis und Nachfrage. Ein Teil der Gleichung (6a) aus Tab. 4.2 wird in Abb. 4.3 veranschaulicht.

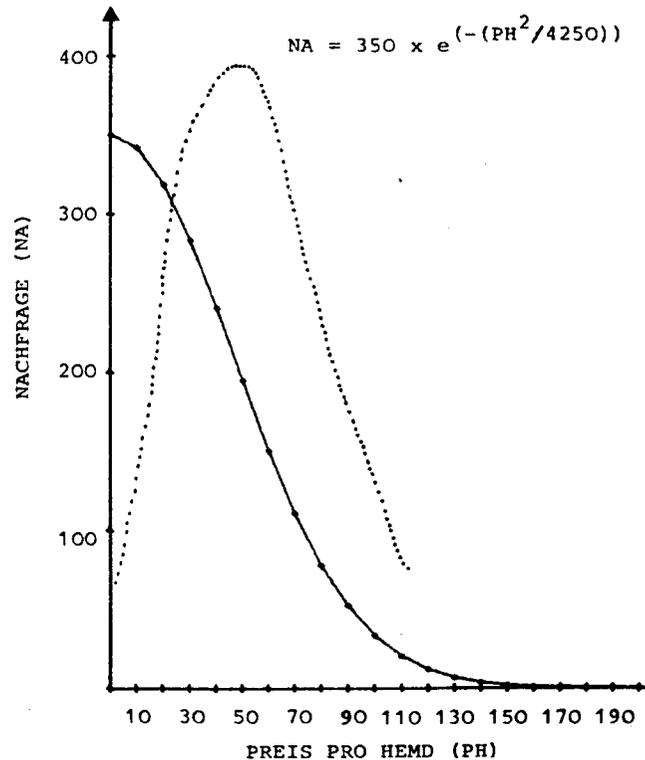


Abb. 4.3: Veranschaulichung eines Teils der Exponentialfunktion, die den Zusammenhang zwischen Hemdenpreis und Nachfrage reguliert (gepunktet: Gewinnfunktion).

Die Exponentialfunktion wird im Bereich $30 \leq PH \leq 70$ einigermaßen stabil bleiben, der Optimalwert (= Wendepunkt) liegt bei 46.10, d.h. bei diesem Wert erreicht der Gewinn $GE = NA * PH$ sein Maximum (9785.88). Bei $PH = 30$ beträgt GE immerhin noch 8496, bei $PH = 70$ noch 7735. Die schwach punktierte Hyperbel in Abb. 4.3 gibt die Gewinnfunktion in Abhängigkeit von PH für den angegebenen Gleichungsausschnitt wieder.

Für die Optimierung von Lohn- und Sozialkosten in Hinblick auf maximale Arbeitszufriedenheit kann folgendes lineare Gleichungssystem erstellt werden (vgl. Gleichung (1) in Tab. 4.2):

- (1) $1/800 * SM + 1/550 * LO = 1.7 - 0.5 + 850/550$
- (2) $SM \geq 0$
- (3) $LO \geq 850$

Will man dieses (Un-)Gleichungssystem optimieren (siehe dazu SCHICK 1981), sucht man nach dem Minimum der Summe von SM und LO unter den gegebenen Beschränkungen; der Wert dieser Zielfunktion beträgt 1516.82 und wird allein durch LO getragen, d.h. SM nimmt dabei den Wert 0 an. Maximale Zufriedenheit erreicht man also mit minimalen Aufwendungen bei $LO = 1516.82$ und $SM = 0$; der maximale Aufwand für maximale Zufriedenheit (d.h. Maximierung der obigen Zielfunktion anstelle Minimierung) beträgt 1815.56, d.h. $LO = 850$ und $SM = 965.56$ erzielen ebenfalls ein ZA von 1.7 Punkten (= ZM). Inwiefern es dem Vorwissen von Probanden entspricht, daß maximale Zufriedenheit der Arbeiter am besten durch Wegfall der Sozialausgaben bei gleichzeitiger Erhöhung der Lohnkosten auf den genannten Betrag erzielt werden kann, sei dahingestellt.

Schließlich muß auf Beziehungen hingewiesen werden, die von vielen Pbn vermutet werden, aber nicht im System implementiert sind (dies tangiert direkt das Problem der Validität des zugrundegelegten Simulationsmodells). Hierzu ist etwa die Hypothese zu zählen, im Laufe eines Jahres gäbe es saisonale Effekte oder auch die Annahme, der Preis für Rohmaterial sei abhängig von der eingekauften Menge. Ebenso kann vermutet werden, daß für Lieferwagen extra Arbeiter als Fahrer freigestellt (und bezahlt) werden müssen oder daß an den Maschinen "Springer" zur Verfügung stehen müssen. Es ließen sich noch eine Menge weiterer, durchaus realistischer Hypothesen anfügen, die ein Pbn sinnvoll aus seinem Vorwissen schöpft, die aber nicht oder gar konträr zu seiner Vermutung im System etabliert sind. Zur Frage des unkontrollierten Einflusses von Vorwissen eine kurze Passage zu einem anderen Untersuchungsinstrument, in der gerade dieser Effekt als wünschenswert dargestellt wird:

"Lohhausen war den Vpn zunächst unbekannt. Sie standen also vor der Anforderung, handeln zu müssen, ohne das System genau zu kennen und vor der Anforderung, sich während des Handelns Informationen über das System beschaffen zu müssen, um sich so sukzessiv ein Bild von dem System aufzubauen. Da allen Vpn irgendwelche kommunalen Institutionen bekannt waren, konnten sie in der Form von Analogieschlüssen ihre Erfahrungen und Kenntnisse über die Struktur von Gemeinden verwenden, um Hypothesen über die Struktur von Lohhausen aufzustellen. - Da die Übertragung (und auch die Überwindung) von vorhandenen Kenntnissen auf neue Situationen ein wesentliches Mittel der Strukturierung von Unbestimmtheit ist, konstruierten wir Lohhausen eben als Lohhausen und nicht als ein zu optimierendes abstraktes mathematisches System. Die Vpn sollten ihre Vorkenntnisse, Einstellungen und Wertsysteme in den Versuch mit einbringen." (DÖRNER et al. 1983, p.136f).

Leider ist genau der Aufbau des Systembildes (genauer: dessen Anknüpfung an bereits vorhandenes Wissen) nicht zum Untersuchungsgegenstand gemacht worden: Weder werden Angaben zur sukzessiven Etablierung eines solchen Systembildes gemacht (das sich ja wohl auch bei den "schlechten" Problemlösern feststellen ließe, die möglicherweise nur die schlechteren Anwender von Systemwissen sind), noch wurde das Vorwissen über "irgendwelche kommunalen Institutionen", die Basis für Analogieschlüsse, erfaßt. Für die bisherigen Studien zum "TAILORSHOP" gilt diese Kritik gleichermaßen.

Die Validität des Programms "TAILORSHOP" als Simulationsmodell für eine Schneiderwerkstatt ist also nach den geschilderten Details anzweifelbar. Welche Möglichkeiten gibt es, diese Frage, die sich für alle Simulationsmodelle stellt, befriedigend zu beantworten? BOSSEL et al. (1982) fordern von einem Simulationsmodell vier verschiedene Gültigkeitsnachweise:

- (1) Strukturgültigkeit im Sinne der Vollständigkeit relevanter Zusammenhänge und deren Entsprechung mit den in der Realität vorhandenen Wirkbeziehungen;
- (2) Verhaltensgültigkeit in dem Sinn, daß unter dem Spektrum möglicher Anfangs- und exogener Bedingungen dynamische Verhaltensweisen resultieren, wie sie real beobachtet werden können;
- (3) empirische Gültigkeit im Sinne des Vergleichs von Simulationsdaten mit Daten des realen Systems; wo derartige Daten nicht vorliegen, sollten Plausibilitätsüberlegungen herangezogen werden;
- (4) Anwendungsgültigkeit in dem Sinne, daß das Simulationsmodell dem Untersuchungszweck entspricht und die dort gewünschte Information liefert.

Natürlich sind im Rahmen psychologischer Studien mit Simulationsmodellen nicht alle genannten Aspekte von gleicher Bedeutsamkeit (vgl. als Beispiel für eine genaue Prüfung aller genannten Validitätsaspekte BOSSEL et al. 1982, p.83f), da man vor allem auf die resultierenden psychologischen Effekte abheben möchte. Ein pointierter Standpunkt könnte lauten: wie das Simulationsmodell aussieht, spielt gar keine Rolle, solange interessante psychologische Daten anfallen. Dieser z.B. von DÖRNER vertretene Standpunkt muß allerdings damit rechnen, daß man ein Scheitern menschlichen Interventionsverhaltens entweder auf die mangelnde bzw. von V_p zu V_p unterschiedliche Kompatibilität von impliziten Vorannahmen und realisiertem Simulationsmodell zurückführt oder aber "chaotisches" Systemverhalten als Ursache individuellen Scheiterns betrachtet. Dies kann nicht beabsichtigt sein.

Die eben geschilderten Validitätsprobleme hat einer der Ko-Autoren der "LOHHAUSEN"-Studie, KREUZIG (1983), recht deutlich gesehen, indem er fragt, wie real die "kleine Welt am Draht" (KREUZIG 1983, p.150) wirklich sei bzw. ob dort nicht eine Pseudo-Welt im Labor errichtet werde. "Die schon bislang mageren Angaben zur Validität in den Testmanualen (gemeint sind hier 'klassische' Testverfahren, J.F.) lassen nichts Gutes ahnen für die -ungleich aufwendigeren- Simulationen."

(KREUZIG 1983, p.151). In diesem Zusammenhang weist KREUZIG ebenfalls darauf hin, daß gänzlich unklar sei, welche realen Aspekte am Modell entfallen dürften, ohne dessen Repräsentativität zu schmälern. Diesem Problem wird man im Rahmen derartiger Simulationsstudien ständig begegnen. Der "TAILORSHOP" ist nicht das einzige Simulationsmodell, das auf die Frage nach seiner Validität eine Antwort schuldig bleibt.

4.2 Eine Simulationsstudie zum Simulationsprogramm

Ein naheliegender Ansatz zur Untersuchung bestimmter Systemeigenschaften und insbesondere zur Überprüfung der Validität der vorgeschlagenen Leistungsindikatoren (Anzahl der Monate mit Aufwärtstrend in den Variablen "Flüssigkapital" nach PUTZ-OSTERLOH bzw. "Gesamtkapital" nach FUNKE; TRENDPO bzw. TRENDFU) besteht darin, das computer-simulierte System "TAILORSHOP" selbst zum Gegenstand einer Simulationsstudie zu machen. Nachfolgend wird gezeigt werden, welche Aufschlüsse bei dieser Vorgehensweise zu erzielen sind. Dabei handelt es sich keineswegs um eine Alternative zur systemtheoretischen Analyse, sondern um eine Ergänzung dieser Sichtweise.

Die Frage, die mit der Simulation des Systems verfolgt wird, bezieht sich auf die Verteilungseigenschaften der beiden konkurrierenden AVn TRENDPO und TRENDFU unter zwei verschiedenen Simulationsbedingungen:

- (a) sämtliche Input-Variablen werden per Zufall und unabhängig voneinander bestimmt (= Randombedingung);
- (b) nur eine einzige Input-Variable wird per Zufall bestimmt, die restlichen Eingriffe werden durch Steuer- bzw. Regelungsaggregate (zur Unterscheidung vgl. CZAYKA 1974, p.50ff) kontrolliert (= "rationale" Bedingung).

Vorausgesetzt, die zur Systemstabilisierung verwendeten Regelungsmechanismen sind wirksam, darf man von einem Leistungsgüteindikator erwarten, daß er unter der hier als "rational" bezeichneten Bedingung tatsächlich bessere Systemhandhabung anzeigt als unter der Bedingung reiner Zufallseingriffe. Die durch eine Simulation erfahrbare Verteilung der Index-Werte auf die möglichen Kategorien dieser Variablen -im Fall der Trendwerte zwischen null (= niemals Aufwärtstrend) und zwölf (= immer Aufwärtstrend) liegend- muß also auf die unterschiedlichen Simulationsbedingungen sensibel reagieren: im Fall der Randomstrategie sollte Kategorie Null die häufigst frequentierte sein, unter rationalen Bedingungen je nach Güte der Steuerglieder höhere Kategorien, im Optimalfall Kategorie zwölf.

Für die Randomstrategie ist es erforderlich, bei den kontinuierlich abstufbaren Eingriffsvariablen einen Wertebereich zu definieren. Prinzipiell wäre es ja möglich, etwa eine Million Hemden einzukaufen oder mit 200000 Lieferwagen zu operieren. Um diese offenkundig unsinnigen Werte auszuschließen, orientiert sich der mögliche Wertebereich an den in der empirischen Studie von FUNKE (1983b) aufgetretenen Werten.

Allerdings wurden offensichtliche Ausreißer (z.B. 20000.- DM Reparaturkosten, nur einmal aufgetreten) dabei ausgelassen. Zwischen P_{10} und P_{90} liegende Werte bestimmten in etwa den Range. Folgende Bereichsbeschränkungen wurden für die Simulation "RANDOM" zugrundegelegt:

Rohmaterial	RM: 0-1000
Hemdenpreis	PH: 40-65
Werbung	WE: 0-7500
Arbeiter 50er	N1: 0-10
Arbeiter 100er	N2: 0-6
Maschinen 50er	A1: 0-10
Maschinen 100er	A2: 0-6
Reparatur	RS: 1000-3000
Lohn	LO: 1000-1300
Soziales	SM: 50-70
Lieferwagen	LW: 1-5.

Bei der Wahl der (nominal skalierten) Geschäftslage könnte man die Verteilung in der Stichprobe von $N = 48$ mit $p(\text{Vorstadt}) = 0.21$, $p(\text{Stadttrand}) = 0.46$ und $p(\text{City}) = 0.33$ für die Simulation verwenden. Wir haben hier jedoch eine konstante Lage angenommen (siehe unten).

Jede der insgesamt 1000 Simulationen erlaubte die zwölfmalige Wahl der 12 zu bestimmenden Eingriffsvariablen per Zufall. Die Berechnungen konnten sämtlich auf einem Kleinrechner durchgeführt werden; die Simulation eines zwölfmonatigen Eingriffs erforderte ca. 10 sek CPU-Zeit, also eine Gesamtdauer von gut drei Stunden für 1000 Simulationen. Die gleichverteilten Zufallszahlen stammen von den eingebauten Pseudo-Zufallszahlengeneratoren der jeweiligen Geräte.

Während es wenig Mühe macht, eine derartige unkontrollierte Eingriffsstrategie zu definieren (lediglich das Problem der Bereichsbeschränkungen muß geklärt werden), bleibt die Herleitung einer als "rational" zu kennzeichnenden, systematischen Eingriffsstrategie immer willkürlich und damit problematisch, solange es kein Optimum der Systemsteuerung gibt - wie im vorliegenden Fall. Aus diesem Grund kommt der Schilderung der hier als "RATIO 1" bezeichneten Simulation ein nicht unwesentlicher Stellenwert zu, da andere Bedingungen denkbar und machbar sind, die mit gleichem Recht die Bezeichnung "rational" für sich beanspruchen können (wie die Kennzeichnung "RATIO 1" bereits andeutet, ist es eine von mehreren, die wir verwendet haben). Anliegen der Simulation muß es daher sein, verschiedene "rationale" Programme auszutesten mit der Absicht, aus der möglichen Unterschiedlichkeit der Befunde Aussagen über die mögliche differentielle Validität der zugrundegelegten AVn zu machen.

Die in RATIO 1 verwendeten Heuristiken gehen davon aus, daß es einen einzigen freien Parameter gibt (d.h. eine einzige Variable schwankt per Zufall) und die übrigen Parameter in Abhängigkeit von diesem ersten bestimmt werden. Diese "zufällige" Variable soll das eingekaufte Rohmaterial RM sein; es handelt sich hierbei um die erste vom Probanden verlangte Entscheidung jeden Monats, auf der andere Ent-

scheidungen aufgebaut werden können. Diese Auswahl ist trotz allem eine willkürliche, da es eine begründbare Rangordnung der verschiedenen Eingriffsmöglichkeiten nicht gibt.

Die der Simulation gemäß dem Modell RATIO 1 zugrundegelegten Entscheidungsregeln laufen im einzelnen wie folgt (neben der umgangssprachlichen wird auch zugleich die Formulierung in Variablenform verwendet, um Eindeutigkeit zu gewährleisten):

Regel 1 (bestimmt Hemdenpreis): Wenn mehr fertige Hemden im Lager sind als Rohmaterial vorhanden ist und der bisherige Preis pro Hemd größer als 40 ist, senke den Preis um 1 Einheit. Überwiegt dagegen das Rohmaterial, ist das Hemdenlager also nicht überfüllt, und liegt der gegenwärtige Preis unterhalb von 65, dann erhöhe den Preis um 1 Einheit.

Formal: $(HL \geq RL) \ \& \ (PH \geq 40) \ \rightarrow \ PH = PH - 1,$
 $(HL < RL) \ \& \ (PH < 65) \ \rightarrow \ PH = PH + 1.$

Auf diese Weise kann ein überfülltes Fertiglager zu verbilligten Preisen "verramscht" werden, während bei reduziertem Lagerbestand an Fertigware eine Preiserhöhung die nicht zu befriedigende Nachfrage bremst.

Regel 2 (bestimmt Werbungskosten): Sinkt die Nachfrage unter 500 Hemden, wird der Etat für Werbungskosten um 100 Einheiten erhöht. Liegt die Nachfrage zwischen 500 und 1000 Hemden, erhöhe um 50 Einheiten. Bei über 1000 nachgefragten Hemden kann die Werbung um 100 Einheiten reduziert werden.

Formal: $(NA \leq 499) \ \rightarrow \ WE = WE + 100$
 $(500 \leq NA \leq 1000) \ \rightarrow \ WE = WE + 50$
 $(NA \geq 1001) \ \rightarrow \ WE = WE - 100$

Der Zweck dieser Maßnahme liegt auf der Hand: es soll nur soviel für WE ausgegeben werden wie eben nötig ist, um eine Nachfrage nach 1000 Hemden zu erzielen.

Regel 3 (bestimmt Arbeiter- und Maschinenzahl): Der Einfachheit halber wird angenommen, daß nur 100er-Maschinen zum Einsatz kommen sollen, da sie bei gleichem Personalaufwand doppelte Produktionsleistung erreichen (also: $A1 = N1 = 0$). Die Maschinen sollen nun in der Zahl zur Verfügung stehen, wie durch die Nachfrage Bedarf entsteht: Besteht Bedarf nach 500 Hemden ($NA = 500$), reichen 5 100er-Maschinen ($A2 = 5$) aus. Entsprechend der Maschinenzahl soll die Personalplanung ausgerichtet sein: Pro Maschine soll genau ein Arbeiter bereitstehen.

Formal: $N1 = A1 = 0$
 $N2 = A2 = \text{INT}(NA/100 + 0.5).$

Selbstverständlich wären hier Alternativen denkbar, bei denen der möglicherweise rasche Wechsel von An- und Verkauf von Maschinen (und die damit möglicherweise verbundenen Wertverluste) vermieden wird. Als grobe Heuristik scheint die Regel jedoch brauchbar.

Regel 4 (bestimmt Reparatur/Service): Solange der aktuelle Maschinenstand unterhalb des maximalen liegt, soll die Reparatur- und Serviceleistung rasch erhöht werden. Bei intakten Maschinen kann eine allmähliche Reduktion vorgenommen werden, solange die Maschinen keine Schäden zeigen.

Formal: $(MA/MM) < .99 \rightarrow RS = RS + 500$
 $(MA/MM) \geq 1 \rightarrow RS = RS - 100.$

Mit dieser Maßnahme kann eine rasche Verringerung der Schäden bewirkt werden, die langsamere Rücknahme bei Erreichen des Schwellenwerts verhindert unnötige Unkosten.

Regel 5 (bestimmt Lohnkosten): Solange die aktuelle Arbeitszufriedenheit unter der maximalen liegt, soll der Lohn erhöht werden. Ist die maximale Zufriedenheit erreicht, kann ein verlangsamter Abbau erfolgen bis zum Erreichen des Schwellenwerts.

Formal: $(ZA/ZM) < .99 \rightarrow LO = LO + 50$
 $(ZA/ZM) \geq 1 \rightarrow LO = LO - 10.$

Ganz ähnlich wie Regel 4 wird in raschem Schritt auf den Sollzustand zugegangen, um dann ganz allmählich wieder die Unkosten zurückzunehmen.

Mit diesen fünf Regeln ist die Heuristik der Simulation **RATIO 1** beschrieben. Festzuhalten bleiben noch die Konstanten $SM = 50$, $GL = 0.5$ (also Cityrand) und $LW = 1$, die aus Gründen der Einfachheit nicht variiert wurden. Für RP , PM und NA_2 gelten die üblichen Zufallsbedingungen.

Tab. 4.5 enthält die Resultate der Simulationsstudie. Neben der **Randombedingung** sind sechs Varianten der "rationalen" Bedingung aufgeführt, die sich durch unterschiedliche Konstanten für $N1$, $A1$, $N2$ und $A2$ auszeichnen (siehe Tabellenfußnote).

Ein Vergleich der Rohwertverteilungen von **TRENDPO** und **TRENDFU** unter den verschiedenen Simulationsbedingungen ist recht aufschlußreich. Unter der **Randombedingung** wäre von einem validen Indikator zu erwarten, daß er keinen hohen Wert anzeigt, da das Eingriffsverhalten ja völlig unsystematisch und ohne Bezug zur aktuellen Systemlage erfolgt. Tatsächlich erhält man für **TRENDFU** einen Mittelwert von 0.37, für **TRENDPO** dagegen einen von 4.21. Unter der "rationalen" Bedingung sollte der Indikator dagegen ausschlagen; dies tritt bei **TRENDFU**, nicht aber bei **TRENDPO** auf. Im Vergleich zu **RANDOM** liegt **TRENDPO** bei **RATIO 1** gar niedriger!

Tab. 4.5: Rohwertverteilungen für die Güteindikatoren nach **PUTZ-OSTERLOH** (1981; **PO**) bzw. **FUNKE** (1983b; **FU**) unter verschiedenen Simulationsbedingungen mit Mittelwert x und Streuung s (jeweils $N = 1000$; KA_{12} = mittleres Flüssigkapital im 12. Monat, GK_{12} = mittleres Gesamtkapital im 12. Monat).

Rohwert	RANDOM ^a		RATIO1 ^b		RATIO2 ^c		RATIO3 ^d		RATIO4 ^e		RATIO5 ^f		RATIO6 ^g	
	PO	FU	PO	FU	PO	FU	PO	FU	PO	FU	PO	FU	PO	FU
0	0	675	5	2	0	0	988	587	998	82	0	0	0	0
1	5	284	55	2	0	0	12	345	2	313	0	0	0	0
2	38	33	190	22	0	0	0	59	0	359	0	0	2	0
3	210	8	310	81	6	3	0	8	0	201	0	1	7	3
4	354	0	244	233	7	8	0	1	0	40	1	2	30	18
5	289	0	146	286	25	15	0	0	0	5	2	3	90	73
6	94	0	42	257	71	26	0	0	0	0	3	11	159	157
7	9	0	8	92	151	119	0	0	0	0	11	34	240	257
8	1	0	0	22	304	250	0	0	0	0	34	125	255	317
9	0	0	0	3	370	555	0	0	0	0	171	821	171	154
10	0	0	0	0	66	24	0	0	0	0	777	3	46	21
11	0	0	0	0	0	0	0	0	0	0	1	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x	4.2	0.4	3.4	5.0	8.1	8.3	0.0	0.5	0.0	1.8	9.7	8.7	7.3	7.3
s	1.1	0.6	1.3	1.3	1.2	1.0	0.1	0.6	0.0	1.0	0.7	0.7	1.5	1.3
KA_{12}	-196736	-103733	+ 25875	- 95326	-249417	+ 95125	+ 66402							
GK_{12}	- 99296	+ 59456	+140780	+ 32734	- 29245	+156271	+111147							

^a absolute Unabhängigkeit der Eingriffsparameter, deren Range aus empirischen Min-/Max-Werten gewonnen wurde

^b nur Rohmaterial zufällig schwankend, ansonsten "rationale" Heuristik wie im Text beschrieben

^c wie b, nur mit $N1 = A1 = 0$ und $N2 = A2 = 6$

^d wie b, nur mit $N1 = A1 = 12$ und $N2 = A2 = 0$

^e wie b, nur mit $N1 = A1 = 0$ und $N2 = A2 = 12$

^f wie b, nur mit $N1 = A1 = 0$ und $N2 = A2 = 3$

^g wie b, nur mit $N1 = A1 = 0$ und $N2 = A2 = 2$

Dennoch könnte man angesichts der weiteren Resultate in Tab. 4.5 versucht sein, vor der Schlußfolgerung zu warnen, damit sei die Überlegenheit von **TRENDFU** gegenüber **TRENDPO** im Rahmen einer Simulationsstudie bestätigt. Setzt man Regel 3 der Strategie **RATIO 1** außer Kraft und arbeitet stattdessen -wie in **RATIO 2** bis **6** geschehen- mit konstanten, aber jeweils unterschiedlichen Werten für Arbeiter- und Maschinenzahlen, treten auf den ersten Blick merkwürdige Effekte auf. Die Überlegenheit von **TRENDFU** schwächt sich deutlich ab, darüberhinaus wird klar, wie massiv diese Konstanten wirken. Vergleicht man die

Verteilungen von RATIO 2 und 3, bei denen die gleiche Produktionskapazität von monatlich 600 Hemden einmal durch 100er-, einmal durch 50er-Maschinen erzeugt wird (wobei sich der Personalbedarf verdoppelt), stellt man ein Absinken von durchschnittlich gut acht Monaten mit Aufwärtstrend auf unter eins fest! Versucht man, die effiziente Strategie RATIO 2 durch Verdoppelung der Produktionskapazität auf 1200 Hemden noch zu steigern (wie in RATIO 4 geschehen), wird man eines besseren belehrt: Beide Indikatoren zeigen durch ihre zugehörigen Verteilungen die Schwäche einer derartigen Maschinenaufrüstung an. Nimmt man dagegen die Zahl der Maschinen auf die Hälfte von RATIO 2 zurück, also eine Kapazität von 300 Hemden pro Monat, erhält man die unter RATIO 5 aufgeführten Verteilungen, deren Mittelwerte die höchsten überhaupt vorkommenden darstellen. Eine noch weitergehende Senkung auf 200 Hemden (vgl. RATIO 6) bringt bereits Verschlechterungen mit sich.

Neben diesen interessanten Befunden zur Auswirkung der Maschinen- und Arbeiterzahl auf die beiden alternativen Güteindikatoren stellt sich die Frage, warum außer unter RANDOM und RATIO 1 keine markanten Unterschiede zwischen TRENDPO und TRENDFU auftreten. Die Antwort darauf ist relativ trivial: Wie bereits in der Begründung für die Ableitung des TRENDFU-Indikators beschrieben, ist TRENDPO sehr anfällig für unterschiedliche Investitionstätigkeiten der Probanden. Aber in RATIO 2 bis 6 wurde ja gerade die mögliche Schwankung in der Maschinenausrüstung durch die Wahl entsprechender Konstanten ausgeschaltet, so daß diese Schwäche von TRENDPO dort gar nicht augenfällig werden konnte. Lediglich RANDOM und RATIO 1 lassen variierende Investitionen zu; dabei unterscheidet sich TRENDPO kaum unter RANDOM bzw. RATIO 1. Von daher sind die Resultate der Simulationsstudien sehr wohl Belege für die Überlegenheit von TRENDFU gegenüber TRENDPO, wenn man die von unterschiedlicher Investitionsneigung NICHT beeinflusste Güte des Managements erfassen möchte. Tätigt man keine Investitionen, macht es keinen Unterschied, welchen Indikator man verwenden möchte; dies dürfte jedoch nicht der Normalfall sein. Der Nutzen einer derartigen Simulation besteht darin, daß bestimmte Fragen etwa bezüglich dessen, was verschiedene Indikatoren eigentlich indizieren oder von welchen weiteren Variablen diese Indikatoren abhängig sein könnten, bereits vor jeder Datenerhebung zu beantworten sind. Darüberhinaus erhält man so ganz interessante Einblicke in das, was nach dem gewählten Indikator als gut oder schlecht bezeichnet werden muß. Bei zukünftigen Studien sollte ein derartiger Simulationsansatz bei der Indikatorbildung zu Rate gezogen werden.

Selbstverständlich kann eine "Simulation der Simulation" auch einen anderen Akzent erhalten: Anstatt das Verhalten von Kennwerten zu untersuchen, fragt man sich nach dem Satz von Regeln, der Verlaufsdaten ähnlich den realen Vpn-Daten liefert. Man könnte sogar daran denken, ein "Expertensystem" (vgl. SAVORY 1985) zu entwickeln, das Heuristiken für den Umgang mit dem Simulationssystem bereitstellt. Die Anzahl der Propositionen eines derartigen Expertensystems, das der Leistung einer "guten" Vp entsprechen sollte, wäre ein weiterer Aspekt zur Bestimmung des objektiven Schwierigkeitsgrads eines Systems.

4.3 Bisherige Studien zum "TAILORSHOP"

Studien mit dem hier vorgestellten Simulationsprogramm liegen vor von PUTZ-OSTERLOH (1981, 1983b) und PUTZ-OSTERLOH & LÜER (1981) sowie von FUNKE (1983b); auf diese Arbeiten soll kurz eingegangen werden. Die Zielvorstellungen der ersten Studie mit dem "TAILORSHOP" faßt PUTZ-OSTERLOH (1981, p.80) wie folgt zusammen:

"Es gilt also, die Anforderungen von komplexen Problemen und von Intelligenztestaufgaben zu beschreiben und zu untersuchen, welche Problemlöseprozesse zu ihrer Überwindung eingesetzt werden. Dabei sollen komplexe Probleme alltäglichen Problemsituationen stärker angenähert sein als es Intelligenztestaufgaben sind."

Drei Anforderungen werden als charakteristisch für komplexe Problemlöseprozesse genannt (PUTZ-OSTERLOH 1981b, p.83):

- (1) das Aufstellen und Ableiten von Problemlösezielen,
- (2) die Auswahl von Handlungen zum Erreichen der Ziele und
- (3) aktive Suche nach Informationen über relevante Systemvariablen.

Gemeinsamkeiten zwischen IQ-Test und komplexem Problemlösen sieht die Autorin in folgenden Anforderungen:

- (4) Analyse der Veränderung von Variablen,
- (5) In-Beziehung-Setzen von (4) mit der Veränderung anderer Variablen,
- (6) Ziehen von Analogieschlüssen sowie
- (7) Aufstellen von Regeln zur Beschreibung von Veränderungen.

Entsprechend diesen Annahmen wird die (psychologische) Hypothese (ohne expliziten Bezug zu einer Theorie der Intelligenz und ohne expliziten Bezug zu einer Theorie des Problemlösens) formuliert, daß Testintelligenz nur dann ein Prädiktor für erfolgreiches Problemlösen sei, wenn hinsichtlich der Anforderungen (1) bis (3) kein Unterschied zwischen Intelligenztestaufgabe und komplexem Problem bestehe, wenn also im komplexen Problem Ziel- und Informationssuche entfalle. Dagegen solle der Prädiktionswert des IQ-Tests gering ausfallen, wenn allein das komplexe Problem die Anforderungen (1) bis (3) stelle, die für die Bearbeitung von IQ-Tests nicht erforderlich seien.

Bevor das Verfahren und die Ergebnisse der Hypothesenprüfung geschildert werden, soll kurz innegehalten und nach der Sinnhaftigkeit der formulierten Hypothese gefragt werden. - Was würde man zu folgendem hypothetisch angenommenen Fall sagen? Ein Forscher behauptet: Eine Variable x, die mit y eine Reihe Gemeinsamkeiten aufweist, korreliert positiv mit dieser. Verändere ich jedoch x und mache es unähnlicher zu y, sinkt der Zusammenhang zwischen beiden Variablen. Diese Trivialität wird dadurch verschleiert, daß man behauptet, man wisse nicht genau, was x sei, noch wisse man genau, was y sei. Wäre es nicht ratsam, zunächst näheres über x und y zu erfahren, bevor man Aussagen über deren Zusammenhang macht (vgl. BRANDTSTÄDTER 1982)? - Zurück zur Un-

tersuchungsrealität.

Zur Prüfung der eben dargestellten Hypothese eines Zusammenhangs zwischen IQ-Test und komplexem Problemlösen für den Fall übereinstimmender Anforderungen werden drei Versuchsbedingungen mit zunehmendem Ausmaß an Übereinstimmung zwischen beiden Variablen realisiert:

- (1) intransparente Bedingung ohne Zielvorgabe,
- (2) intransparente Bedingung mit Zielvorgabe sowie
- (3) transparente Bedingung.

Die Hypothese wird dahingehend präzisiert, daß in der genannten Reihenfolge die Korrelation zwischen Problemlöseerfolg und Intelligenztestwerten anwachsen (= statistische Hypothese; zum Verhältnis wissenschaftlicher zu statistischen Hypothesen vgl. WESTERMANN & HAGER 1982). Die drei genannten Bedingungen unterschiedlicher Gemeinsamkeit werden so definiert, daß unter (1) den Vpn kein Ziel vorgegeben wird, wohl aber alle Eingriffsmöglichkeiten zur Veränderung des Systems genannt werden, bei (2) zusätzlich zu (1) das Ziel der Problembearbeitung genannt wird und bei (3) zusätzlich der aktuelle Zustand aller beeinflussbaren Systemvariablen beschrieben wird (nach PUTZ-OSTERLOH 1981, p.84-85). Außerdem erhält Gruppe (3) in der Instruktion Hinweise über einige Beziehungen (welche?) zwischen den Variablen.

Die Einzelversuche mit dem "TAILORSHOP" (24 simulierte Monate) wurden durch V1 geleitet, die entsprechend den Versuchsbedingungen den Vpn Auskünfte gaben und die vom Probanden gewünschten Maßnahmen in den Taschenrechner eingaben. Als Maß der Intelligenz diente der Summenscore der "Advanced Progressive Matrices" von RAVEN (1965). Als Erfolgsmaß wurde die Anzahl der Monate mit einem Anstieg der Variable "Flüssigkapital" gegenüber dem Vormonat ausgezählt (=TRENDPO).

Von N = 83 Probanden waren infolge Eingabefehlern, Ausfall technischer Geräte und unzulässiger Informationsweitergabe durch den V1 letztlich nur 48 verwertbar (= 57.8 %), was die Autorin selbst problematisiert. Da Probanden von Gruppe (1) durchweg als Ziel die Kapitalvermehrung bzw. Gewinnmaximierung nannten -so PUTZ-OSTERLOH (1981, p.90)-, wurden sie zusammen mit den Vpn aus Gruppe (2) zu einer Kategorie "Intransparenz" (1+2) zusammengefaßt.

Tau-Rangkorrelationen zwischen APM-Score und TRENDPO, dem Kapitalanstiegswert, liefern für (1+2; N = 35) einen Wert von -.22, für (3; N = 13) einen Wert von -.01! Die Hypothese, nach der der "Erfolg im Umgang mit diesem Kleinsystem ... mit der Leistung bei den APM-Aufgaben positiv interkorreliert sein" müßte (PUTZ-OSTERLOH 1981, p.84), bestätigt sich nicht, wenngleich das prognostizierte Zusammenhangsmu-ster in seiner Rangordnung mit den Daten übereinstimmt.

Weitere Befunde aus dieser "Pilot"-Studie sollen hier nicht referiert werden wegen der Bedenken bezüglich der Stichprobenselektivität. Hier sind in der Replikationsstudie von PUTZ-OSTERLOH & LÜER (1981) Verbesserungen anzutreffen. Die Bezeichnung "Replikation" ist nicht

ganz unumstritten: LÜER (schriftl. Mitteilung) hält die Arbeit nicht für eine Replikation der Studie von PUTZ-OSTERLOH (1981) "im methodischen Sinne"; m.E. werden jedoch nicht nur Ergebnisse repliziert ("... Replikation der Ergebnisse aus der ersten Untersuchung ..."; PUTZ-OSTERLOH & LÜER 1981, p.322), sondern bereits in der Anlage der Studie wird eine Replikations-Absicht formuliert (vgl. die Beschreibung von "Fragestellung und Hypothesen", in der es heißt, die Studie diene "... als Replikation und Überprüfung der vorangegangenen Untersuchung ..."; PUTZ-OSTERLOH & LÜER 1981, p.316). Legt man WINERS (1970, p.213) strenge Definition zugrunde:

"A replication of an experiment is an independent repetition under as nearly identical conditions as the nature of the experimental material will permit.",

kann dies als teilweise zutreffend bezeichnet werden, weil es explizit bei PUTZ-OSTERLOH & LÜER heißt, daß eine Gruppe (I) "unter einer Instruktionsbedingung wie bei PUTZ-OSTERLOH (1981)" auftreten soll. Eine interexperimentelle konzeptuelle Replikation (vgl. BREDEKAMP 1980, p.34) liegt insofern vor, als die UV "Transparenz" nunmehr durch Vorgabe einer Grafik mit der Variablenvernetzung realisiert wird. Von einer intraexperimentellen Replikation würde man sprechen, wenn entweder UV oder AV oder beide in einem Experiment verschiedenartig realisiert werden; dieser Fall liegt nicht vor. Da die konzeptuelle Replikation für BREDEKAMP

"von größter Wichtigkeit (ist), ... ist unbedingt zu empfehlen, in jedem Experiment Alpha, Beta und Effektgröße festzusetzen und das benötigte n zu schätzen, damit überhaupt Aussagen über die Replizierbarkeit der Befunde möglich sind." (1980, p.35).

Wie eben gesagt, werden in der Replikationsstudie zwei Bedingungen eingerichtet: Bedingung I entspricht Gruppe 3 (dies vermute ich jedenfalls; es wird im Text nicht genau beschrieben, welche der drei Gruppen (1) bis (3) in der Folgestudie als Bedingung I bezeichnet wird). Bedingung G dagegen enthält zusätzlich die Vernetzungsgrafik (mit Ausnahme der "internen" Variablen und der Vorzeichen identisch mit dem Schema aus Abb. 4.1). Als Indikator für Testintelligenz dient wiederum der APM von RAVEN (1965); als Gütekriterium für den Problemlöseerfolg wird erneut der Kapitalanstiegswert TRENDPO herangezogen. 70 studentische Vpn (davon 68 verwertbar) bearbeiteten den "TAILORSHOP" in Taschenrechnerform (also wieder V1-vermittelt) über 15 simulierte Monate.

Als Ergebnis zeigt sich bezüglich der Hypothese, unter I bestehe eine Nullkorrelation zwischen TRENDPO und APM-Score, unter G dagegen lasse sich ein "systematischer" Zusammenhang erkennen, ein Tau von 0 für die I- und ein Tau von 0.31 für die G-Bedingung. Die Schlußfolgerung der Autoren:

"Damit ist (a) die früher gefundene Korrelation von 0 bei der Instruktionsbedingung an einer neuen Stichprobe repliziert wor-

den; (b) unsere Haupthypothese über die Wirkung eines Schaubildes zur Reduktion von Intransparenz gestützt worden. Der Erfolg der Systemsteuerung unter Einbeziehung einer grafischen Abbildung steht in systematischer Beziehung zum Lösungserfolg bei APM-Aufgaben." (PUTZ-OSTERLOH & LÜER 1981, p.319).

Kritisch vermerkt sei an dieser Stelle, daß die Umsetzung der wissenschaftlichen Hypothese (= wH) in eine statistische (= sH) an dieser Stelle schwerwiegende Annahmen bzw. mögliche Fehler enthält (zum Umsetzungsproblem vgl. WESTERMANN & HAGER 1982). Aus der vorgelegten Arbeit von PUTZ-OSTERLOH (1981) geht folgende wH hervor:

wH₁: Mit wachsender Transparenz eines Problems ist mit einem Anstieg der Korrelation zwischen Intelligenztest und Problemlöseerfolg zu rechnen.

Aus der Arbeit von PUTZ-OSTERLOH & LÜER (1981) kann man eine wH₂ herauslesen, die als Spezialfall von wH₁ angesehen werden muß:

wH₂: (a) unter Transparenz: Korrelation zwischen IQ-PL > 0,
(b) unter Intransparenz: Korrelation zwischen IQ-PL = 0.

Einmal abgesehen von der Problematik, die unter wH₂ Teil (b) formulierte Punkthypothese teststark zu prüfen (Powerangaben wurden nicht gemacht), ist auch zu fragen, ob aus der "Theorie" überhaupt eine derartig präzise Prognose ableitbar ist oder ob nicht eher wH₁ zu prüfen ist. Die Autoren haben die spezifischere wH₂ derart in die sH₂ überführt, daß (a) $Tau_{IQ-PL} > 0$ unter Transparenz und (b) $Tau_{IQ-PL} = 0$ unter Intransparenz; dabei kam es zu dem oben geschilderten und scheinbar bestätigenden Befund. Will man dagegen die allgemeinere wH₁ in eine geeignete sH₁ umsetzen, prüft man -dies haben die Autoren unterlassen- einseitig auf Unterschiedlichkeit der beiden Tau-Koeffizienten. Nach dem bei LIENERT (1973², p.652) beschriebenen schärferen Verfahren habe ich das Konfidenzintervall um die Tau-Koeffizienten bestimmt und jeweils einseitig geprüft (für Tau₁: $0.00 \pm 1.65 * 0.239$; für Tau₂: $0.31 \pm 1.65 * 0.234$; Alpha = 0.05). Die Prüfung zeigt, daß für beide Tau-Werte gilt: Berechnet man das Konfidenzintervall des jeweiligen Tau, liegt das andere Tau innerhalb der Grenzen des Intervalls! D.h.: bei einer Prüfung der Hypothese wH₁ -und dies ist die zur Zeit besser begründbare- mittels sH₁ kommt man zu dem Schluß, daß die gefundenen Tau-Werte sich statistisch nicht bedeutsam unterscheiden. Die oben referierte Schlußfolgerung von PUTZ-OSTERLOH & LÜER (1981) steht unter diesem Blickwinkel mit den Daten nicht mehr in Einklang.

Neben diesem -möglicherweise zweifelhaften- Hauptbefund gibt es natürlich noch eine Reihe anderer Befunde aus den Protokollen des lauten Denkens, zu dem die Vpn (übrigens auch in der Studie von PUTZ-OSTERLOH 1981) angehalten wurden. Verglichen werden Daten von erfolgreichen bzw. erfolglosen Vpn (am Median des Gütekriteriums dichotomisiert) hinsichtlich (a) Fragen nach Variablen und Maßnahmen, (b) Hypothesen über Variablen, (c) Planungen und Zielen, (d) Analysen von

Variablen, (e) Berechnungen, (f) Entscheidungen und (g) V1-Äußerungen (vgl. PUTZ-OSTERLOH 1981, p.87f). Die Fülle signifikanter (und auch nichtsignifikanter) Befunde wollen wir dem Leser zur geflissentlichen Lektüre überlassen, zumal der Wert der inferenzstatistischen Überprüfungen durch fehlende Alpha-Adjustierung und fehlende Teststärkeangaben in Ermangelung konkreter Hypothesen stark geschmälert wird. Daß die untersuchten Variablenbereiche von Interesse sind, soll nicht geleugnet werden; jedoch mangelt es an einer Theorie des Problemlösens, die konkrete Prognosen zuließe.

Interessant ist ein im letzten Teil der Arbeit von PUTZ-OSTERLOH & LÜER (1981, p.331f) gemachter Hinweis auf eine Studie mit einem "biologischen Tiefseesystem" namens "Urgos": Deutlich wird nämlich die Annahme der Autoren, Komplexität werde allein durch die Anzahl der beteiligten Variablen bestimmt (vgl. die Ausführungen in Kap. 3.4).

"Gegenüber dem System der Schneiderwerkstatt war dieses System in seiner Komplexität drastisch reduziert: Es bestand aus nur sechs Variablen (drei Tier- und drei Pflanzenarten), die durch fünf Relationen miteinander verknüpft waren." (p.331, Unterstreichungen von mir).

Wohl wegen seiner "Einfachheit" werden die Ergebnisse dazu in vier kurzen Absätzen geschildert; eine positive Korrelation zwischen der Systembewältigung und APM-Intelligenztest läßt sich nachweisen.

In der Studie von PUTZ-OSTERLOH (1983b) wird der Zusammenhang PL-Erfolg und Testintelligenz nur noch in der Einleitung erwähnt mit dem Tenor:

"Die Intelligenz, wie sie durch herkömmliche Intelligenztests erfaßt werden kann, hat jedenfalls keinen Einfluß auf die Leistungen bei der Bearbeitung komplexer Probleme (...)." (PUTZ-OSTERLOH 1983b, p.104).

Angesichts der deutlichen Schwächen der Arbeiten, die als Beleg angeführt werden (ausschließlich vier Arbeiten der eigenen Forschungsgruppe), mutet diese Feststellung merkwürdig an, zumal noch zwei Jahre zuvor von "logisch wie auch inhaltlich schwer verständlichen Nullkorrelationen zwischen Intelligenztest- und Problemlöseleistungen" (PUTZ-OSTERLOH & LÜER 1981, p.330) gesprochen wurde. - Gegenstand dieser letzten Studie sind nun Verhaltensunterschiede, die im Zusammenhang mit beobachteten Leistungsunterschieden stehen. In 15 simulierten Monaten sollten je 30 Vpn unter einer Kontroll- (= KG), einer Reflexions- (= RE) bzw. einer Rückmeldungsbedingung (= RÜ) den "TAILORSHOP" leiten. Reflexion bedeutete, daß die Vpn nach Durchführung von Maßnahmen zur Selbstreflexion aufgefordert wurden. In der Bedingung RÜ erhielten die Vpn in den ersten fünf Monaten eine Erfolgsmeldung unabhängig von den erzielten Resultaten. Folgende drei Hypothesen liegen vor:

- (1) "Die Fähigkeit zum Bearbeiten komplexer, intransparenter Probleme läßt sich erfassen, indem man lediglich das Verhalten zu Beginn der Bearbeitung näher analysiert."
- (2) "Rückmeldung über eine gute Problemlöseleistung sollte besonders bei mißerfolgsmotivierten Personen dazu führen, daß sie Rückschläge bei der Systemsteuerung weniger als persönliche Mißerfolge erleben und deshalb ebenso wie erfolgsmotivierte Personen weiter nach neuen Lösungsmöglichkeiten suchen."
- (3) "Die systematische Aufforderung zur Selbstreflexion sollte bei Personen dazu führen, daß sie ihr Verhalten wirksamer und flexibler steuern und Mißerfolge zum Anlaß nehmen, ihr Verhalten zu ändern." (alle Zitate aus PUTZ-OSTERLOH 1983b, p.106).

Zur Prüfung von (1) wird eine Reihe von Indikatoren aus dem ersten Drittel der Problembearbeitung mit den Daten aus den beiden restlichen Dritteln korreliert. Das Muster der Daten ist nur an einer Stelle einheitlich: Die durchschnittliche Summe "systemangepaßter Maßnahmen" im ersten Drittel steht mit dem Problemlöseerfolg des zweiten und dritten Drittels in allen Gruppen in positivem Zusammenhang, während die Variable "Erfolg im ersten Drittel" mit der Variable "Erfolg in der Restzeit" nur bei Gruppe RÜ signifikant korreliert, dies allerdings mit dem höchsten r (= .59) aller vorgelegten Korrelationen. Diese Inkongruenz der Befunde von "Erfolg" einerseits, "systemangepaßten Maßnahmen" andererseits läßt verwundern: Sollten nicht systemangepaßte Maßnahmen auch per definitionem zum Erfolg führen? Dazu später. - Trotz dieses eindrucksvollen Zusammenhangs zwischen "Erfolg im ersten Drittel" und "Erfolg in der Restzeit" in der RÜ-Gruppe schreibt die Autorin:

"Nicht der erzielte Erfolg im ersten Abschnitt der Problembearbeitung, sondern die dabei eingesetzten Problemlöseprozesse (damit meint sie die systemangepaßten Maßnahmen, J.F.) sind die konsistente Basis für die Güte künftiger Leistungen." (PUTZ-OSTERLOH 1983b, p.109).

Was hat es mit diesen "systemangepaßten Maßnahmen" auf sich? Es handelt sich dabei um Bewertungen der von der Vp getroffenen Entscheidungen, wobei nach den in Tab. 4.6 beschriebenen Regeln Punkte vergeben wurden (PUTZ-OSTERLOH, schriftl. Mitteilung).

Wie man sieht, werden die Maßnahmen weitgehend kontextlos und entsprechend pauschal beurteilt. Die in (1), (6) und (9) vorgenommene positive Bewertung einer Erhöhung der Variablen EH, LO und SM entspricht nicht dem vernünftigen Vorgehen, z.B. Rohmaterial dann in Mengen einzukaufen, wenn es preiswert ist oder den Lohn auf den Optimalwert von 1510 zu setzen bei gleichzeitiger Senkung von SM. Bewertung (2) ist schlicht falsch: sie entspricht zwar der alltäglichen Meinung, aber wie Abb. 4.3 demonstriert, gibt es nur einen Wert von PH, der optimal ist; alles übrige bringt Verluste und ist daher unangemessen (und zwar umso mehr, je weiter sich PH vom Optimum entfernt,

Tab. 4.6: Bewertung von Maßnahmen in der Studie von PUTZ-OSTERLOH (1983b; Angaben laut schriftl. Mitteilung; '+' = Erhöhung, '-' = Senkung).

Nr.	Variable ¹	Punkt	Bedingung
(1)	EH ²	1 0	wenn EH > 0 wenn EH = 0
(2)	PH	1 0	wenn + und HL leer, wenn - und HL voll wenn - und HL leer, wenn + und HL voll
(3)	WE	1 0	wenn + und HL voll und WE < 6000 wenn - und HL leer und WE < 6000 wenn + und HL leer wenn - und HL voll
(4)	A1,N1;A2,N2	1 0	wenn Gleichgewicht wenn Ungleichgewicht
(5)	RS	1 0	wenn - und MA = MM, wenn + und MA < MM wenn + und MA = MM, wenn - und MA < MM
(6)	LO	1 0	wenn + wenn -
(7)	GL	1 0	wenn City wenn nicht City
(8)	LW	1 0	wenn + und HL voll wenn -
(9)	SM	1 0	wenn + wenn -

¹zur Bedeutung der Kürzel vgl. Tab. 4.1

²EH=eingekauftes Rohmaterial

nach oben wie nach unten). Die Bewertung (3) legt eine falsche Konstante bei der Begrenzung des oberen Werts von WE zugrunde, nämlich 6000 anstatt 4500 (vgl. Kap. 4.1). Die Bewertungen (4), (5) und (7) sind akzeptabel. Bewertung (8) versucht der merkwürdigen Relation LW-NA Rechnung zu tragen, obwohl nicht geklärt ist, inwiefern die Investitionskosten in Höhe von 10000 Einheiten durch den um ca. 100 Hemden erhöhten Verkauf gedeckt werden; hierzu ist die Berücksichtigung der Laufzeit wie auch der laufenden Kosten von 500 Einheiten pro LW notwendig.

Was hier also erfaßt wurde, ist unklar; jedenfalls kann es sich nicht ausschließlich um "systemangepaßte Maßnahmen" handeln, die hierbei als "gut" bewertet werden. Zu einem guten Teil reflektiert das Maß dasjenige Eingriffsverhalten, das nach Ansicht des Bewerter "gut" wäre, aber nach Systemkriterien zu schlechten Resultaten führt. Die Interpretation der "konsistenten Basis für die Güte künftiger Leistungen" (siehe oben) sollte auch deshalb mit dem entsprechenden Vorbehalt vorgenommen werden.

Zur Prüfung der zweiten und dritten Hypothese führte die Autorin einfaktorische Varianzanalysen mit der UV "Gruppenzugehörigkeit" und der AV "Problemlöseerfolg" (= TRENDPO) sowohl insgesamt wie auch nur für das erste Drittel durch, die jedoch keinen signifikanten Effekt nachwies. Bei einem Alpha von 0.05, dem gewählten $n = 30$ und einem $u = 3-1 = 2$ liegt -so meine aposteriori-Berechnungen- nach COHEN (1977², Table 8.3.13, p.313) bei einem angenommenen $f = .40$ ("large effect") eine Power von 0.93 (also ein Beta von 0.07) für diesen Test vor; bei einem angenommenen "medium effect" ($f = .25$) beträgt Beta allerdings schon 0.45. Da die Autorin keine Angabe zur vermuteten Effektstärke macht, könnte somit auch ein $f = .10$ ("small effect") gemeint sein mit einem Beta von 0.88. Unabhängig davon stellt sich erneut die Frage, ob nicht sinnvoller zwei t-Tests hätten durchgeführt werden sollen, nämlich jeweils der Vergleich der KG mit einer der Experimentalgruppen RÜ bzw. RE. Doch anstelle derartiger Überlegungen findet man in der Arbeit 107 Korrelationskoeffizienten (davon 46 als signifikant ausgewiesen), wobei die "detaillierte Darstellung der Einzelergebnisse" zu "Motiv"-Kennwerten schon mangels Signifikanz -es war nicht möglich, "irgendwelche Systematiken" (p. 111) zu erkennen- unterblieb!

Die Untersuchung von FUNKE (1983b) sollte dazu dienen, die behauptete Unwirksamkeit der Variablen "Testintelligenz" auf die Bearbeitung des "TAILORSHOP" nachzuprüfen. Zu diesem Zweck wurde ein zweifaktorieller Versuchsplan mit den Faktoren "Testintelligenz" (Stufen: niedrig/hoch) und "Transparenz" (Stufen: Variablenvernetzung verfügbar/nicht verfügbar) realisiert, mit dem Hypothesen über Haupteffekte und die Wechselwirkung überprüft werden konnten. Als abhängige Variable dienten die beiden in Kap. 4.1 und 4.2 beschriebenen Trendmaße TRENDPO und TRENDFU. Die Durchführung der Untersuchung ist im Detail bei FUNKE (1983b) nachzulesen; wichtig ist, daß die Vpn nach einer zweimonatigen "Probezeit", die der Instruktion diente, selbständig mit dem System interagieren konnten, ein V1-Einfluß wie in bisherigen Studien also weniger wirksam werden konnte. Die möglichen Eingriffsbereiche präsentierte das Programm der Vp in einer fixen Abfolge. Nach Abschluß aller Eingriffe erschien der für diesen Simulationstakt resultierende Zustand auf dem Monitor als Grundlage der Planung kommender Eingriffe.

Die Datenanalyse lieferte für die AV TRENDPO keinerlei bedeutsame Effekte. Bei Verwendung der AV TRENDFU mußte dagegen die Annahme eines ausbleibenden Intelligenzeffekts verworfen werden: Es zeigt sich, daß die Gruppe mit den höheren RAVEN-Scores auch bessere Simulationsergebnisse erzielt, vorausgesetzt man verwendet den Indikator TRENDFU.

Die Zweifel an der Brauchbarkeit des Indikators TRENDPO erhöhen sich durch die Ergebnisse einer kleinen Test-Retest-Studie, die von FUNKE (1983b) beschrieben wird: Für TRENDPO ergibt sich über ein 14-Tages-Intervall eine Test-Retest-Rangkorrelation von 0.20, für TRENDFU von 0.80! So problematisch Retest-Reliabilitäten bei der wiederholten Bearbeitung komplexer Probleme sein mögen, so unwahrscheinlich ist es, daß die Rangfolgen wiederholter Messungen zu Null korrelieren. Sollte

letzteres aber wirklich der Fall sein, wäre ein derartiger Indikator wenig geeignet, fehlende Zusammenhänge zwischen Intelligenztestleistungen und Problemlösefähigkeiten zu demonstrieren.

4.4 Zur Kontroverse um die Interpretation der Ergebnisse

Durch die eben beschriebenen Befunde von FUNKE (1983b) angeregt, schrieb PUTZ-OSTERLOH (1983a) eine Replik, auf die wiederum eine Stellungnahme von FUNKE (1984) folgte. Gegenstand der Diskussion war die Frage, ob die von mir gefundenen Ergebnisse tatsächlich Zweifel an den bisher vorgelegten Resultaten begründen. PUTZ-OSTERLOH greift folgende wichtige Punkte auf:

- (1) die Veränderung der Untersuchungssituation;
- (2) die Operationalisierung von Lösungsgüte;
- (3) die Transparenz der Problemstellung.

Diese Aspekte wollen wir nachfolgend erläutern und unsere diesbezüglichen Standpunkte darstellen.

Bezüglich der Veränderung der Untersuchungssituation im Vergleich zu den Studien von PUTZ-OSTERLOH (1981) und PUTZ-OSTERLOH & LÜER (1981) gibt es - und hier hat PUTZ-OSTERLOH (1983a) recht mit ihrem Hinweis - tatsächlich einige diskussionsfähige Argumente:

- (a) Die Notwendigkeit des aktiven Erfragens von Informationen über einen V1 gibt es bei FUNKE nicht: Die Vpn erhalten die monatlichen Resultate ihres Managements auf dem Monitor des Rechners präsentiert.
- (b) Die Vpn brauchen sich nicht aktiv für bestimmte Maßnahmen zu entscheiden, sondern werden hinsichtlich jedes möglichen Eingriffsbereichs (vom Computer) gefragt, ob sie eingreifen möchten oder nicht. Damit sei - so PUTZ-OSTERLOH (1983a, p.307) ein "Vergessen" von Entscheidungen genausowenig möglich wie der Wunsch nach nicht realisierbaren Maßnahmen.
- (c) Die Vpn erhalten bei FUNKE zusätzliche Informationen, die das System "transparenter" machen.

Alle drei Veränderungen laufen - so PUTZ-OSTERLOH (1983a) - darauf hinaus, die Untersuchungssituation stärker an die Intelligenztestsituation anzunähern; die von FUNKE (1983b) gefundene Korrelation zwischen PL und IQ von 0.28 läge nahe an der Korrelation, die von PUTZ-OSTERLOH unter Transparenz gefunden wurde.

Dieser Argumentation könnten und würden wir uns anschließen unter der Voraussetzung, daß Einigkeit über den Terminus "Transparenz" besteht; da dies jedoch nicht der Fall ist (siehe unten zu Punkt 3), bleiben obige Aussagen zwar richtig, die Schlußfolgerungen sind jedoch in bezug auf die Transparenz der Situation arbiträr und nicht entscheidbar.

Die Diskussion um die angemessene Operationalisierung der Lösungsgüte wird durch die Bemerkungen PUTZ-OSTERLOH's (1983a) nicht bereichert: Sie macht darauf aufmerksam, bereits seit zehn Jahren von der Nichtoptimierbarkeit des Systems "TANALAND" zu wissen. Dem liegt offenbar der Trugschluß zugrunde, was an "TANALAND" erkannt worden sei, ließe sich auf andere, ähnliche Systeme generalisieren. Ein Trugschluß ist dies insofern, als unter Umständen die Änderung von ein oder zwei Gleichungen eines (nicht-)linearen Gleichungssystems oder die Änderung von Beschränkungen aus einem vordem nicht-optimierbaren ein optimierbares System macht. Daß auch im Fall nicht-optimierbarer Systeme eine zumindest relative Bewertung der Systemhandhabung - relativ im Vergleich von Personen einer Stichprobe - möglich ist, soll nicht bestritten werden; mir scheint es jedoch nach wie vor (vgl. FUNKE 1981) wünschenswert, unabhängig von Vpn-Daten (apriori) etwas über die Qualität der Problemlösung zu sagen. THALMAIERS (1979) Arbeit über die "MONDLANDUNG" ist m.E. ein gutes Beispiel für einen derartigen Ansatz. Voraussetzung hierfür ist allerdings eine klare Zielvorgabe an die Vp.

Unter diesen zweiten Diskussionspunkt - Gütekriterien - fällt auch die Frage nach Reliabilität und Validität der Gütemaße. Das sowohl im "TAILORSHOP" wie auch in "LOHHAUSEN" angewendete Verfahren der Beurteilung von Systemdatenverläufen durch "Experten" (d.h. durch die Versuchsleiter) muß einer genaueren Überprüfung unterzogen werden. Es geht also darum, ob man aus den hohen Übereinstimmungskoeffizienten verschiedener Beurteiler und deren Korrelation mit einem "objektiven" Indikator auf die Validität dieses Indikators schließen darf.

Zunächst die Resultate einer kleinen Studie, in der mit einem künstlichen Datensatz von 48 "Vpn" operiert wurde (Ausführlicheres hierzu wurde bereits in Kap. 3.8 berichtet). Jeder Datensatz bestand aus einer Sequenz von 10 geglätteten Zufallszahlen, die so etwas wie den Erfolg der Vp beschreiben sollte. Wir können nunmehr fragen, ob Rater die hier per Zufall erzeugten "Kurven" treffsicher beurteilen und wie die Korrelation zu einem "objektiven" Gütekriterium (nehmen wir dafür die Zahl der Einheiten mit Aufwärtstrend) ausfällt. Aufgrund des erzeugenden Algorithmus wissen wir, daß zwei aufeinanderfolgende Meßwerte absolut unabhängig voneinander sind, die Reliabilität eines summarischen Trendmaßes also Null beträgt.

Vier "Experten" (Spieler des "TAILORSHOP") und zwei "Naive" beurteilten die "Kapital"kurven hinsichtlich ihrer Güte auf einer fünfstufigen Skala. Die Auswertung der erhaltenen Daten erfolgte mittels des Übereinstimmungskoeffizienten r_{cc} von HORST (1949; zit. nach SIXTL 1982², p.149f) ganz analog zu der bei PUTZ-OSTERLOH (1981) beschriebenen Prozedur. Das r_{cc} der sechs Beurteiler beträgt 0.941, die Korrelation der durchschnittlichen Beurteilung (schulnotenskaliert) zum Trendmaß beträgt -0.685, d.h. niedrige und damit gute Noten indizieren einen guten hohen Trendwert. Darf man aus diesen Daten folgern, einen guten Indikator für das Systemverhalten gefunden zu haben? Was bringt das vollkommen "zufällige" Schwanken der Rohdaten zum Ausdruck? Ich meine, daß als einziger Schluß aus diesen Daten folgt, daß Rater den

Trendindikator reliabel schätzen, keineswegs aber, daß es sich hierbei um einen reliablen oder gar validen Indikator des Systemverhaltens handelt! Aus der hohen Interraterkorrelation und der Korrelation mit dem "Außen"-Kriterium folgt nicht die geringste Aussage über die Gültigkeit des gewählten Kriteriums für eine angemessene Systembeschreibung.

Eine Validitätsprüfung kann nur über andere Wege erfolgen (vgl. HAND & SIMS 1975, RUBEN & LEDERMAN 1982), sollte dabei auch die "Fällen" der klassischen Testtheorie im Auge haben (z.B. bei Verletzungen der Annahme unkorrelierter Fehler; vgl. ZIMMERMAN & WILLIAMS 1977). STERNBERG's (1982, p.229) Postulate der Quantifizierbarkeit, Reliabilität, Konstruktvalidität und empirischen Validität bei Intelligenztestaufgaben sollten auch für die Meßinstrumente im Bereich komplexen Problemlösens gelten. Die "augenscheinliche" Validität mancher Güteindikatoren muß gründlicher belegt sein, bevor weitreichende Schlußfolgerungen daraus gezogen werden. Vor allem sollte man nicht über Validität reden, solange "das Problem der Reliabilitätsüberprüfung des Gütekriteriums noch nicht eindeutig geklärt" (PUTZ-OSTERLOH 1983a) ist. Wenn Vpn von Mal zu Mal andere Strategien verwenden (wie PUTZ-OSTERLOH vermutet) und damit andere Gütewerte erzielen, scheint das Konzept eines traits "guter Problemlöser" doch äußerst problematisch.

Der dritte und letzte Diskussionspunkt betrifft das Konzept der Transparenz einer Problemstellung: Dazu schreibt PUTZ-OSTERLOH (1981, p.84):

"... müssen die Pbn unter der intransparenten Bedingung den Zustand der Systemvariablen ... erfragen. Dabei können sie allerdings aus dem Katalog der möglichen Maßnahmen alle Variablen, über deren Zustand sie Informationen benötigen, ableiten."

Wesentliches Merkmal ist hier die aktive Informationsbeschaffung. Ähnlich PUTZ-OSTERLOH & LUER (1981, p.312):

"Als intransparent gelten ... Probleme, deren Variablen erst durch aktives Eingreifen des Problemlösers für ihn erkennbar und analysierbar sind, da ihm zu Beginn nicht genügend Informationen über die Variablen vorliegen."

In der Kurzform:

"Ein Problem ist für einen Problemlöser dann intransparent, wenn ihm nicht alle relevanten Daten vorgegeben sind." (PUTZ-OSTERLOH 1983b, p.101).

Diese Formulierungen machen die Notwendigkeit einer Präzisierung deutlich, denn ab welchem Zeitpunkt will man etwa Transparenz in bezug auf die für den "TAILORSHOP" wichtige Relation zwischen Werbung und Nachfrage gelten lassen:

(a) wenn die Vp global erfährt, Werbung steigere die Nachfrage,

- (b) wenn die Vp die genaue Relation erfährt, wieviel Nachfrage durch Erhöhung der Werbungskosten um eine Einheit zusätzlich erzeugt wird,
- (c) wenn die Vp darüberhinaus den Grenzwert erfährt, von dem ab die Relation nicht mehr gilt?

Das Beispiel macht deutlich, daß es offenkundig ein Kontinuum der Transparenz und damit des zur Problemlösung notwendigen Wissens gibt. Dieses Kontinuum kann nun nach mehreren Ebenen aufgebrochen werden. Transparenz eines Systems kann - zu einem gegebenen Zeitpunkt - hinsichtlich folgender Aspekte bestehen: (a) Zielkriterium bzw. -kriterien, (b) Variablenzahl, (c) Variablenzustand und (d) Variablenverknüpfung. Intransparenz bezüglich eines oder mehrerer dieser Aspekte kann zudem temporärer oder permanenter Natur sein: im ersten Fall besteht die Möglichkeit, während des Experiments die notwendigen Informationen einzuholen, im zweiten Fall werden bestimmte Systemdetails dem Probanden gegenüber auf Dauer geheimgehalten. Neben dem objektiven Grad an Transparenz (wie weit also der V1 seine Karten aufdeckt) gilt es, auch den subjektiven Grad an Transparenz (also die Nutzung der objektiv verfügbaren Daten) zu berücksichtigen. Diesen sollte man der Präzision wegen besser als "Grad des Systemwissens" bezeichnen und die Bezeichnung "Transparenz" für objektive Charakteristika der Problemsituation reservieren.

Daß die so bestimmte Transparenz nicht ausschließlich durch eine Vp-V1-Interaktion vermittelbar ist, wie sie von PUTZ-OSTERLOH unter Hinweis auf einen "Bewegungsfreiraum für die Vp" gefordert wird, dürfte klar sein: Auch ohne sozialpsychologische Störgrößen kann die Vp interaktiv von einem Rechner Informationen erfragen und sich durch ein Datennetz "hangeln", ähnlich der Art, wie sie sich vom V1 Auskünfte erbittet. Studien über Informationssammlung bei komplexen Problemen haben übrigens schon GLASER, DAMRIN & GARDNER (1954) mit ihrer "tab item"-Methode durchgeführt (siehe auch RIMOLDI 1960, McGUIRE & BABBOTT 1967, KARLINS et al. 1967). Auch in jüngster Zeit gibt es Arbeiten, in denen die Diagnose von Fehlern in komplexen Systemen wie Schiffen, Flugzeugen oder Atomkraftwerken untersucht wird (z.B. RASMUSSEN & ROUSE 1981). Die Kommunikation mit einem V1 ist dort in der Regel ersetzt durch die (realistische) Kommunikation mit dem System selbst. Dies scheint mir den gleichen Bewegungsfreiraum für die Vp zu gewähren, ohne daß V1-Effekte massiv wirksam werden könnten wie beim bisherigen Vorgehen, wo gleichzeitig dem V1 großer Spielraum für sein Verhalten gelassen wird (sog. Effekt des "laxen" Vorgehens; nach BREDEKAMP 1980, p. 41).

Vergleicht man einmal einige Studien miteinander, in denen die "Schneiderwerkstatt" eingesetzt wurde, ergibt sich hinsichtlich der vier genannten Ebenen von Transparenz folgendes Bild, das Tab. 4.7 veranschaulicht.

Tab. 4.7: Vergleich der Studien von PUTZ-OSTERLOH (1981, =PO), PUTZ-OSTERLOH & LÜER (1981, =PO&L) und FUNKE (1983b, =FU) hinsichtlich der Transparenzmerkmale unter zwei verschiedenen experimentellen Bedingungen; zur Ergänzung die Studien von PUTZ-OSTERLOH (1983b) und DÖRNER et al. (1983).

Transparenz hinsichtlich	"Transparenz"			"Intransparenz"			PO 1983b	DÖ et al.
	PO	PO&L	FU	PO	PO&L	FU		
a) Ziel- kriterium	+ ^a	+	0	+	+	0	+	--
b) Variablen- anzahl	+	+	++	0 ^b	0 ^b	++	0 ^b	-
c) Variablen- zustand	+	+	++	-	-	++	-	- ^c
d) Verknüp- fung	+	+	+	--	--	--	--	--

^a --=sehr niedrig, -=niedrig, 0=mittel, +=hoch, ++=sehr hoch

^b es werden alle Eingriffsmöglichkeiten genannt, Katalog möglicher Maßnahmen liegt vor

^c automatische Mitteilung über Zustand der "Alarmmatrix" (vgl. DÖRNER et al. 1983, p.130)

Direkt vergleichbar sind die Studien von PUTZ-OSTERLOH & LÜER (1981) und FUNKE (1983b), da dort jeweils Transparenz- und Intransparenzbedingungen realisiert wurden. Bei PUTZ-OSTERLOH (1983b) wie bei DÖRNER et al. (1983) lagen jeweils "intransparente" Bedingungen vor. - Die in Tab. 4.7 eingetragenen Ausprägungsgrade beruhen auf Einschätzungen, da in den genannten Publikationen nicht alle Details aufgeführt sind. Deutlich sichtbar wird der Unterschied zwischen FUNKE (1983b) und den übrigen Arbeiten auf der Ebene "Intransparenz": die sehr hohe Transparenz bezüglich des Variablenzustands bei FUNKE durch das monatliche Display der Spielresultate zeigt die Richtigkeit des Arguments von PUTZ-OSTERLOH (1983a), soweit es diesen Aspekt der Transparenz betrifft. Bezüglich der Intransparenz der Variablenverknüpfung jedoch zeigen sich keine Unterschiede bei den fünf genannten Studien. FUNKEs Arbeit prüft somit den reinen Effekt der Verknüpfungskenntnis, die beiden anderen Studien dagegen konfundieren die Aspekte (c) und (d).

Transparenz bzw. Intransparenz beeinflussen direkt das Ausmaß der Problemschwierigkeit. Hier hat HUSSY (1984b) versucht, durch eine hierarchische Konzeption des Schwierigkeitskonzepts Klarheit in die Zusammenhänge zu bringen. Abb. 4.4 verdeutlicht seine Annahmen.

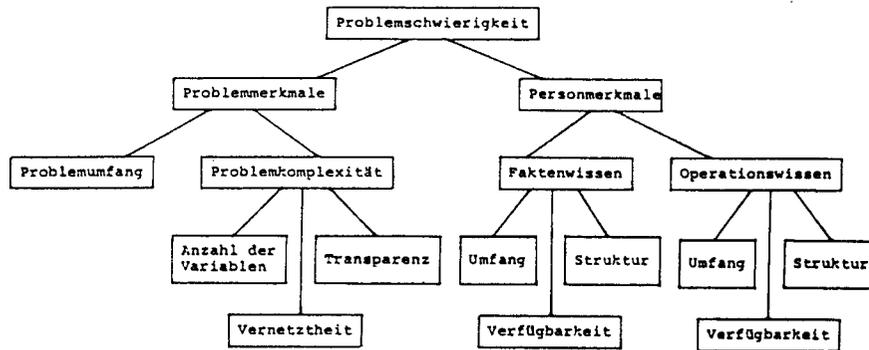


Abb. 4.4: Analyse des Begriffs "Problemschwierigkeit" nach HUSSY (1984b, p.4).

HUSSY unterteilt den Begriff der Problemschwierigkeit in die zwei Bereiche "Problemmerkmale" und "Personmerkmale", die jeweils weiter differenziert werden. Auf seiten des Problems sind es Umfang (= Zahl der Lösungsalternativen) und Komplexität (letztere unterteilt in Variablenzahl, Vernetztheit und Transparenz), auf seiten der Person werden Fakten- und Operationswissen unterschieden jeweils nach den Kategorien von Umfang, Struktur und Verfügbarkeit. Dieses Schema ist insofern problematisch, als natürlich die aufgezeigten Komponenten nicht unabhängig voneinander variieren können und darüberhinaus erst bei klarer Definition Sinn machen; HUSSY weist explizit darauf hin, "daß diese Determinanten in vielfältiger Weise miteinander verknüpft sind" (1984a, p.123). Für Paradigmen der komplexen Problemlöseforschung läßt sich die Problemmplexität als Funktion der vier Determinanten Variablenzahl, Variablenvernetzung, Transparenz und Eigendynamik darstellen, die Problemschwierigkeit auf seiten der Problemmerkmale als Funktion von Problemmplexität, Problemumfang und Barriertyp (vgl. HUSSY 1984a, p.228). Meiner Meinung nach werden diese begrifflichen Bemühungen - so wertvoll sie sind, um den Gegenstandsbe- reich zu markieren - jedoch solange scheitern, solange keine präzi- sieren Definitionen zu den einzelnen Komponenten vorgelegt werden (vgl. dazu die Ansätze in Teil III dieser Arbeit). Auch der Hinweis auf die "vielfältigen" Verknüpfungen muß durch genauere Beschreibung der funk- tionalen Beziehungen ersetzt werden; diese wiederum können nur Ausfluß einer Theorie kognitiver Prozesse sein, deren Fehlen bereits in Kap. 3.1 konstatiert wurde.

Daher zum Schluß ein Wort zur kritisierten "Theoriearmut": PUTZ- OSTERLOH verweist auf DÖRNER's Theorie, aus der sich "sehr wohl Hypo- thesen für empirische Untersuchungen ableiten und diese auch bestä- tigen lassen" (1983a, Hervorhebung von mir). Den heuristischen Wert wollen wir nicht bestreiten, doch läßt die Präzision der Prognosen noch insofern viel zu wünschen übrig, als ihre Falsifikation vielfach nicht möglich ist. Hinzu kommt die Tendenz, nichtsignifikante Befunde erst gar nicht oder nur unvollständig mitzuteilen, selbst wenn sie sich auf explizit formulierte Hypothesen beziehen. HERRMANN, der sich

über den starken Einfluß nur schlecht überprüfbarer Theorien in der Kognitionspsychologie wundert (1982), warnt davor: "Solange man lediglich 'Fakten' sammelt, kann man für eine Weile meinen, irgendwann springe aus den Fakten schon von ganz allein die Wahrheit, der Zugang zur wahren Wirklichkeit, heraus." (HERRMANN 1983, p.98). Über die gelegentliche "Nützlichkeit der Fiktionen" wollen jedoch auch wir nicht hinwegsehen, die "Wende" allerdings ist von dort nicht zu erwarten.

Teil III: Weiterführende Perspektiven für eine Theorie des Umgangs mit dynamischen Systemen

5 Problemlösen als Konstruktion von Kausalmodellen

Nachdem im vorangegangenen Kapitel anhand eines konkreten Beispiels Probleme aufgezeigt wurden, die das Verständnis der in Kap. 3 formulierten kritischen Thesen vertiefen sollten, geht es in diesem Kapitel um die eigenen Vorstellungen über die Fortentwicklung der denkpsychologischen Forschungsbemühungen. Ausgangspunkt der Überlegungen wird sein, über eine genaue Beschreibung des Aufgabenmaterials (1) die für menschliche Steuerungsabsichten wesentlichen objektiven Wirkgrößen eines dynamischen Systems zu identifizieren und damit (2) zu einer Analyse derjenigen kognitiven Operationen zu gelangen, die für den Umgang mit derartigen dynamischen Systemen bedeutsam sind.

Das formale Inventar, das eine Beschreibung von Wirkungsgefügen im Sinne der Formulierung kausaler Theorien ermöglicht, wird von uns herangezogen werden, um die Probleme, die von Probanden in denkpsychologischen Untersuchungssituationen bearbeitet werden, präzise beschreiben zu können. Über diesen Weg der exakten Aufgabenbeschreibung und über die Beschreibung des Wegs, den der Problemlöser beschreitet, wenn er ein ihm unbekanntes Wirkungsgefüge exploriert, soll zugleich ein Hilfsmittel zur Hypothesenbildung auf seiten des Forschers bereitgestellt werden, das der konstatierten "Theoriearmut" abhelfen könnte. Einen Überblick über den Forschungsstand hinsichtlich des Konzepts kausaler Abhängigkeit findet man bei STEYER (1983c, p.61f), der darauf hinweist, daß die beiden Hauptströmungen - experimentalpsychologischer versus pfadanalytischer Ansatz - bislang Lücken in dem Unternehmen aufweisen, "eine von diesen Kontrolltechniken (gemeint sind Randomisierung, Konstanthaltung bzw. Parallelisierung, J.F.) zunächst unabhängige Definition kausaler Abhängigkeit" zu geben.

Zunächst wird in Kap. 5.1 eine Terminologie vorgestellt, die der formalen Beschreibung stochastischer Mechanismen dient. Dabei interessiert insbesondere die Tauglichkeit derartiger formaler Beschreibungsmittel als Beschreibungen von Eigenschaften solcher Probleme, wie sie in der neueren Denkpsychologie als Untersuchungsinstrumente eingesetzt werden. Daran schließt sich die Frage an, welche Parallelen zwischen den herkömmlichen Beschreibungsmitteln der Denkpsychologie und den hier eingeführten Konzepten bestehen (Kap. 5.2). Mit der Darstellung einer Fallstudie in Kap. 5.3 sowie der exemplarischen Anwendung des neuen Modells auf bekannte Untersuchungsinstrumente in Kap. 5.4 wird der Nutzen dieser Beschreibungsform dokumentiert. Schließlich wird der Grundstein für eine Theorie der Konstruktion kausaler Modelle in dynamischen Prozessen gelegt (Kap. 5.5). Sieben Annahmen bilden den Kern dieser Theorie, die den konstruktiven Weitergang unserer Arbeiten markiert. Abschließende Bemerkungen in Kap. 5.6 befassen sich mit parallelen Modellvorstellungen anderer Autoren.

5.1 Ein Begriffsinventar für Kausalmodelle in der Denkpsychologie

Spricht man in Anlehnung an ZWICKER (1981, p.18) dann von einem System, "wenn darunter ein durch Beobachtungen aufweisbarer Zusammenhang verstanden wird", kann man ein Modell dieses Systems erstellen, um etwa noch unbekannte Eigenschaften des Systems zu erkennen. Solche Modelle können verbal (z.B. sprachliche Umschreibung), ikonisch (z.B. Landkarte) oder symbolisch ausfallen; unter den symbolischen sind es vor allem die mathematischen Modelle, die hier interessieren. Die Absicht des Systemanalytikers ist dabei die "Entwicklung und Gewinnung von Informationen, welche dem Systemuntersucher bislang nicht bekannt waren, durch geeignete Methoden aber aus dem Systemmodell erschlossen werden können." (ZWICKER 1981, p.20). Kennzeichnend für dynamische Modelle, mit denen wir uns nachfolgend ausschließlich beschäftigen, sind nach ZWICKER (1981, p.22f) drei Merkmale, nämlich

- (1) vollsymbolisierte Repräsentation der abzubildenden Zusammenhänge,
- (2) Kennzeichnung der Ereignisse oder Zustände des Modells durch einen Zeitindex, und
- (3) Vorliegen mindestens einer zeitinvarianten Verknüpfung zweier zeitlich gegeneinander verzögerter Ereignisse.

Alle drei Merkmale müssen erfüllt sein, wenn die Bezeichnung "dynamisches Modell" gelten soll. Die dritte Forderung bedeutet übrigens nichts anderes als die nach dem Vorliegen mindestens einer "Wenn-Dann"-Beziehung, bei der die "Dann"-Komponente in einem gleichbleibenden zeitlichen Abstand von der "Wenn"-Komponente eintritt.

Die Verwendung der Ausdrücke "Wenn" und "Dann" darf nicht fehlinterpretiert werden: natürlich wird damit eine kausale Beziehung angesprochen, aber keineswegs als solche definiert. Eine Definition von Kausalität darf sich ja nicht allein auf die im Modell repräsentierten Variablen beziehen, soll also nicht eine Geschlossenheitsannahme machen, sondern muß alle potentiellen Störvariablen einbeziehen, also auch solche, die in einem Modell über ein System nicht explizit enthalten sind (vgl. STEYER 1983c, p.63). Doch dies ist für den hiesigen Kontext weniger bedeutsam.

Strukturgleichungsmodelle (vgl. DUNCAN 1975; GOLDBERGER 1964, 1971; JÖRESKOG 1978; u.v.a.) werden ebenso zur Beschreibung kausaler Zusammenhänge verwendet wie dies in der Theorie latenter Variablen (vgl. BENTLER 1980; HAMERLE 1982; LAZARSFELD 1959; u.v.a.) der Fall ist. Eine systematischere und stärker formale Behandlung kausaler Abhängigkeit zwischen exogenen und endogenen Variablen gibt STEYER (1982b), der derartige Abhängigkeiten in Form multivariater autoregressiver Prozesse (AR-Prozesse) ausdrückt. Die Theorie der AR-Prozesse, auf die wir ausführlicher eingehen, hat vielfältige Anwendungsmöglichkeiten:

"Sie reichen vom klassischen Experiment über Modelle latenter Variablen (Theorie psychologischer Tests) bis zu komplizierten kausalen Modellen mit beobachtbaren und mit latenten Variablen.

Auch zur Beschreibung und Prognose von Entwicklungsverläufen können sie dienen, ebenso wie zur Analyse von Zeitreihen, dem Gebiet, in dem sie bisher am meisten verwendet wurden. Alle genannten Modelle können kausale Modelle sein, die nicht nur rein stochastische Zusammenhänge beschreiben sollen, sondern unsere Vorstellung ausdrücken, wie die Daten erzeugt wurden, oder, wie die beobachteten statistischen Zusammenhänge erklärt werden können." (STEYER 1982a, p.2.46).

Ich möchte zur Liste dieser Anwendungsmöglichkeiten eine weitere hinzufügen, nämlich die, den Problemraum des Problemlösers formal mit Hilfe von AR_k -Prozessen zu beschreiben und festzustellen, wie Probanden derartige Prozesse "entschlüsseln".

Grundlage der weiteren Überlegungen bildet eine formale Struktur, die nicht nur die Beziehung zwischen je zwei Variablen X und Y beschreibt, sondern

"die gesamte Untersuchungssituationsklasse oder das gesamte Experiment mit allen darin vorkommenden Ereignissen ... Diese formale Struktur wird mit (Ω, \mathcal{A}, P) bezeichnet und heißt 'Wahrscheinlichkeitsraum'. Dabei ist Ω die Menge der in der betreffenden Untersuchungssituation vorkommenden Elementarereignisse, \mathcal{A} eine Sigmaalgebra daraus bildbarer einfacher und zusammengesetzter Ereignisse, und P ist eine Funktion, die jedem dieser Ereignisse $A \in \mathcal{A}$ eine feste Wahrscheinlichkeit zuordnet" (STEYER 1983c, p.63f).

Die formale Theorie, die ich hier und im weiteren zugrundelege, findet man ausführlich dargestellt in den Arbeiten von STEYER (1982a, b; 1983a, b, c); daher genügt an dieser Stelle eine knappe Schilderung des Ansatzes.

Auf einem gemeinsamen Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) kann ein stochastischer Prozeß definiert werden, der aus einer Familie $\{Z(t)_i, t \in T\}$ stochastischer Vektoren $Z(t)$ besteht. Unter einem stochastischen Vektor (oder Zufallsvektor) versteht man eine \mathcal{A} -meßbare Abbildung der Ereignisse E in die reellen Zahlen R , d.h. der stochastische Vektor - und jede einzelne stochastische Variable - muß eine Verteilung besitzen.

"Setzt man voraus, daß die Indexmenge T diskret und streng konnex geordnet ist, so ist ein solcher stochastischer Prozeß ein ideales Werkzeug zur Formulierung kausaler Theorien. Man kann leicht formulieren, welche Variable in $Z(t_i)$ welche Variable in $Z(t_j)$ beeinflußt. Die Richtung der Beeinflussung ist durch die Geordnetheit der $t_i, t_j \in T$ festgelegt." (STEYER 1982a, p.1.7).

Die Definition zweier Arten von Kausalität, der sog. "causal" bzw. "weak causal linear stochastic dependence", setzt neben der Vorgeordnetheit der X - vor die Y -Variablen die Erfüllung der Invarianz-Bedingung bzw. der Durchschnitts-Bedingung voraus (vgl. STEYER 1983a, b).

Damit wollen wir uns an dieser Stelle nicht weiter beschäftigen; statt dessen soll das Konzept des R -variaten stochastischen Prozesses $[Z(t), t \in T]$ vertieft werden.

Jeder R -variate stochastische Prozeß $[Z(t), t \in T]$ besteht aus R univariaten stochastischen Subprozessen $\{Z_r(t), t \in T\}$, $r \in \{1, 2, \dots, R\}$. So können R verschiedene Variablen zu jedem Punkt $t \in T$ betrachtet werden.

Ein autoregressiver Prozeß erster Ordnung (AR_1 -Prozeß) läßt sich nach STEYER (1982b, p.3) durch folgende drei Eigenschaften charakterisieren:

- (1) Der Set T ist diskret und besteht aus den natürlichen Zahlen (=geordnete Indexmenge).
- (2) Für alle $t \in T$ kann man die Abhängigkeit der R Variablen des Vektors $Z(t)$ von den R Variablen der vorangegangenen Zeitpunkte $Z(t-1)$, $Z(t-2)$, ..., durch multiple Regressionsgleichungen bestimmen.
- (3) In einem AR_1 -Prozeß wird die Abhängigkeit des Vektors $Z(t)$ von den vorangegangenen Vektoren $Z(t-1)$, $Z(t-2)$, ..., ausreichend durch $Z(t-1)$ beschrieben.

Die letzte Eigenschaft kann in der Schreibweise für bedingte Erwartungen (vgl. STEYER 1984) auch so formuliert werden:

$$E[Z(t)|Z(t-1), \dots, Z(0)] \quad \text{fs} = A Z(t-1), \quad (5.1)$$

wobei A eine reellwertige $R \times R$ -Matrix und $E[Z(t)|Z(t-1), \dots, Z(0)]$ die bedingte Erwartung von $Z(t)$ unter $Z(t-1), \dots, Z(0)$ bedeutet; die Bezeichnung "fs=" bedeutet "fast sicher gleich" (vgl. dazu STEYER 1983c, Fußnote 2 auf p.64). Die Matrix A enthält somit die entscheidende Information, welche Variablen in $Z(t)$ wie stark von Variablen in $Z(t-1)$ abhängen.

Natürlich können auch autoregressive Prozesse höherer Ordnung betrachtet werden. Bei einem autoregressiven Prozeß k -ter Ordnung (AR_k -Prozeß)

"handelt es sich im wesentlichen um einen R -variaten stochastischen Prozeß, bei dem die Vorhersage seiner Zustände zu einem Punkt t_i von den letzten k Zuständen bereits optimal ist, und zwar in dem Sinn, daß die Hinzunahme weiter vorgeordneter Zustände die Vorhersage nicht verbessern würde." (STEYER 1982a, p.2.15).

Im Unterschied zum AR_1 -Prozeß werden nunmehr k A -Matrizen benötigt: In einem AR_k -Prozeß gibt es k Matrizen A_k der direkten AR_k -Parameter, wobei $k \in \{1, 2, \dots, K\}$. Einige Eigenschaften solcher AR_k -Prozesse beschreibt STEYER (1982a, b); sie seien daher hier nur kurz angedeutet:

- (1) Jeder R-variate AR_k -Prozeß läßt sich ohne Informationsverlust als $R * k$ -variater AR_1 -Prozeß darstellen.
- (2) Die Frage der Stabilität eines AR_k -Prozesses läßt sich durch die Suche nach einem Gleichgewichtspunkt beantworten (für spezielle Klassen von AR_1 -Prozessen befriedigend geklärt).
- (3) Die Residuen haben bestimmte Eigenschaften (bedingter und unbedingter Erwartungswert Null, Unkorreliertheit mit vorangegangenen Residuen etc.).

Bislang haben wir von den stochastischen Vektoren $Z(t)$ eines AR_k -Prozesses gesprochen, ohne eine weitere Differenzierung der stochastischen Variablen vorzunehmen; dies soll nun nachgeholt werden. Zunächst einmal soll zwischen endogenen und exogenen Teilprozessen unterschieden werden: ein stochastischer Vektor $Z(t)$ enthalte x exogene und y endogene Variablen, die in den Teilvektoren $X(t)$ bzw. $Y(t)$ zusammengefaßt sind. Als exogene oder x -Variablen seien solche bezeichnet, die "nie von anderen Teilprozessen abhängig (sind), sondern höchstens von ihren eigenen vorangegangenen Zuständen" (STEYER 1982a, p.2.36) determiniert werden. Endogene oder y -Variablen sind dagegen immer auch von anderen Teilprozessen abhängig. Die Prozesse werden als deterministisch bezeichnet, wenn dem stochastischen Vektor $Z(t)$ ein Fehlervektor $F(t)$ zugeordnet wird, dessen Elemente den Wert Null annehmen; ansonsten hat man es mit einem stochastischen Prozeß zu tun, der wie folgt beschrieben wird:

$$Z(t+1) = A Z(t) + F(t+1). \quad (5.2)$$

Durch die Unterscheidung endogener und exogener Teilprozesse läßt sich die Matrix A des direkten AR_1 -Prozesses in vier Teilmatrizen zerlegen:

$$A = \begin{bmatrix} A_{xx} & I & A_{xy} \\ \text{-----} & \text{-----} & \text{-----} \\ A_{yx} & I & A_{yy} \end{bmatrix} \quad (5.3)$$

Entsprechend der pfadanalytischen Notationsweise (vgl. OPP & SCHMIDT 1976) bezeichnet das erste Subskript in (5.3) die "abhängigen", das zweite Subskript die unabhängigen Variablen. Die Parameter der Teilmatrix A_{xx} geben somit die Stärke an, mit der die exogenen Teilprozesse von Zeitpunkt zu Zeitpunkt aufrechterhalten werden; nimmt man hier eine Einheitsmatrix an, bleiben einmal gegebene Werte von X -Variablen konstant erhalten (sofern keine weitere Änderung vorgenommen wird), nimmt man eine Nullmatrix an, haben einmal gegebene Werte von X -Variablen eine "Schockwirkung", d.h. sie nehmen im nächsten Zeittakt den Wert Null an, wenn sie nicht gezielt aufrechterhalten werden. Andere als Diagonalmatrizen sind für A_{xx} nicht erlaubt, da gemäß der Definition einer exogenen Variablen diese nur von sich selbst abhängig sein kann.

Die Teilmatrix A_{xy} ist eine Nullmatrix, da gemäß unserem Verständnis von exogenen und endogenen Variablen die exogenen X -Variablen nicht von den endogenen Y -Variablen beeinflusst werden können:

Die Teilmatrix A_{yx} enthält die Parameter der direkten AR_1 -Abhängigkeit der Y - von den X -Variablen; hier sind keine Beschränkungen vorgesehen. Die Teilmatrix A_{yy} gibt schließlich die wechselseitige Abhängigkeit der endogenen Variablen untereinander an, was für die Bestimmung der Stabilität eines Prozesses von entscheidender Bedeutung ist (vgl. STEYER 1982b). Hat man etwa eine spezielle Klasse von AR_1 -Prozessen vor sich mit den Eigenschaften (a) Teilmatrix A_{xx} ist eine Einheitsmatrix sowie (b) Teilmatrix A_{xy} ist eine Nullmatrix (definitionsgemäß) und (c) beliebige Teilmatrizen A_{yx} und A_{yy} , dann ist dieser AR_1 -Prozeß stabil (d.h. besitzt einen Gleichgewichtspunkt), wenn eine der drei folgenden Aussagen zutrifft (vgl. STEYER 1982a, p.2.39):

- (1) $\lim_{n \rightarrow \infty} A_{yy}^n = \text{Nullmatrix}$, $n \in \mathbb{N}$,
- (2) $A_{yy} = \text{Nullmatrix}$,
- (3) $A_{yx} = \text{Nullmatrix}$ und A_{yy} idempotent (d.h. $A_{yy}^2 = A_{yy} = A_{yy}^3$ usw.).

Soviel zu den Teilmatrizen mit den direkten AR_1 -Parametern. - Nun noch einige wichtige Begriffsklärungen. Zunächst einmal sollen statische und dynamische R-variate stochastische Prozesse unterschieden werden. Zu Beginn haben wir drei Merkmale genannt, die nach ZWICKER (1981) kennzeichnend für ein "dynamisches Modell" sind. Daß einem dynamischen Modell (nach ZWICKERs Definition) ein statischer Prozeß zugrundeliegen kann, wird an einem bei STEYER (1982a, p.2.10f) entlehnten Beispiel deutlich.

Man denke sich ein einfaches pfadanalytisches Modell mit einseitiger Abhängigkeit, wie es in Abb. 5.1 veranschaulicht ist.

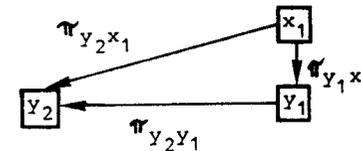


Abb. 5.1: Pfadmodell mit einseitiger Abhängigkeit zwischen drei Variablen (modifiziert nach STEYER 1982a, p.2.10).

Bei dem x_1 aus Abb. 5.1 handele es sich um Schwangerschaftsdauer, bei y_1 um Geburtsgewicht und bei y_2 um die Anzahl von Krankheitstagen bis zum fünften Lebensjahr. Schematisch läßt sich dieses Pfadmodell wie in Abb. 5.2 darstellen.

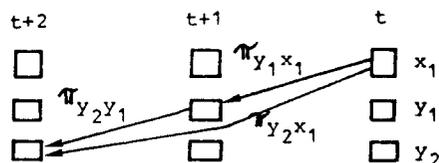


Abb. 5.2: Schematische Darstellung des AR₂-Prozesses aus Abb. 5.1.

Warum ist dieses Pfadmodell nun als statisch einzustufen? Durch die Definition der Variablen ist in diesem Modell festgelegt, daß für jede Variable des stochastischen Prozesses genau ein $t \in T$ existiert, an dem sie von Null unterschieden ist; zu keinem anderen als zu diesem t kann sie also einen Wert größer Null annehmen. Das Geburtsgewicht kann etwa (wie auch die übrigen Variablen) nur zu einem Zeitpunkt t eintreten und dann nicht mehr. Diese Eigenschaft macht das Modell zu einem degenerierten Prozeß, der als statisch bezeichnet wird.

Das in Abb. 5.1 und 5.2 gezeigte Modell wird zu einem dynamischen, wenn man die Variablen so wählt, daß sie zu mehreren Zeitpunkten von Null verschiedene Werte annehmen können (z.B. wenn x_1 Dauer des Arbeitstages, y_1 Menge des Bierkonsums und y_2 Länge der Tiefschlafperiode), dann können etwa täglich Daten anfallen. In einem solchen Fall bezeichnet man das Modell als dynamisch. Wie schon einleitend erwähnt, werden wir uns ausschließlich mit dynamischen Prozessen beschäftigen.

Nun zur Unterscheidung von direkten und totalen Effekten. In dem in Abb. 5.1 bzw. 5.2 gezeigten Beispiel liegt ein AR₂-Prozeß zugrunde, für den man zwei Matrizen der direkten AR₁-Prozesse benötigt, die mit 1π und 2π bezeichnet werden sollen (für einen AR_k-Prozeß braucht man folglich k Matrizen der direkten AR_k-Parameter):

$$1\pi = \begin{bmatrix} 0 & 0 & 0 \\ \pi_{y_1x_1} & 0 & 0 \\ 0 & \pi_{y_2y_1} & 0 \end{bmatrix} \quad 2\pi = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \pi_{y_2x_1} & 0 & 0 \end{bmatrix} \quad (5.4)$$

Will man die Abhängigkeiten, denen y_2 ausgesetzt ist, angemessen beschreiben, muß man die direkten Wirkungen beider Teilprozesse berücksichtigen; den Parameter für den sog. totalen Effekt von x_1 auf y_2 erhält man, indem man wie folgt vorgeht:

$$1B^{(2)} = 1\pi^2 + 2\pi, \quad (5.5)$$

wobei $1B^{(2)}$ die Matrix des totalen Effekts aus allen AR₁-Teilprozessen für den gesamten AR₂-Prozeß darstellt:

$$1B^{(2)} = \begin{bmatrix} 0 & 0 & 0 \\ \pi_{y_1x_1} & 0 & 0 \\ 0 & \pi_{y_2y_1} & 0 \end{bmatrix} * \begin{bmatrix} 0 & 0 & 0 \\ \pi_{y_1x_1} & 0 & 0 \\ 0 & \pi_{y_2y_1} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \pi_{y_2x_1} & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \pi_{y_2y_1} & \pi_{y_1x_1} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \pi_{y_2x_1} & 0 & 0 \end{bmatrix} \quad (5.6)$$

Diese etwas umständliche Schreibweise läßt sich abkürzen, wenn man die Schreibweise in Form bedingter Erwartungen verwendet (vgl. STEYER 1984). Die hier gewählte Präsentationsform macht die allgemeinen Eigenschaften multivariater autoregressiver Prozesse jedoch deutlicher, da die kompletten Parametermatrizen aufgeführt sind und nur für das gewählte, sehr einfache Beispiel eine Menge Leerzellen besitzt. In Kap. 5.4 werden kompliziertere Matrizen vorgestellt. Doch zunächst zu ersten psychologischen Implikationen der hier vorgetragenen Überlegungen.

5.2 Demonstration einiger typischer Formen von autoregressiven Prozessen

In diesem Abschnitt soll versucht werden, an einem einfachen Beispiel mit drei exogenen oder x - und drei endogenen oder y -Variablen Grundtypen der AR-Prozesse in Beziehung zu setzen zu Eigenschaften komplexer Probleme, wie sie etwa von DÖRNER et al. (1983) skizziert wurden. Doch zunächst wird etwas über die Eigenschaften von Operatoren (der "Pfeile" bzw. genauer: der Parameter aus den Matrizen der AR_k-Prozesse) zu sagen sein, die DÖRNER (1976, p.21f) unter den Stichworten (a) Wirkungsbreite, (b) Reversibilität, (c) Größe des Anwendungsbereichs, (d) Wirkungssicherheit und (e) materielle und zeitliche Kosten vorstellt. Im Rahmen des von mir verfolgten Ansatzes lassen sich für einige dieser Aspekte Präzisierungen vornehmen. Anstelle des Begriffs "Operator" wird nachfolgend öfter die Bezeichnung "Interventionsvariable" gewählt; gemeint sind damit die exogenen Systemvariablen.

Abb. 5.3 macht zunächst einmal Unterschiede zwischen Interventionsvariablen in bezug auf deren Wirkungsbreite deutlich: Während der "Schmalbandoperator" (vgl. Abb. 5.3a) genau eine Wirkung nach sich zieht, sind es beim "Breitbandoperator" (vgl. Abb. 5.3b) mehrere Wirkungen, die er verursacht. Dies macht Nebenwirkungsanalysen erforderlich. "Äquivalente Operatoren" (vgl. Abb. 5.3c) sind solche, die sich auf die gleiche Zielvariable richten; hierbei mag die Stärke der Wirkung ein Unterscheidungsmerkmal sein, da z.B. Interventionsvariable x_2 in der doppelten Stärke wie x_1 oder x_3 appliziert werden muß, um die gleiche Wirkung bei y_1 zu erzielen.

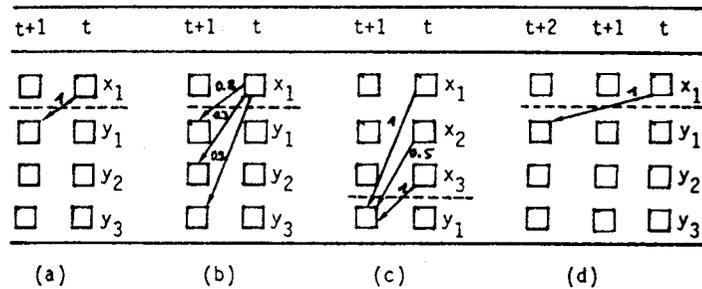


Abb. 5.3: Unterschied zwischen (a) einem "Schmalbandoperator" und (b) einem "Breitbandoperator" sowie (c) "äquivalenten Operatoren"; (d) wie (a), nur zeitverzögert.

Das Konzept der Wirkungsstärke erlaubt es bei Breitbandoperatoren überhaupt erst, Haupt- von Nebenwirkungen zu separieren (eine Unterscheidung, die zwar nicht mathematisch, aber psychologisch sinnvoll ist; nicht zu verwechseln mit der Unterscheidung von direkten und indirekten Wirkungen wie in Kap. 5.1 beschrieben): Im Beispiel (b) ist die Hauptwirkung von x_1 die auf y_1 , während x_1 bei y_2 und y_3 deutlich schwächere Effekte erzeugt, die daher als "Neben"-Wirkungen bezeichnet werden dürfen. Wären alle Gewichte der x_1 -Interventionsvariablen für die unterschiedlichen y -Variablen gleich stark, wäre das Konzept der Haupt- und Nebenwirkung unangemessen; sinnvoller wäre es, dann von "multipler Wirkung" zu sprechen.

Terminologisch erreicht man hier höhere Präzision, wenn man hinsichtlich ihrer Wirkung uni- von multieffektiven Interventionsvariablen unterscheidet, hinsichtlich der Abhängigkeit endogener von exogenen Variablen uni- von multikausalen Variablen. In Abb. 5.3a wäre Variable x_1 uneffektiv, y_1 unikausal; in Abb. 5.3b wäre x_1 multieffektiv, y_1 , y_2 und y_3 jeweils unikausal; in Abb. 5.3c wären x_1 , x_2 und x_3 je uneffektiv, y_1 multikausal.

Das Konzept der Reversibilität läßt sich in Begriffen der hier zugrundegelegten Modellvorstellung ausdrücken durch den zulässigen Wertebereich, den die x -Variablen annehmen dürfen: Liegen Beschränkungen vor (z.B. nur positive Zahlenwerte erlaubt), ist die Interventionsvariable irreversibel; ist dagegen auch die Eingabe negativer x -Werte zugelassen, nennen wir die Interventionsvariable reversibel. Der Grad der "Balance" von Restriktionen im positiven bzw. negativen Wertebereich für die x -Variable kennzeichnet die Leichtigkeit der Reversibilität. Die Beschränkung: $10 < x < -1$ bedeutet etwa, daß die Rücknahme des von x erzeugten Effekts im Extremfall nur mit 1/10 der Geschwindigkeit vorgenommen werden kann wie die Erzeugung der Wirkung.

Die Größe des Anwendungsbereichs einer Interventionsvariablen wird nach DÖRNER bestimmt durch die Menge an Bedingungen, an die sein Einsatz geknüpft ist. Den größten Anwendungsbereich hat somit die x -Variable, deren Wirkung unter allen Umständen gilt. Einschränkende Voraussetzungen der Anwendbarkeit sind im Rahmen unserer Modellvor-

stellungen formal charakterisierbar durch zwei- oder mehrwertige Moderator- bzw. Indikatorvariablen. Für die Pfeildarstellung bedeutet dies, daß prinzipiell so viele Pfeildarstellungen existieren wie es Stufen der Indikatorvariable gibt (bzw. bei mehreren Indikatoren: wie es mögliche Produkte der Indikatorvariablen gibt); für die Matrixschreibweise ist pro Bedingung eine Matrix erforderlich. Diese Obergrenze kann durch bestimmte Bedingungskonstellationen allerdings auch erheblich unterschritten werden. Dies soll an einem kleinen Beispiel demonstriert werden.

DÖRNER (1976, p.56f; vgl. PUTZ-OSTERLOH 1974) schildert das Interpolations-Problem der "Käferzüchtung": Eine den Zoologen bislang unbekannte Gattung von Käfern unterscheidet sich auf acht Merkmalsdimensionen ($y_1 - y_8$), die jeweils zwei oder drei Ausprägungsformen annehmen können (insgesamt $2 \cdot 3 \cdot 2 \cdot 2 \cdot 3 \cdot 2 \cdot 2 \cdot 2 = 576$ unterschiedliche Fälle). Durch zehn unterschiedliche Bestrahlungsformen ($x_1 - x_{10}$) kann gezielt das Erbgut manipuliert werden (und damit die Zugehörigkeit eines Käfers zu einer der 576 möglichen Erscheinungsformen). Aufgabe des Probanden ist es, einen gegebenen Käfertyp durch Wahl der richtigen Bestrahlungsform (bzw. einer Sequenz davon) in einen gewünschten Käfertyp umzuwandeln. Die Tatsache, daß bestimmte Bestrahlungsformen nur bei bestimmten Käferformen wirken, macht dieses Beispiel zu einem, das die Größe des Operator-Anwendungsbereichs demonstriert: nur vier der zehn x -Variablen wirken universell, drei sind an das Vorliegen einer, drei an das Vorliegen zweier Bedingungen geknüpft (siehe genauer DÖRNER 1976, p.57). Die x -Variablen wirken in zwei Fällen auf eine y -Variable, in acht Fällen auf je zwei y -Variablen. Als generelle Regel wird vereinbart, daß die Interventionsvariablen nur dann angewendet werden können, wenn die durch sie erzeugten y -Werte nicht schon vorhanden sind. Auch diese Problemstruktur kann in Form von AR-Prozessen ausgedrückt werden, wobei die x -Variablen ebenso wie die y -Variablen dicho- bzw. trichotom codiert sind. Ob man für jede Randbedingung ein eigenes Kausaldiagramm erstellt oder ob man eine Gruppe von sog. W -Variablen einführt (vgl. STEYER 1983c, der damit die Störvariablen eines Experiments kennzeichnet), die Indikatorfunktion in bezug auf die einzuhaltenden Bedingungen ausüben, ist letztlich eine Frage der Ökonomie der Darstellung, nicht aber des Prinzips. Der Problemlöser, der sich in der Situation des Käferzüchters sieht, muß die "Indikationsstellung" selbst vornehmen; dazu benötigt er die Liste der Bestrahlungsformen und deren Voraussetzungen. Die notwendigen Kausaldiagramme hält der Proband somit nicht im Kopf präsent, sondern greift auf sie in Form eines externen Speichers zurück.

Die Wirkungssicherheit einer Interventionsvariablen kann im Rahmen der stochastischen Modellierung kausaler Systeme einfach durch Angabe der entsprechenden Fehlervarianz beschrieben werden. Neben deren Größe (z.B. in Streuungseinheiten der Interventionsvariablen) können hier auch Angaben zur Verteilungsform des Fehlers (z.B. Normalverteilung) gemacht werden. Die Fehlertheorie ist expliziter Bestandteil stochastischer Modelle; nimmt man im Extremfall ein Modell ohne Fehler an, hat man ein deterministisches System vor sich. Das Ausmaß des Fehlers in nicht-deterministischen Modellen können wir unschwer mit dem Be-

griff des "Rauschens" in Verbindung bringen: je mehr Rauschen über eine x-y-Relation gelegt wird, umso schwieriger fällt dem Beobachter die Aufnahme dieser Relation in sein internes Modell. Im Extremfall mag man wegen der hohen Fehlervarianz eine vorhandene Relation als nicht-existent vermuten; dabei spielt natürlich eine entscheidende Rolle, über welchen Zeitraum man das System beobachten kann. Bei genügend langer Beobachtungszeit wird der Beobachter sein Modell des AR-Prozesses so konstruieren, daß in dem subjektiv vermuteten Ursache-Wirkungs-Verhältnis eine Fehlertheorie enthalten ist. Dies scheint eine psychologisch höchst interessante Fragestellung: unter welchen Bedingungen bilden Menschen Fehlertheorien über die von ihnen vermuteten Wirkbeziehungen? Die ständig falsch gehende Uhr wäre kein geeignetes Beispiel, da dort ja ein systematischer Trend vorliegt, sofern die Uhr nicht an einem Tag drei Minuten vor- und am nächsten zehn Minuten nachgeht; Hypothesen über die systematisch falsch gehende Uhr lassen sich unter Rückgriff auf den Grad des autoregressiven Prozesses beschreiben (siehe unten).

Die materiellen und zeitlichen Kosten einer Interventionsvariable kann man als Faktoren auffassen, die bei Vorliegen mehrerer äquivalenter Interventionsvariablen (siehe oben) das Präferenzverhalten bei ihrer Auswahl steuern. Wenn wie in Abb. 5.3c etwa die Interventionsvariablen x_1 und x_3 mit gleicher Stärke auf y_1 wirken, kann leicht anhand der Kosten für eine x_1 - bzw. x_3 -Einheit entschieden werden, welche Interventionsvariable zu präferieren sein sollte. Unterschreitet der Aufwand für x_2 die Hälfte des Aufwands für die billigste x_1 - oder x_3 -Einheit, wäre Interventionsvariable x_2 zu empfehlen. Ganz allgemein kann hier das Produkt aus Wirkungsstärke pro Einheit und Kosten pro Einheit sowie Kosten pro Zeiteinheit (bei verzögerten Prozessen) die Präferenzliste äquivalenter Interventionsvariablen erstellen helfen. Dies wird schwieriger bei Interventionsvariablen mit unterschiedlichen Nebenwirkungen, die aber prinzipiell ebenfalls einer Kosten-Nutzen-Rechnung zugänglich gemacht werden können, sowie schwierig dort, wo Kosten und Nutzen nur ungenügend quantifiziert werden können. In anderen Fällen kann mit Hilfe mathematischer Verfahren wie etwa der Optimierungsalgorithmen das günstigste Vorgehen bestimmt werden. Die zeitlichen Kosten können dann zu einem Entscheidungskriterium werden, wenn - wie in Beispiel (a) und (d) der Abb. 5.3- eine gleich starke Interventionsvariable einmal sofort und einmal verzögert wirkt: nach dem Sprichwort "Zeit ist Geld" wird man (a) vorziehen.

Nach diesen Ausführungen zu den Eigenschaften von Operatoren im Sinne von Interventionsvariablen und deren Beschreibbarkeit im Rahmen von Kausalmodellen sollen nunmehr Eigenschaften komplexer Systeme untersucht werden, wie sie von DÖRNER et al. (1983, p.19-47) skizziert worden sind. Wir werden uns dabei der (a) Komplexität, (b) Vernetztheit, (c) Eigendynamik und (d) Polytelie widmen. Fragen bezüglich der Unbestimmtheit der Zielsituation und der Intransparenz der Problemstellung sind überwiegend von der Untersuchungssituation abhängig und werden daher nur am Rande behandelt (zur Präzisierung des Transparenz-Konzepts vgl. Kap. 4.4).

Wie bereits an anderer Stelle (vgl. Kap. 3.4) ausführlicher dargestellt, handelt es sich beim Begriff "Komplexität" um einen schillernden Ausdruck, der an Genauigkeit erheblich zu wünschen übrig läßt. HUSSY (1984) führt ersatzweise den Begriff der "Problemschwierigkeit" als zentrales Konzept seiner differenzierten Überlegungen ein und versteht Komplexität als eines von mehreren Bestimmungsstücken der Schwierigkeit, ohne dabei jedoch eine Festlegung zu treffen, ob nun statische, dynamische oder Kontrollkomplexität gemeint sei.

Die differenzierten Sichtweisen von Komplexität vor Augen, wollen wir an dieser Stelle zunächst einmal der Auffassung folgen, wonach die Anzahl der Variablen eine maßgebliche Rolle spielen könnte. Im Rahmen der Kausalmodellierung können wir einerseits die Gesamtzahl der Systemvariablen als Indikator der Komplexität wählen, zum anderen - auf einer höheren Auflösungsebene- die Zahl der x- und y-Variablen sowie deren Verhältnis zueinander. Die absolute Größe sagt etwas aus über die Menge der beteiligten Komponenten, die im zugrundeliegenden System "mitspielen"; das Verhältnis von x- zu y-Variablen gibt ersten Einblick in die Kontrollkomplexität. Lautet etwa das Verhältnis der x- zu den y-Variablen 1:5, muß man den Systemkontrolleur bedauern, der fünf y-Variablen allein durch Variation einer x-Variable steuern soll; im umgekehrten Fall würde man dagegen ein hohes Ausmaß an Kontrollierbarkeit des Systems erwarten können.

Diese Vermutungen können aber erst durch einen Blick auf die "Vernetztheit" der x- und y-Variablen erhärtet werden. Abb. 5.4 zeigt ein Beispiel dafür, wie das Verhältnis von x- zu y-Variablen fehlinterpretiert werden kann, wenn keine Vernetzungsangaben vorliegen.

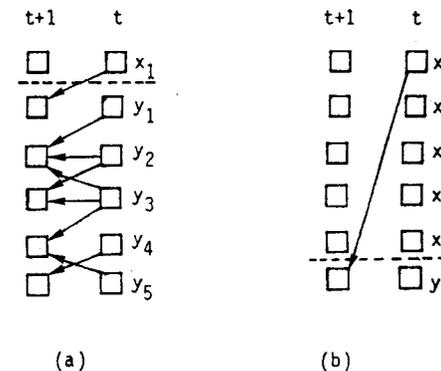


Abb. 5.4: Zwei Beispiele (a) und (b) für Systeme gleicher Variablenzahl, unterschiedlichen Ausmaßes von x- und y-Variablen und gleicher Vernetzung der x- mit den y-Variablen.

Das in Abb. 5.4 gezeigte Beispiel (b) enthält eine Reihe "unwirksamer" x-Variablen, die man im Prinzip aus dem System entfernen kann;

dadurch würde die Proportion der x- und y-Variablen von 1:5 im Fall (a) tatsächlich nicht umgekehrt in Fall (b), sondern 1:1 betragen.

Abb. 5.4a verdeutlicht noch etwas anderes, nämlich indirekte Wirkungen, wechselseitige Wirkungen und Eigendynamiken. Unter wechselseitiger Wirkung wollen wir hier die wechselseitige Beeinflussung von y-Variablen untereinander verstehen; der im Kontext der Varianzanalyse bekannte Ausdruck "Wechselwirkung" darf nicht damit verwechselt werden. Wechselwirkungen dieser Art werden in multivariaten AR_k-Prozessen durch entsprechende Wechselwirkungsvariablen repräsentiert. Der Begriff der Eigendynamik soll für die Relation einer Variablen mit sich selbst gelten. Als indirekte Wirkung verstehen wir ganz allgemein einen Effekt einer y-Variablen auf eine andere y-Variable; wechselseitige Wirkung und Eigendynamik sind somit Spezialfälle der indirekten Wirkung.

Dies sei wieder an einem einfachen, selbstkonstruierten Beispiel demonstriert. Man denke sich ein System mit den drei x-Variablen "Dünger" (=x₁), "Schädlingsfresser" (=x₂) und "Gift" (=x₃) sowie den drei y-Variablen "Blätterzahl" (=y₁), "Käfer" (=y₂) und "Wasserverschmutzung" (=y₃). Zwischen diesen sechs Variablen mögen die unterschiedlich differenzierten Beziehungen bestehen, die Abb. 5.5 veranschaulicht.

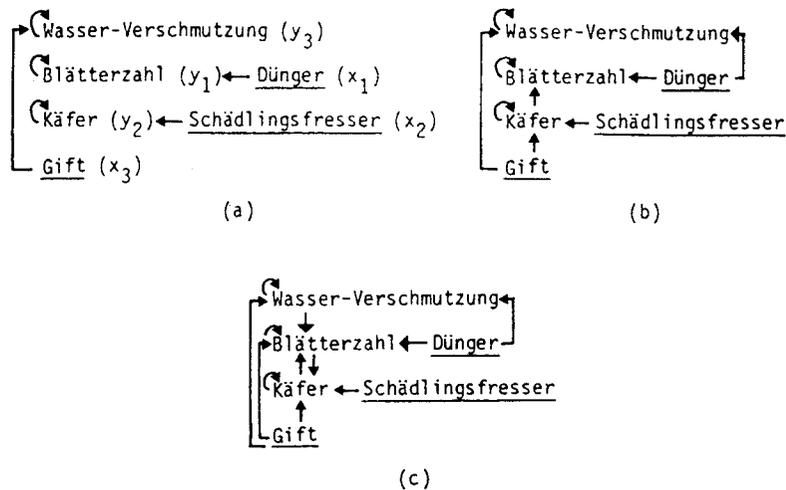


Abb. 5.5: Drei verschiedene Vernetztheitsgrade eines simplen ökologischen Modells mit drei exogenen (=unterstrichen) und drei endogenen Variablen.

Ganz offenkundig repräsentieren die drei Modelle in Abb. 5.5 den unterschiedlichen Vernetztheitsgrad eines Systems: Während in Fall (a) ein reines Modell unieffektiver unikausaler Variablen vorliegt (jeder x- ist eine y-Variable zugeordnet), werden in Fall (b) multieffektive multikausale Variablenrelationen (bei "Dünger" und "Gift") sowie eine

indirekte Wirkung (von "Käfer" auf "Blätterzahl") eingeführt. Fall (c) enthält zusätzlich eine wechselseitige Beziehung (zwischen "Blätterzahl" und "Käfer").

Auf einer formalen Ebene lassen sich diese drei Modelle darstellen in Form von Matrizen des zugrundeliegenden autoregressiven Prozesses (AR-Prozeß); im hier gewählten Beispiel sollen aus Gründen der Einfachheit nur AR₁-Prozesse gelten, d.h. die x-Variablen zum Zeitpunkt t wirken sich unverzögert auf die y-Variablen zum Zeitpunkt t+1 aus, nicht jedoch direkt auf t+2. Eine indirekte Wirkung auf t+2 besteht jedoch in den Fällen (b) und (c), nämlich vermittelt über Variablen zum Zeitpunkt t+1. Abb. 5.6 enthält die drei entsprechenden A-Matrizen.

	x ₁	x ₂	x ₃	y ₁	y ₂	y ₃
x ₁	1					
x ₂		1				
x ₃			1			
y ₁	1			1.1		
y ₂		-2			1	
y ₃			0.5			0.97

(a)

	x ₁	x ₂	x ₃	y ₁	y ₂	y ₃
x ₁	1					
x ₂		1				
x ₃			1			
y ₁	1			1.1	-0.05	
y ₂		-2	-1		1	
y ₃	0.01		0.5			0.97

(b)

	x ₁	x ₂	x ₃	y ₁	y ₂	y ₃
x ₁	1					
x ₂		1				
x ₃			1			
y ₁	1		-0.01	1.1	-0.05	-0.05
y ₂		-2	-1	-0.05	1	
y ₃	0.01		0.5			0.97

(c)

Abb. 5.6: Drei Matrizen der AR₁-Prozesse für die Modelle aus Abb. 5.5. Die Variablen x₁-x₃ sowie y₁-y₃ sind im Text benannt; Zeilenvariablen sind Spaltenvariablen zeitlich nachgeordnet.

Aus Abb. 5.6 wird nun der genaue Wirkzusammenhang deutlich; während Abb. 5.5 lediglich angibt, an welchen Stellen Kausalbeziehungen vorliegen, kann man nunmehr Richtung und Stärke der jeweiligen Relation ablesen (dies hätte in Abb. 5.5 durch Beschriftung der Pfeile eben-

falls erreicht werden können). Die Zeilenvariable ist der Spaltenvariable zeitlich nachgeordnet: Nimmt man die Einträge einer Zeile heraus, kann man erkennen, von welchen Variablen die Zeilenvariable wie stark beeinflusst wird. Die oberen beiden Quadranten jeder Matrix sind trivial: Die A_{xx} -Teilmatrix (links oben) ist eine Einheitsmatrix (die x -Variablen beeinflussen sich nicht gegenseitig und die Stärke der x -Eingaben wird unverändert übernommen und über die Zeit hinweg aufrechterhalten), die A_{xy} -Matrix (rechts oben) ist immer eine Nullmatrix, da die Ursachen (= x -Variable) nicht von den Wirkungen (= y -Variable) abhängig sein können. Interessant sind somit nur die unteren beiden Teilmatrizen A_{yx} und A_{yy} . In der A_{yx} -Teilmatrix (links unten) sind die Parameter für die Wirkung der x - auf die y -Variablen enthalten. Im Fall (a) sehen wir für jedes x genau ein y , das beeinflusst wird; in Fall (c) gibt es nur für x_2 eine einzelne Wirkung, x_3 hat dort drei - unterschiedlich starke - Wirkungen. Die vierte Teilmatrix, A_{yy} (rechts unten), regelt die Beziehungen der y -Variablen untereinander: Während die Diagonalelemente eine Angabe über die Konstanz der jeweiligen y -Variable machen (bei konstanten Randbedingungen bedeuten Werte größer als 1 Wachstums-, kleiner als 1 Schrumpfungsprozesse; der Wert 1 bedeutet eine unveränderte Übernahme des entsprechenden vorangegangenen y -wertes), liefern die nichtdiagonalen Elemente Auskunft über wechselseitige Systemprozesse zwischen den y -Variablen.

Haben sich die bisherigen Ausführungen mit den Aspekten der Komplexität, Vernetztheit und Eigendynamik in Strukturmodellen beschäftigt, muß nunmehr noch die Frage der Polytelie der Problemstellung, also der simultanen Berücksichtigung mehrerer, u.U. einander widersprechender Ziele untersucht werden. Diese ergibt sich nicht aus dem Modell selbst, sondern aus dem Auftrag, den der "Systemanalytiker" (also der Proband) entweder vom Versuchsleiter erhält oder sich selbst auferlegt. Eine polytelische Zielvorgabe durch den Versuchsleiter kann etwa darin bestehen, dem Probanden die Erreichung bestimmter y -Zielwerte aufzutragen. Polytelisch ist diese Vorgabe dann, wenn mehrere y -Variablen zu optimieren sind, die Systemstruktur die Erreichung eines bestimmten y -Zustands unabhängig vom Zielwert einer anderen y -Variable jedoch behindert (siehe Abb. 5.5b, wo Gift nicht nur die Käfer reduziert, sondern zugleich die Wasserverschmutzung erhöht). Selbstverständlich können auch monotelische Zielvorgaben gemacht werden, etwa nur den Wert einer y -Variable auf eine bestimmte Höhe zu bringen und dann konstant zu halten. Dies sind Varianten der experimentellen Prozedur, die mit dem Kausalmodell selbst, um das es geht, nichts zu tun haben. Genauso verhält es sich mit der Frage nach der Transparenz: Auch hierbei handelt es sich nicht um einen Aspekt des Strukturmodells, das erkannt werden soll, sondern vielmehr um die Frage, wie x - oder y -Variablen erfaßt werden können (z.B. ob der Versuchsleiter bestimmte Informationen zurückhält). Dies muß an anderer Stelle diskutiert werden; hier ging es zunächst einmal um die Demonstration einiger Möglichkeiten, die sich aus der Darstellung komplexer Systeme in Form von multivariaten autoregressiven Prozessen ergeben.

5.3 Konstruktion von Kausalmodellen: Eine Fallstudie

Im Rahmen erster Erprobungen mit dem Untersuchungsgegenstand "Kausalmodelle" wurde eine Explorationsstudie durchgeführt, die nachfolgend kurz beschrieben werden soll.

Zwei mit Kausalmodellen nicht völlig unvertraute, mathematisch interessierte Probandinnen sahen sich gemeinsam vor folgendes Problem gestellt: Ein ihnen unbekanntes Kausalsystem (drei x - und drei y -Variablen) sollte in seiner Kausalstruktur "erkannt", d.h. identifiziert und auf einen definierten Zielzustand hingeführt werden. Die Aufgabenstellung war in folgender Weise vorstrukturiert: Zunächst einmal konnte man in vier Durchgängen mit jeweils acht Zeittakten die Wirkungsweise des Systems ohne jede Zielvorgabe erfahren. Im fünften Durchgang sollte dann das System von dem -zu Durchgangsbeginn jeweils gleichen- Ausgangszustand in den bereits vorher bekanntgegebenen Zielzustand überführt werden. Nach jedem Durchgang zeichneten die Probandinnen die von ihnen vermutete Kausalstruktur auf ein Blatt Papier. Notizen während der Durchgänge wurden erlaubt (und auch reichlich genutzt).

Dieses Vorgehen hat Parallelen zu der von EYFERTH et al. (1982) gewählten Prozedur, bei dem die Probanden eine fiktive "WELT" (ein künstliches System mit vier verschiedenen Elementen, deren Verhalten - ausgedrückt in Bewegungsformen von Quadraten auf dem Bildschirm eines Kleincomputers- zu erkennen ist) beobachten und handelnd begreifen sollten. Dort wurde ebenfalls der schrittweise Aufbau der Systemrepräsentation anhand von Daten des lauten Denkens untersucht. HESSE (1983), der für verschiedene Fragestellungen aus dem Bereich der Wissensrepräsentation interessante Versuchspläne beschreibt, kommt mit seinem fünften Plan (p.67f) zu ähnlichen Überlegungen wie wir, wenn er sich für die Formen des Lernens bzw. der "Produktionskompilation" bei wiederholter Bearbeitung komplexer Probleme interessiert. KLUWES (1983) Forschungsprojekt über die "Ausbildung subjektiver Ordnungsstrukturen durch Erfahrungen beim Umgang mit umfangreichen Systemen" kommt ebenfalls dieser Fragestellung nahe.

Bevor die Ergebnisse einer explorativen Fallstudie geschildert werden, soll kurz das vom Versuchsleiter vorgegebene System mit seinen Vernetzungsstrukturen und den für jeden Durchgang gleichen Ausgangsbedingungen dargestellt werden (vgl. Abb. 5.7).

Das (willkürlich) gewählte System besteht aus drei exogenen (=x-) Variablen und drei endogenen (=y-) Variablen. Fünf Kausalpfeile gehen von den exogenen auf die endogenen Variablen, drei Pfeile kennzeichnen autoregressive Prozesse innerhalb der nicht miteinander vernetzten y -Variablen. Drei Strukturgleichungen beschreiben das System vollständig (die hochgestellten Indizes repräsentieren den Zeitindex t):

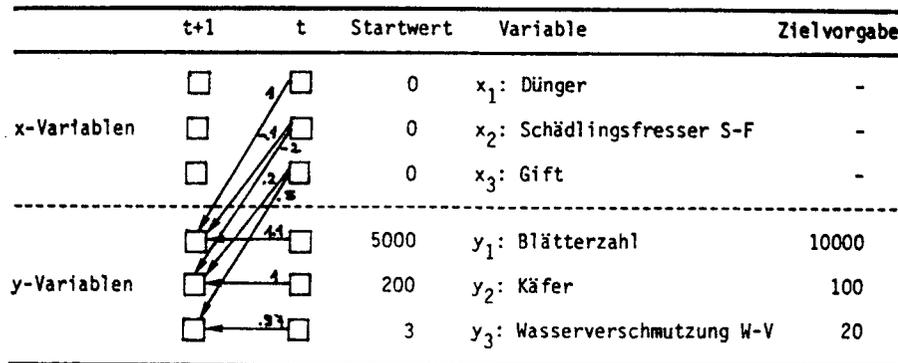


Abb. 5.7: Detaillierte Angaben über das vorgegebene System (weitere Angaben im Text).

$$y_1^{t+1} = 1.10 * y_1^t + 1.00 * x_1^t + (-0.1) * x_2^t \quad (5.7)$$

$$y_2^{t+1} = 1.00 * y_2^t + (-2) * x_2^t + 0.2 * x_3^t \quad (5.8)$$

$$y_3^{t+1} = 0.97 * y_3^t + 0.50 * x_3^t \quad (5.9)$$

Das System ist in einen ökologischen Kontext eingebettet: Man versetze sich in die Rolle eines Kleingärtners, der die Blätterzahl eines Baumes, die Zahl der in diesem Baum lebenden Käfer sowie das damit verbundene Wassernetz (Wasser-Verschmutzung, W-V) überwachen soll. Er kann Dünger geben, Schädlingsfresser (S-F) placieren sowie Gift verabreichen. Die Zusammenhänge zwischen diesen sechs Variablen muß er als gänzlich ungeklärt annehmen: Weder darf er z.B. davon ausgehen, daß das Gift (schädliche ?) Käfer tötet, noch weiß er etwas darüber, wozu die Schädlingsfresser gut sein sollen. Er wird lediglich gebeten, eine "technokratische" Haltung einzunehmen und die Beziehungen zwischen den Variablen zu erkunden, um nach der Explorationsphase sein erworbenes "Struktur"- bzw. "Funktions"-wissen erfolgreich im Sinne der (multiplen) Zielvorgabe einsetzen zu können. Der Proband befindet sich somit in der Rolle des Wissenschaftlers, der Beobachtungen an einem Untersuchungsgegenstand macht und dabei eine Theorie über diesen Gegenstand(sbereich) entwirft und prüft, um anschließend einen "Auftrag" erfüllen zu können. Die Zielvorgabe für das Beispiel zeigt die rechte Spalte von Abb. 5.7, deren Werte genauso willkürlich gewählt wurden wie die Startwerte, mit denen das System zu jedem Durchgang beginnt.

Die Ziffern an den Kausalpfeilen in Abb. 5.7 geben die Stärke der Wirkung an, mit der der Wert einer Variable zum Zeitpunkt t den Wert der "Ziel"-Variable zum Zeitpunkt t+1 beeinflusst; es sind dies die Koeffizienten der Strukturgleichungen (5.7), (5.8) und (5.9). Zum Beispiel ist der Pfeil $x_1^t \rightarrow y_1^{t+1}$ von der Stärke 1, so daß eine Gabe von 100 Düngereinheiten (x_1) die Blätterzahl (y_1) um 100 Einheiten ver-

größert. Die Wirkung von x_2 auf y_1 beträgt -0.1 , d.h. 10 Einheiten von x_2 vermindern y_1 um 1 Einheit. Die autoregressiven Prozesse innerhalb der y-Variablen geben Auskunft über die Eigendynamik des Systems, also über das, was geschieht, wenn keine Eingriffe erfolgen. In diesem Fall erhöht sich die Variable y_1 "automatisch" von Zeittakt zu Zeittakt um 10%, y_2 bleibt konstant und y_3 schrumpft jeweils um 3%. Wechselseitige Wirkungen zwischen den y-Variablen existieren nicht, das System realisiert darüberhinaus nur AR_1 -Prozesse, d.h. es gibt keine zeitverzögerte Wirkung der x- auf die y-Variablen. Außerdem wurde kein Fehler auf die Wirkungsmechanismen zugelassen, das System reagiert also deterministisch.

Der Versuchsablauf läßt sich wie folgt beschreiben. Zunächst erhalten die Probanden auf dem Bildschirm eines Kleincomputers CBM-8032 die Werte der x- und y-Variablen zum Zeitpunkt t=1 und werden gebeten, nunmehr zu entscheiden, welche Stärke sie den Eingriffsvariablen x_1 bis x_3 zuordnen wollen. Voreingestellt ist der Wert Null, der durch bloßes Betätigen der Return-Taste übernommen werden kann; ansonsten ist der gewünschte numerische Wert einzutippen und ebenfalls durch die Return-Taste abzuschließen. Die Reihenfolge der Abfrage der x-Variablen bleibt während des ganzen Versuchs konstant. Eine Beschränkung des Bereichs möglicher Eingaben nach oben oder unten ist nicht vorgesehen, es können also auch negative Werte eingegeben werden. Darauf wurden die Probandinnen hingewiesen. Genauso können auch die y-Variablen jeden beliebigen Wert annehmen, also auch negativ werden, eine Möglichkeit, die auf den ersten Blick merkwürdig -da kontrafaktisch- erscheint, aber dem Modellkonstrukteur die Möglichkeit bietet, die Konsequenzen überhöhter Maßnahmen festzustellen (z.B. negative Blätterzahl infolge starker Giftgabe). Auf seiten der x-Variablen bietet dies einen reversiblen Einsatz der Interventionsvariablen, d.h. Maßnahmen können rückgängig gemacht werden (sofern dies nicht mit unerwünschten Nebenwirkungen verbunden ist). Der Proband erhält so die (unnatürliche) Möglichkeit, das vor ihm befindliche System in alle von ihm gewünschten Richtungen zu beeinflussen, es steht vollständig und uneingeschränkt unter seiner Kontrolle. Die "Unnatürlichkeit" diesbezüglich gilt übrigens auch für andere Untersuchungsparadigmen der Forschung zum komplexen Problemlösen, in denen der "Manager" oder "Bürgermeister" ja ebenfalls mit geradezu autokratischen Fähigkeiten ausgestattet wird, auch wenn dort nicht jede Entscheidung problemlos (d.h. folgenlos) rückgängig zu machen ist.

Die im Verlauf eines Durchgangs anfallenden Systemdaten (gerundet jeweils auf die nächste ganze Zahl) werden säulenartig nebeneinander auf dem Bildschirm präsentiert (jede Säule entspricht einem Zeittakt); die Zahl der auf dem Monitor simultan verfügbaren Zeittakte eines Durchgangs beträgt -aus technischen Gründen- maximal fünf, minimal eins. Dadurch kann man das Ausmaß der Gedächtnisbelastung ein Stück weit variieren, was insbesondere bei der Einführung zeitverzögerter Wirkungen nicht ohne Einfluß auf den empfundenen Schwierigkeitsgrad bleiben dürfte. Da wir diese Fragestellung momentan nicht in den Mittelpunkt unserer Untersuchung stellen wollen, wird die maximal verfügbare Zahl von fünf nebeneinander präsentierten Zeittakten ge-

wählt. Nach jedem der insgesamt fünf Durchgänge mit jeweils acht Zeittakten wurden die Probandinnen gebeten, das von ihnen vermutete Wirkungsnetz in freier Niederschrift zu Papier zu bringen; anstelle numerischer Angaben zu den Wirkstärken waren auch Angaben in Form von Pfeilen mit Plus- oder Minus-Zeichen zur Charakterisierung einer vermuteten Relation erlaubt.

Soviel zu den Vorbemerkungen, nun zu den Resultaten dieser Fallstudie. Abb. 5.8 zeigt die fünf sukzessive erstellten Kausaldiagramme (a) bis (e) sowie die korrekte, aufzufindende Struktur (f). Die Kausaldiagramme (a) bis (e) stellen Transformationen der freien Niederschriften dar, die weniger systematisch angefertigt worden waren.

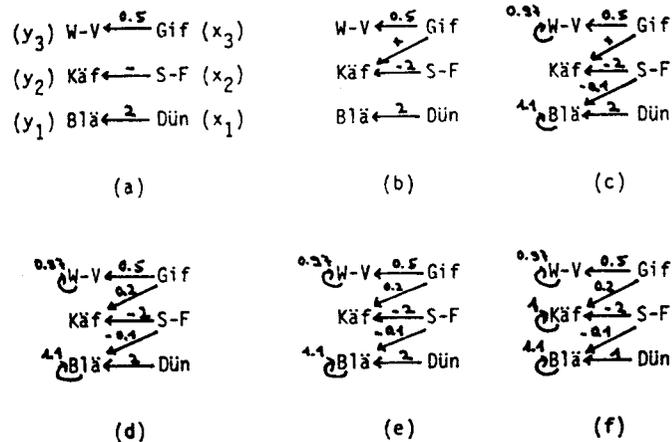


Abb. 5.8: Die fünf in Anschluß an den jeweiligen Durchgang angefertigten Kausaldiagramme (a)-(e) sowie die tatsächliche Struktur (f). Zahlen über den Pfeilen bedeuten Angaben über die Wirkstärke, in einigen Fällen stehen "+" und "-" als Indikatoren für die vermutete Wirkrichtung.

Im ersten Durchgang (Abb. 5.8a) werden einfache Kausal"ketten" identifiziert: Die "Haupteffekte" der x-Variablen werden bereits eindeutig erkannt, also sowohl Richtung als auch Vorzeichen der Pfeile von x auf y. Die Stärke der x_3y_3 -Relation ist bereits in ihrer Höhe richtig, die von x_1 auf y_1 liegt mit 2 doppelt so hoch wie der tatsächliche Parameter. Erstaunlich ist, daß die stärkste Wirkung des ganzen Systems (x_2 auf y_2 mit Gewicht 2) nur in ihrer Richtung, nicht aber in ihrer Höhe bekannt ist. Dies gelingt im zweiten Durchgang (Abb. 5.8b), in dem zusätzlich die erste Nebenwirkung einer x-Variable aufgedeckt wird (zumindest in ihrer Richtung); man könnte davon sprechen, daß nunmehr das Kausal"netz" ins Auge gefaßt wird. Der dritte Durchgang (Abb. 5.8c) deckt eine weitere Nebenwirkung, nämlich die von x_2 auf y_1 , auf, und zwar gleich in korrekter Richtung und Stärke. Außerdem werden die Autokorrelationen von y_1 und y_3 in korrekter Weise erfaßt. Mit Durchgang 4 (Abb. 5.8d) ist die Aufdeckung der Kausal-

struktur für die Probandinnen beendet. Die letzte "Unschärfestelle" (Relation x_3y_2) wird geklärt, die Probephase ist damit abgeschlossen. Für Durchgang 5 (Abb. 5.8e), in dem das vorgegebene Systemziel angestrebt werden soll, liegt eine fast vollständige Repräsentation des implementierten Systems vor, die auch im letzten Durchgang nicht mehr verändert wird, obwohl der Zielzustand (vgl. Abb. 5.7) nicht ganz exakt erreicht wird: im achten Zeittakt ist ein Systemzustand von 10307 Blättern, 100 Käfern und einer W-V von 20 erreicht. Im dritten Zeittakt dieses Durchgangs wurde übrigens die größte Zielnähe erreicht ($y_1=9707$, $y_2=101$, $y_3=19$), sofern man zur Bewertung der Zielnähe ein ungewichtetes Abstandsmaß verwendet.

An zwei Stellen weicht die vermutete von der tatsächlichen Systemstruktur ab. Der Gewichtungsfaktor für Dünger wurde -übrigens gleich von Anfang an- mit 2 doppelt so hoch wie der wahre Wert 1 angesetzt; außerdem findet sich in den Aufzeichnungen keine Angabe zur Autokorrelation der Käfer (Gewicht 1). Da der vorgegebene Zielzustand dennoch gut angenähert wurde, bleibt dafür nur eine Erklärung: Die Probanden haben mit dem wahren Wert für die x_1y_1 -Beziehung gerechnet, aber den falschen Wert notiert. Die Stabilität der Käfer war darüberhinaus zu trivial, um notiert zu werden, damit brauchte nicht "gerechnet" werden. Eventuell handelt es sich hierbei auch um ein Beispiel dafür, daß zwar Funktions-, nicht aber Strukturwissen vorliegt: Die Probandinnen wissen, wie sie Dünger in bezug auf Blätterzahl einsetzen müssen, ohne daß sie jedoch dieses Wissen richtig angeben können (so die Vermutung). Dies wäre allerdings insofern überraschend, als ja selbst "schwierige" Beziehungen (etwa die 3%ige Schrumpfung von W-V aufgrund der y_3y_3 -Beziehung) in ihrem korrekten Zahlenwert erfaßt wurden. Das Problem, das hierdurch deutlich gemacht wird, betrifft den Unterschied zwischen dem Kausalwissen im Kopf der Probanden und dem durch das -ungenau- Meßinstrument dokumentierte Strukturwissen.

Zu erwähnen bleibt übrigens noch die Tatsache, daß die gesamte, hier berichtete Fallstudie ca. zwei Stunden konzentrierte Arbeit in Anspruch nahm. Die nicht aufgezeichneten Daten des lauten Denkens zeigten, daß permanent Hypothesen über das Kausalmodell gebildet und geprüft wurden; ständige Überraschungen sorgten dabei für die Aufrechterhaltung der Motivation ebenso wie das gelegentliche Eintreffen prognostizierter Ereignisse, wodurch eine subjektive Abschätzung des erreichten Ausmaßes an Systemkontrolle möglich wurde.

5.4 Anwendungsmöglichkeiten

Daß das geschilderte Begriffsinventar der ARK-Prozesse tauglich ist für die Beschreibung verschiedenster Untersuchungsinstrumente der Denkpsychologie, soll an zwei Anwendungsfällen demonstriert werden.

Zunächst einmal betrachte ich das von FUNKE (1981) vorgeschlagene System "MONDLANDUNG". Abb. 5.9 enthält das Pfeildiagramm für dieses dynamische System.

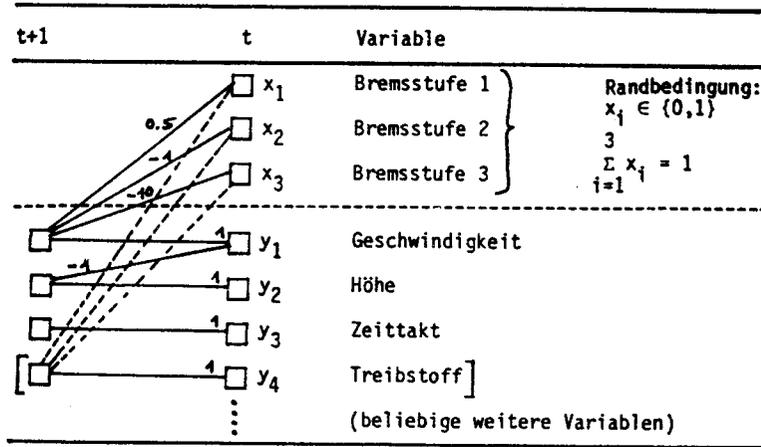


Abb. 5.9: Pfeildiagramm einer Variante der "MONDLANDUNG" von FUNKE (1981).

Drei exogene Variablen (die drei dem Probanden zur Auswahl stehenden Bremsstufen) stehen vier endogenen Variablen gegenüber, wobei die unterschiedlichen Bremsstufen, von denen gemäß Nebenbedingung zu jedem Zeitpunkt nur eine gewählt werden kann (und diese eine auch nur den Wert 1 annehmen kann), lediglich auf die Variable "Geschwindigkeit" Einfluß nehmen bzw. je nach experimenteller Variante auch den Treibstoff reduzieren. Indirekt wirkt die Bremsstufe natürlich auch auf die Variable "Höhe", während "Zeittakt" völlig unbeeinflusst von der Wahl des Probanden fortschreitet.

Die im Pfeildiagramm der Abb. 5.9 enthaltenen Relationen lassen sich auch in Form folgender Strukturgleichung darstellen:

$$y_1^{t+1} = y_1^t * 1 + x_1^t * 0.5 - x_2^t * 1 - x_3^t * 10 \quad (5.10)$$

$$y_2^{t+1} = y_2^t * 1 - y_1^t * 1 \quad (5.11)$$

$$y_3^{t+1} = y_3^t * 1 + 1 \quad (5.12)$$

$$y_4^{t+1} = y_4^t * 1 - x_1^t * 0.01 - x_2^t * 0.04 - x_3^t * 0.07 \quad (5.13)$$

Natürlich müssen die in Abb. 5.9 genannten Nebenbedingungen erfüllt sein. - Die Matrix der Parameter für diesen AR₁-Prozeß zeigt Abb. 5.10.

	x_1	x_2	x_3				
				y_1	y_2	y_3	y_4
x_1							
x_2							
x_3							
y_1	0.5	-1	-10	1			
y_2				-1	1		
y_3						1	
y_4	-0.01	-0.04	-0.07				1

Abb. 5.10: AR₁-Matrix für das System "MONDLANDUNG" aus Abb. 5.9.

Wie aus den drei unterschiedlichen Darstellungsformen gleichermaßen ablesbar, handelt es sich bei "MONDLANDUNG" um ein System, in dem mit einer Variable (hier aus formalen Gründen auf drei exogene Größen aufgeteilt) eine zweite beeinflusst werden kann (bzw. je nach Variante auch noch weitere, Nebenbedingungen repräsentierende Variablen). Das System enthält eine indirekte Wirkung der x - auf die y -Variablen ("Bremsstufe" wirkt vermittelt über "Geschwindigkeit" auf "Höhe"), die wohl auch die Schwierigkeit dieses Problems ausmacht.

Der Vergleich mit einem zweiten, von der Variablenzahl und vom Grad der Differenzgleichungen her gesehen ähnlichen System namens "HAMURABI" (GEDIGA, SCHÖTTKE & TUCKE 1982) zeigt nun den entscheidenden Vorteil einer solchen formalen Darstellung der Untersuchungsinstrumente. Abb. 5.11 enthält zunächst wieder das Pfeildiagramm.

Wiederum sind es drei exogene Variablen, mit denen nunmehr fünf endogene Variablen beeinflusst werden können. Der Proband bestimmt zum einen die "Nahrung pro Einwohner"; damit reduziert er den Kornspeicher und erhöht die Population. Ackerland kann ge- oder verkauft werden: Dies senkt oder erhöht den Kornspeicherinhalt und erhöht oder senkt den Umfang des Ackerlands. Schließlich kann der Proband über die Bebauung des Ackerlands entscheiden; damit beeinflusst er Ernte und Saat. Die Strukturgleichungen dieses Systems lauten:

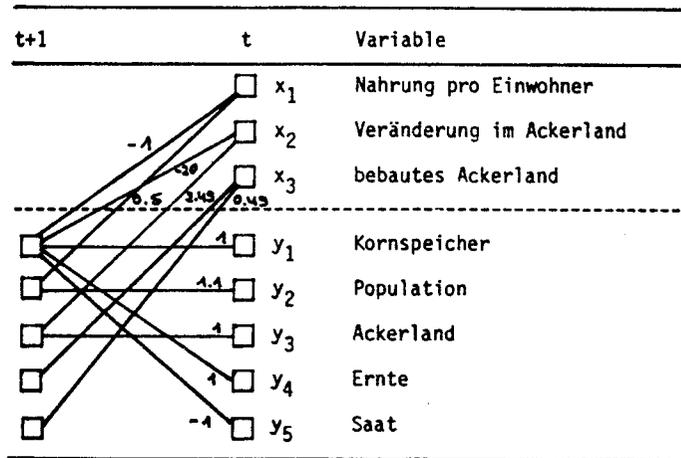


Abb. 5.11: Pfeildiagramm des Systems "HAMURABI" von GEDIGA, SCHÖTTKE & TUCKE (1982).

$$y_1^{t+1} = y_1^t * 1 + y_4^t * 1 - y_5^t * 1 - x_1^t * 1 - x_2^t * 20 \quad (5.14)$$

$$y_2^{t+1} = y_2^t * 1.1 + x_1^t * 0.5 \quad (5.15)$$

$$y_3^{t+1} = y_3^t * 1 + x_2^t * 1 \quad (5.16)$$

$$y_4^{t+1} = x_3^t * 2.49 \quad (5.17)$$

$$y_5^{t+1} = x_3^t * 0.49 \quad (5.18)$$

Auch hier kann natürlich die Parametermatrix des modellierten AR₁-Prozesses erstellt werden. Sie ist in Abb. 5.12 dargestellt.

Nachdem für beide Systeme, "MONDLANDUNG" und "HAMURABI", eine einheitliche Darstellungsweise vorliegt, kann man nun versuchen, die charakteristischen Unterschiede herauszufinden (wobei die Nebenbedingungen nicht vergessen werden dürfen). Ganz offenkundig handelt es sich bei "MONDLANDUNG" um das einfachere System: Die dortige Av_x-Teilmatrix (vgl. Abb. 5.10) enthält zwar ebensoviel Parameter wie "HAMURABI" (vgl. Abb. 5.12), jedoch ist sofort ersichtlich, daß die Interventionsvariablen nur auf zwei endogene statt auf fünf Variablen wirken. Außerdem findet sich in der Av_y-Teilmatrix der "MONDLANDUNG" nur eine wechselseitige Abhängigkeit, bei "HAMURABI" sind es zwei. Schließlich besitzt "HAMURABI" eine Eigendynamik in Form eines Diagonalelements der Av_y-Matrix vom Wert größer Eins: Die Population (y₂) wächst ceteris paribus je Zeiteinheit um 10%, also ein exponentieller

	x ₁	x ₂	x ₃					
				y ₁	y ₂	y ₃	y ₄	y ₅
x ₁								
x ₂								
x ₃								
y ₁	-1	-20					1	-1
y ₂	0.5					1.1		
y ₃		1					1	
y ₄			2.49					
y ₅			0.49					

Abb. 5.12: AR₁-Matrix für das System "HAMURABI" aus Abb. 5.11.

Verlauf. Bei "MONDLANDUNG" kann jeweils nur ein uni- bzw. bi-effektives x gewählt werden, bei "HAMURABI" sind es drei bi-effektive Interventionsvariablen, die die V_p berücksichtigen muß. Ähnlich sind sich die beiden Systeme hinsichtlich der Zentralität einzelner endogener Variablen: y₁ kann in beiden Fällen als multikausale Variable bezeichnet werden, die sich von den sonstigen uni- bzw. bikausalen endogenen Variablen abhebt. - Soviel zum Vergleich zweier Systeme, der durch diese formale Darstellung ermöglicht wird.

5.5 Bausteine für eine Theorie der Konstruktion von Kausalmodellen über dynamische Prozesse

Nachdem in den vorangegangenen Passagen die Terminologie und Methodik multivariater AR_x-Prozesse als Beschreibungsmittel für dynamische Simulationsmodelle vorgestellt wurden, soll nunmehr versucht werden, einen theoretischen Bezugsrahmen darüber zu entwickeln, wie Problemlöser solche Kausalmodelle "konstruieren". Dabei handelt es sich um keinen einmaligen Akt, sondern um einen aktiven Prozeß, der von einigen Autoren (z.B. HAYES-ROTH, KLAHR & MOSTOW 1981) auch als "knowledge refinement" bezeichnet wird. Die Situation, um die es geht, läßt sich folgendermaßen veranschaulichen: Der Problemlöseforscher entwirft ein Kausalmodell mit dynamischen Prozessen über einen bestimmten Gegenstandsbereich (oder über ein künstliches System) und legt es dem Probanden zur Bearbeitung vor. Der Proband konstruiert seinerseits aus den beobachteten Systeminformationen ein Modell über den Prozeß, der den vom Simulationsprogramm erzeugten Daten zugrundeliegt. Diese Modellkonstruktion des Probanden, deren Veränderung im Laufe der Bearbeitung sowie die Nähe des vom Probanden konstruierten Modell zum Modell des Forschers werden untersucht. Um die aktive Tätigkeit des Probanden zu betonen, soll von ihm auch als dem naiven Systemkonstruk-

teur gesprochen werden. Der wissenschaftliche Systemkonstrukteur -und als solchen könnte man analog dazu den Problemlöseforscher bezeichnen, der das System vorgibt- spielt im Rahmen der hier zu untersuchenden Problemlöseprozesse keine Rolle.

Die im folgenden gemachten Annahmen stellen keine im strengen Sinne prüfbare Theorie dar, da sie zum einen nicht formalisiert sind, zum anderen die Frage der empirischen Prüfbarkeit teilweise ungeklärt ist. Dies soll weiteren Arbeiten vorbehalten bleiben. Es geht mir darum, einen allgemeinen Rahmen abzustecken, innerhalb dessen komplexes Problemlösen im Sinne der Konstruktion von Kausalmodellen, die einem dynamischen System zugrundeliegen, begreifbar wird. Damit ist zugleich der Gültigkeitsanspruch dieser Annahmen festgelegt: er besteht für solche komplexen Probleme, die sich in Form multivariater AR_k -Prozesse beschreiben lassen. Wo eine derartige objektive Problemstruktur vorliegt, sollen die nachfolgenden Annahmen über den Prozeß der Konstruktion und Modifikation eines subjektiven Modells gelten.

Die Prognosen, die aus den Annahmen ableitbar sind, scheinen mir nicht trivial zu sein; inwiefern diese Skizze als Theorie-Entwurf tauglich ist, muß der Leser selbst entscheiden. Im Sinne einer Integration sollen -soweit möglich- Parallelen zu anderen Erklärungsansätzen gezogen werden, um eine Einordnung des vorgeschlagenen Rahmenkonzepts zu ermöglichen. Einige Fragen und Probleme dieses Konzepts (z.B. anthropologische Voraussetzungen) werden im Anschluß an die nachfolgend vorgestellten Annahmen diskutiert.

Annahme 1: Der naive Systemkonstrukteur entwickelt ein operatives Abbild (ein Kausalmodell) über die Dynamik der ablaufenden Prozesse, das in Form von Matrizen der dem zu steuernden System zugrundeliegenden multivariaten AR_k -Prozesse darstellbar ist (Repräsentationsannahme).

Dies ist eine relativ allgemeine Annahme, da zunächst keine Entscheidung über die Art der individuellen Systemrepräsentation getroffen wird (etwa in Form von Netzwerken, Schemata oder Produktionssystemen), sondern lediglich postuliert wird, daß das vorhandene Wissen über ein dynamisches System derartig dargestellt werden kann. Die Frage nach einem Meßinstrument des Systemwissens läßt sich daher so beantworten: stimmt Annahme 1, sollte man Aussagen bzw. Aufzeichnungen der Pbn in Form von Matrizen multivariater AR_k -Prozesse (oder gleichwertig: in Form von Strukturgleichungen) bringen können. Datenquellen dürfen verschiedener Natur sein, solange es möglich ist, daraus qualitative oder quantitative Angaben über einzelne Parameter der ablaufenden AR_k -Prozesse zu entnehmen.

Hierbei stellt sich die Frage, ob nicht eine Einführung der naiven Systemkonstrukteure in die Terminologie multivariater AR_k -Prozesse vorteilhaft wäre, um eine leichtere Übersetzbarkeit von Äußerungen der Probanden in die gewünschte Form zu erreichen (vgl. hierzu Kurt LEWINS Beitrag "Die Erziehung der Versuchsperson zur richtigen Selbstbeobachtung und die Kontrolle psychologischer Beschreibungsangaben" in dem

von METRAUX 1981 herausgegebenen Band 1 der Werkausgabe). Dabei handelt es sich um die prinzipielle Frage, ob nicht der bisherige Untersuchungsansatz im Bereich komplexen Problemlösens durch einen anderen Zugang nutzbringend ergänzt werden muß: die Studie von Problemlösungsprozessen trainierter Probanden. Das Versagen von naiven Probanden in kognitiv schwierigen Situationen festzustellen ist die eine Seite, die Seite, die am wenigsten überraschende Ergebnisse liefert, wenn man an die Schwierigkeiten denkt, die Probanden bereits in einfachen Situationen erfahren. Der andere Zugang besteht darin, das Lösungsverhalten der Probanden zu untersuchen, die ein vorheriges Training durchlaufen haben und in bezug auf grundsätzliche Mechanismen und Prinzipien sowie in bezug auf eine zu verwendende Terminologie als "wissend" zu bezeichnen wären. Eine derartige zweigleisige Strategie ist zum Teil in den Studien realisiert, die Novizen und Experten vergleichend untersuchen (siehe etwa SIMON & SIMON 1978), allerdings stammt das Expertenwissen in den bislang untersuchten, zumeist aus dem naturwissenschaftlichen Bereich kommenden Problemen nicht aus einem wiederholten Umgang mit einem System, sondern eher aus einer Art "Hintergrund"-Wissen. So wäre etwa zu fragen, wie sich die "naive LOHHAUSEN-Vp" von einer "erfahrenen LOHHAUSEN-Vp" unterscheidet. Dabei dürfte die zentrale Rolle des Gedächtnisses hervortreten, wie sie auch in Studien über Schachexperten bereits nachgewiesen wurde (z.B. von CHASE & SIMON 1973).

Das Trainingsargument ist übrigens noch aus einer anderen Perspektive heraus interessant: reale Manager etwa werden nicht "aus dem Stand heraus" zu Entscheidungen befugt, sondern erst einem ausführlichen Training unterworfen, weil man sich die Fehler eines "naiven" Managers nicht leisten kann. Dies ist im politischen Kontext anders. Dort gibt es Fälle, wo ein entsprechendes Training über die Tragweite politischer Entscheidungen nicht vorliegt. Konsequenzen eines solchen Verzichts auf Expertise werden uns in politischen Entscheidungen alltäglich vorgeführt.

Überlegungen zum Training komplexer Operatoren werden von RUPPELL & RUSCHSTROER (1984) angestellt: Im Rahmen des von ihnen entwickelten "Curriculums zur Ausbildung der produktiven Intelligenz" werden den Pbn Mikro- und Makro-Operationen komplexer Verarbeitungsprozesse beigebracht, die zu beachtlichen Leistungssteigerungen führen. Inwieweit dieses Training auch zu einem rascheren Aufbau eines brauchbaren subjektiven Kausalmodells über ein dynamisches Simulationssystem führt, wird derzeit untersucht.

Wie der naive Systemkonstrukteur sich ein operatives Abbild verschafft, wird im Detail noch zu klären sein; daß hierbei die von DÖRNER et al. (1983, p.420f) postulierten Dependenz- und Effektanzalysen im Sinne von Kausalanalysen "rückwärts" (= wovon ist x abhängig?) und Kausalanalysen "vorwärts" (= was bewirkt x?) einsetzen, dürfte klar sein. Ob allerdings die Vpn ein "Gespür" für Kovariationen besitzen müssen und "einfach so" Informationen sammeln, wie von DÖRNER et al. (1983, p.425) angenommen, kann geklärt werden, wenn man Annahme 2 untersucht (siehe weiter unten). Die in Annahme 1 formulierte Be-

hauptung, der Systemkonstrukteur bemühe sich um ein operatives Abbild der Prozedurdynamik, wird wohl auch von DÖRNER et al. geteilt, wenn sie schreiben:

"Ein Realitätsausschnitt ist ein Gefüge von kausalen oder stochastischen (= Wahrscheinlichkeits-)Beziehungen zwischen aktiven und passiven Elementen." (1983, p.26).

Unter aktiven Elementen werden dabei solche verstanden, die eine Eigendynamik besitzen und sich partiell selbst determinieren; die Bezeichnung "endogene Variable" (vgl. Kap. 5.1) scheint mir allerdings angemessener. An anderer Stelle wird dafür der Begriff "Selbsttransformation" gewählt, der im Unterschied zur "Fremdtransformation" die Eigendynamik einer Variable kennzeichnen soll. Bezüglich des Abbilds, das sich der Akteur verschafft, schreiben DÖRNER et al. (1983), es enthalte Elemente und deren Relationen sowie Eingangs- (d.h. beeinflussbare) und Ausgangs- (d.h. feststellbare) Variablen. Außerdem wird zugelassen, daß der Akteur sich ein falsches Abbild machen kann, indem er das System entweder falsch, unvollständig oder mit überflüssigen Elementen versehen darstelle (vgl. Annahme 5).

Ebenfalls kompatibel mit Annahme 1 sind OESTERREICHs (1981) Vorstellungen von einem "Handlungsfeld", das aus Handlungen, Wahrscheinlichkeiten und Konsequenzen von Handlungen zusammengesetzt wird. Da von OESTERREICH keine Angaben darüber gemacht werden, wie ein derartiges Handlungsfeld konkret erfaßt werden könnte, füllt die in Annahme 1 getroffene Behauptung, ein in der Zeit ablaufender Prozeß lasse sich in Form von Parametermatrizen der multivariaten AR_k -Prozesse darstellen, diese Lücke.

Auf störende Effekte impliziter Kausalmodelle weisen NISBETT & WILSON (1977) in ihrem "anti-introspektionistischen" Argumentationsgang hin. Eine der Quellen des Mißtrauens gegenüber Verbaldaten sehen diese Autoren darin, daß Vpn Aussagen über Reiz-Reaktions-Zusammenhänge gemäß ihren impliziten Kausalmodellen treffen. In den Fällen, in denen die tatsächlichen Reaktionen den plausiblen Reaktionen entsprechen, könnte man daher durchaus den Eindruck gewinnen, Verbaldaten seien zuverlässige Informationen; richtige Berichte seien dagegen nicht zu erwarten, wenn die Ursachen für eine beobachtete Wirkung im apriori-Kausalmodell des Auskunftgebenden nicht vorgesehen sind. Eine derartige Argumentation betont übrigens die Notwendigkeit der differenzierten Vorwissenenerfassung im Sinne der Erfassung präexperimenteller Kausalmodelle, um den "Überraschungswert" und damit die Schwierigkeit des zu bearbeitenden Systems für die Vp richtig einschätzen zu können.

Der wissenschaftstheoretische Status von Annahme 1 ist ein besonderer, insofern als sie die Voraussetzung darstellt, von deren Erfüllung die folgenden Annahmen ausgehen. Sie als Axiom zu bezeichnen und damit auszudrücken, daß es sich um einen nicht-disponiblen Teil der hier präsentierten Annahmen handelt, scheint mir nicht richtig, zumal auch die nachfolgenden Annahmen nicht als Theoreme gelten können, die aus

diesem Axiom deduzierbar wären. Es handelt sich eher um eine unbeschränkte universelle Hypothese (vgl. GROEBEN & WESTMEYER 1975), die durchaus -genauso wie die folgenden Annahmen- einer empirischen Prüfung unterzogen werden sollte.

Ein Weg zur empirischen Prüfung der Gültigkeit von Annahme 1 könnte darin bestehen, Widersprüche zwischen einem naiven Kausalmodell und den daraus ableitbaren Konsequenzen aufzudecken. So ist z.B. errechenbar, welcher Systemzustand unter den Prämissen eines spezifizierten Modells etwa nach 20 Zeittakten resultiert. Läßt man dies von einem naiven Modellkonstrukteur prognostizieren, gibt es drei Möglichkeiten: (a) die Prognose stimmt, (b) sie stimmt nicht, (c) der Proband sieht sich zu einer derartigen Prognose außerstande. Fall (a) wäre unproblematisch, da er Annahme 1 stützt. Fall (b) und (c) machen dagegen Hilfsannahmen notwendig, so etwa die, daß man einen Ableitungsfehler auf seiten der Vp zulassen müßte bzw. daß Probanden nicht alle möglichen Implikationen eines Kausalmodells überblicken. Eine Variation des Prognosezeitraums könnte darüber Aufschluß geben, wie weit jeweils die Vorausrechnungen mit dem Modell übereinstimmen. Vorausgesetzt wird dabei die prinzipielle Bewußtseinsfähigkeit des subjektiven Modells.

Annahme 2: Der Systemkonstrukteur nimmt die Parameter des AR_k -Prozesses in einer bestimmten Reihenfolge in sein Modell auf. Erhalten die Input-Variablen gleiche Eingabewerte, werden die Parameter für die x-y-Beziehungen in absteigender Reihenfolge gemäß ihrer Stärke in das subjektive Modell aufgenommen. Besitzen bestimmte Input-Variablen keine Varianz (nimmt der Proband dort keine Eingriffe vor), können die entsprechenden Wirkungen dieser Variablen nicht abgeschätzt werden (Sequenzannahme).

Im Unterschied zu Annahme 1 wird hier der Prozeß des Wissenserwerbs fokussiert. Dafür entscheidend sind zwei Aspekte: (a) starke Effekte werden leichter als schwache erkannt -man kann hier durchaus an Maße der Effektstärke sensu COHEN (1977²) denken, die die Stärke einer Beziehung etwa in Form der durch sie aufgeklärten Varianz angeben-, (b) das Erkennen eines Effekts setzt die entsprechende Variation der exogenen Variable voraus.

Man kann sich den Vorgang des Wissenserwerbs über ein zunächst unbekanntes System durchaus analog zum wissenschaftlichen Erkenntnisprozeß vorstellen. Probanden, die an Studien mit computersimulierten Systemen teilnehmen, schlüpfen vorübergehend in die Rolle eines Wissenschaftlers, der experimentierend Eigenschaften des von ihm untersuchten Objekts erkunden soll:

"Die Entdeckung bzw. Überprüfung kausaler Relationen weist zwei Aspekte auf: einen aktiven, der darin besteht, Systeme durch Herstellung ihrer Anfangszustände 'in-Bewegung-zu-setzen' und einen passiven, der darin besteht, daß man das Systemverhalten, ohne es zu stören, beobachtet. Das wissenschaftliche Experiment kann als eine systematische Kombination dieser beiden Komponenten aufgefaßt werden (...)." (SCHULZ, MUTHIG & KOEPLER 1981, p.69).

Je nach Untersuchungsabsichten können bezüglich der beliebigen Eingreifbarkeit in das zu untersuchende System Restriktionen vorgenommen (vgl. etwa das Konzept der Reversibilität in Kap. 5.2) und dadurch das Ausmaß der Experimentiermöglichkeiten eingeschränkt werden.

PITT (1983) hat z.B. ihren Pbn vier chemische Substanzen (Flüssigkeiten) vorgelegt und ihnen die Aufgabe gestellt herauszufinden, welche Substanzkombinationen (unter Beimischung eines Indikators) zu einer bestimmten Färbung führen. Aufgabe war es also, die Zusammenhänge zwischen den Substanzen und der Färbung in Erfahrung zu bringen, wobei sie keine Restriktionen vorgab. In einer Variante dieses von PIAGET stammenden Experiments, das ebenso wie das erste aus Daten des lauten Denkens von Novizen unterschiedlichen Alters sowie Experten Hinweise über die Existenz von 24 postulierten Subroutinen geben sollte, wurde die Prozedur leicht verändert: Mittels vier Testlösungen sollte die Zusammensetzung einer unbekanntes Flüssigkeit identifiziert werden, wobei diesmal eine Beschränkung des freien Experimentierens insofern vorgenommen wurde, als die Reihenfolge der Prüfung eine entscheidende Rolle spielte (vgl. die Ähnlichkeit zu dem in Kap. 5.2 beschriebenen Käferzuchtungsproblem von DÖRNER bzw. PUTZ-OSTERLOH).

Daß Effekte nur erkannt werden können, wenn die entsprechenden ursächlichen Variablen Varianz besitzen, dürfte trivial erscheinen. Im Bereich exogener Variablen erzeugt der Proband diese Varianz durch seine Eingriffe. Endogene Variablen können per definitionem nicht vom Pbn direkt festgelegt werden; dennoch kann ein Avv-Teilprozeß durchaus erkannt werden, ohne daß ein aktives Eingreifen des Pbn erforderlich wäre.

Die Sequenzannahme 2 ist an eine Voraussetzung geknüpft, die sehr restriktiv formuliert ist: Die Input-Variablen müssen gleiche Ausprägungen bzw. genauer: gleiche Varianz zwischen den Zeitpunkten t und $t+1$ besitzen, um die Annahme einer leichteren Identifizierbarkeit von Variablenbeziehungen mit stärkeren Parametern postulieren zu können. Nun kann man gerade dies bezweifeln, zumal eine optimale Systemidentifikation anderes verlangt (vgl. die Arbeiten von ARIMOTO & KIMURA 1971; MEHRA 1974a, b; zur Frage der optimalen Regulation nichtlinearer stochastischer Systeme: GUSTAFSON & SPEYER 1975). Über den Prozeß der Festlegung von Inputsignalen zum Systemtest ist bislang jedoch nichts bekannt. Im Vorgriff auf These 6, die die Rolle des kontextspezifischen Vorwissens herausstellt, kann aber vermutet werden, daß Vorwissen auch bei diesem Prozeß eine wesentliche Einflußgröße darstellt.

Annahme 3: Ist der Grad der Differenzgleichungen größer Eins (liegen also zeitverzögerte Wirkungen vor), werden zunächst die Parameter des AR_1 -Prozesses in das subjektive Modell aufgenommen, dann die Parameter des AR_2 -Prozesses usw. Oberhalb eines bestimmten Grades der Differenzgleichungen gelingt eine subjektive Repräsentation nicht mehr (Zeitverzögerungsannahme).

Für diese Annahme besteht keine direkte empirische Evidenz, aber die Arbeiten von ROUSE (1973, 1977) und VAN HEUSDEN (1980) lassen

derartiges vermuten. In den genannten Arbeiten von ROUSE hatten Probanden Punkte einer Zeitreihe vorherzusagen, deren Ordnungsgrad zwischen zwei und vier variierte; bei den von VAN HEUSDEN ausgewählten Systemen handelte es sich um Differenzgleichungen dritter Ordnung. Da die Untersuchungsabsichten dort andere waren als die hier verfolgten, können die Resultate dieser Arbeiten nicht direkt zur Stützung von Annahme 3 beitragen.

Daß man als Systemkonstrukteur zunächst von unverzögerten Ursache-Wirkungs-Verhältnissen ausgeht, dürfte einsichtig sein. Viele der alltäglich zu beobachtenden Prozesse besitzen eine derartige Wirkcharakteristik, obwohl es gerade im technischen Bereich Prozesse gibt, bei denen zeitverzögerte Effekte bekannt sind, z.B. das langsame Erwärmen eines Heizkörpers oder einer Herdplatte, die Reaktion eines Computers auf Befehle im Multi-User-Betrieb usw. Es erscheint sinnvoll anzunehmen, daß Menschen zunächst einfache Kausalstrukturen annehmen und prüfen, ehe sie - aufgrund von Mißerfolgen oder anderer Erfahrungsquellen - ihre Modelle um zeitverzögerte Effekte bereichern (diese Annahme läßt sich übrigens durch einen Blick auf psychologische Theorienbildungen erhärten). Daß es hierbei keine beliebigen Ausdehnungen bezüglich des angenommenen Grades der Differenzgleichungen geben wird, folgt aus der begrenzten Speicherkapazität des menschlichen Gedächtnisses. Entscheidende Rolle spielen das Vorwissen des Systemkonstruktors bzw. die möglicherweise in einer Instruktion enthaltenen Hinweise auf zeitverzögerte Effekte; eine "Voreinstellung" wird man in die durch eine Instruktion spezifizierte Richtung erwarten dürfen.

Annahme 4: Der Systemkonstrukteur beginnt mit der Konstruktion einfacher Modelle. Direkte Wirkungen werden früher als indirekte Wirkungen angenommen, deterministische Modelle früher als stochastische konzipiert (Ökonomieannahme).

Die Annahme von indirekten Wirkungen (vgl. Kap. 5.1) gehört aus einem Ökonomieaspekt heraus genauso wenig wie die Annahme stochastischer Beziehungen zu dem, was der naive Systemkonstrukteur spontan erwartet. So wächst die Zahl möglicher indirekter Wirkungen schon bei kleiner Variablenzahl erheblich über das Vorstellbare hinaus. Ganz im Sinne des Auffindens eines optimalen Prädiktorensatzes aus einer langen Liste potentieller Vorhersagevariablen steht zu Beginn einer Systemidentifikation die Überlegung, ob nicht einfache Haupteffekt-Modelle (d.h. nur direkte Wirkungen) zur Prädiktion reichen. Erst bei deren Ungenügen mögen indirekte Wirkungen vermutet und geprüft werden. Ein "deszendierendes" Vorgehen, das von einem saturierten Modell ausgeht, scheint dagegen unökonomisch.

Ein Beispiel für eine indirekte Wirkung findet sich in Abb. 5.11 ("HAMURABI"), wo die exogene Variable x_3 (Ausmaß des bebauten Ackerlandes) direkt auf y_4 (Saat) und y_5 (Ernte) wirkt, indirekt auf y_1 (Kornspeicher). Ganz im Sinne von Annahme 4 findet man in der Arbeit von GEDIGA, SCHÖTTKE & TÜCKE (1982, p.42) einen entsprechenden Hinweis: Die Probanden wurden nach Abschluß der Simulationsspiele gebe-

ten, die perzipierte Wichtigkeit der Variablen einzuschätzen, und gelangen dabei zu einer relativ guten Übereinstimmung mit einer als "objektiv" bezeichneten Rangreihe. Einzige Ausnahme: "Die Variable Ackerland wurde in ihrer Wertigkeit stark unterschätzt."! Die einzige Variable des Systems, von der wichtige indirekte Wirkungen ausgehen, wird somit fehlergeschätzt!

Da unter den Begriff "indirekte Wirkung" auch die Eigendynamik einer Variable (also der $y_1 y_1$ -Parameter aus der A_{VV} -Teilmatrix) fällt, bedarf es hier einer Präzisierung der obigen Annahme. Wie aus der in Kap. 5.3 geschilderten Fallstudie hervorgeht, war es für die Probanden selbstverständlich, daß die $y_1 y_1$ -Parameter den Wert 1 besaßen, also die Variablen über die Zeit hinweg stabil bleiben. Die im Rahmen der Fallstudie von 1 abweichenden Gewichte der Variablen y_1 und y_3 (vgl. Abb. 5.8f) werden im dritten Durchgang in das Modell aufgenommen. Annahme 4 gilt somit auch für Gewichte ungleich 1; die Vermutung zeitlich stabiler Systemvariablen gehört dagegen zu den "Voreinstellungen", mit denen die Probanden die Systemidentifikation beginnen (vgl. auch Annahme 6).

Eng damit verbunden ist die in Annahme 4 ebenfalls postulierte standardmäßige Konstruktion eines deterministischen Modells, das erst unter bestimmten Bedingungen in ein stochastisches Modell überführt wird. Dafür spricht wiederum zunächst ein Ökonomieprinzip: Die Hinzunahme eines Fehlermodells bedeutet für die subjektive Repräsentation einen erhöhten Aufwand, der erst dann betrieben werden muß, wenn sich deterministische Modelle als "untauglich" erwiesen haben. Hierbei ist zu prüfen, unter welchen Bedingungen das "Ungenügen" (im vorletzten Absatz erwähnt) bzw. die "Untauglichkeit" eines Modells vom naiven Systemkonstrukteur festgestellt wird, eine Frage nach dem Evaluationsprozeß und den Evaluationskriterien. Ganz offenkundig muß man -analog zur wissenschaftlichen Prozedur bei der Gültigkeitsprüfung eines Kausalmodells- auf das Konzept des Modelltests rekurren und untersuchen, welche Form von Modelltest von Probanden gewählt wird. Möglicherweise bildet der Grad der Koinzidenz von Prognose und Erfahrung den Maßstab dafür, wann indirekte Wirkungen oder ein Fehlermodell in das naive Modell aufgenommen werden. Daß bei einem solchen "naiven Modelltest" andere (und möglicherweise wechselnde) Regeln gelten werden als für Tests gemäß methodologischen Konventionen, dürfte anzunehmen sein, zumal die subjektiven Entscheidungsregeln in den seltensten Fällen explizit festgelegt sind.

Annahme 5: Der naive Modellkonstrukteur muß kein exaktes Abbild des zugrundeliegenden Kausalmodells erstellen, um handlungsfähig zu sein. Auch unvollständige Modelle können nützlich sein (Unvollständigkeitsannahme).

Denkt man an umfangreiche Variablensätze im Zusammenhang mit dynamischen Modellen, wird man sich fragen müssen, wie derartige "Gebilde" subjektiv abbildbar sein sollen, wo doch seit MILLER (1956) die magische Zahl "5-7 Einheiten" die Grenze der kurzfristigen Speicherbarkeit markiert. Zum einen ist dagegenzuhalten, daß der Erwerb von Wissen

über komplexe Systeme sehr lange Zeiträume in Anspruch nimmt (also die weitreichenden Langzeitspeicher-Fähigkeiten angesprochen werden), wie man am Beispiel der Ausbildung von Piloten oder Raumfahrern sehen kann, wo ein Überlernen geradezu überlebenswichtig sein mag. Zum anderen ist dagegenzuhalten, daß ein Modell nicht komplett oder akkurat sein muß, um nützlich zu sein (vgl. JOHNSON-LAIRD 1983, p.474)! Selbstverständlich muß der Astronaut nicht notwendig wissen, welche Chips sich in den ihm zur Verfügung stehenden Rechnern befinden; ebensowenig muß er etwas über die Details der Materialien wissen, aus denen sein Raumanzug besteht (z.B. genaue Angaben über Reißfestigkeit, Farbechtheit etc.). Auch unvollständiges Wissen kann somit zu brauchbaren Handlungen befähigen. Mit wachsender Komplexität wachsen die Lücken, die mein Wissen über den Gegenstand bzw. Gegenstandsbereich aufweist; jedoch kann auch ein unscharfes Modell ein brauchbares Modell sein. Daß mit sinkendem Systemwissen auch sinkende Kontrollierbarkeit von Systemprozessen einhergeht, dürfte selbstverständlich sein.

Zu klären wäre, ob nicht die Theorie unscharfer Begriffe ("fuzzy concepts"; vgl. ZADEH 1965, 1973) für die Präzisierung von Annahme 5 nutzbar gemacht werden könnte. Die "fuzziness" kann sowohl bezüglich der Parameter wie auch bezüglich der Variablen eines AR_k -Prozesses bestehen. Dabei muß nicht unbedingt auf die mit der Theorie unscharfer Begriffe oft verbundene "fuzzy logic" zurückgegriffen werden, die sehr präzise Angaben über logische Verknüpfungen unscharfer Begriffe macht. Man könnte sich überlegen, eine Theorie unscharfer AR_k -Prozesse zu entwickeln, die dem Aspekt Rechnung trägt, daß ein Modellkonstrukteur anstelle präziser Parameter oftmals nur sehr grobe Angaben über die Wirkungscharakteristik macht (etwa in der Form "wirkt stark", "schwach", etc.). Dies soll weiteren Arbeiten vorbehalten bleiben.

Annahme 6: Handelt es sich um kontexteingebundene Systeme, werden zu Beginn Parameter, die aus dem Vorwissen des Individuums stammen, als Startwerte für die Stärke und Richtung der Kausalrelationen gewählt. Bei abstrakt (d.h. inhaltsfrei) präsentierten Systemen ist die Parametermatrix bei erstmaliger Konfrontierung mit dem System eine Nullmatrix (Kontextannahme).

Mit dieser Annahme soll der Tatsache Rechnung getragen werden, daß menschliche Problemlöser nicht "ahistorisch" an ein Problem herangehen, sondern dort, wo es ihnen möglich und sinnvoll scheint, ihr Vorwissen einbringen wollen. Für die Präsentation kontexteingebundener Systeme (etwa Kleinbetrieb, Kleinstadt, Raumfahrzeug, Regent, Entwicklungshilfe, Ökosystem etc.) besagt Annahme 6, daß Probanden zu Versuchsbeginn nicht mit Nullmatrizen beginnen, keine irgendwie gearteten Annahmen über Kausalbeziehungen machen bzw. "keine Ahnung" haben, sondern im Gegenteil soviel wie möglich von ihrem (impliziten) Kausalvorwissen aktualisieren und als erste Schätzung der Parameter für die AR_k -Prozesse verwenden. Diese "Startwerte" werden in der konkreten Erfahrung mit den Systemabläufen mehr oder weniger stark modifiziert werden müssen, andernfalls würde man wohl kaum von einem Problem sprechen können: Während man bei den üblichen Problemstellungen dann

von einem Problem spricht, wenn die direkte Überführung des Ausgangs- in den Zielzustand durch eine Barriere verhindert wird, sollte man bei Systemen den Grad der Abweichung der subjektiven Schätzungen der Parameter von deren objektiver Höhe heranziehen. Dadurch läßt sich auch erklären, warum ein- und dasselbe System von zwei Personen als unterschiedlich schwierig empfunden werden kann bzw. wieso bei wiederholtem Durchlauf manche Systeme leichter beherrschbar scheinen.

Natürlich kann nicht behauptet werden, semantisch eingebettete Systeme ließen sich prinzipiell leichter steuern als abstrakt präsentierte; in diese Richtung kann die Arbeit von HESSE (1982) geradezu fehlinterpretiert werden. Gerade die Startwerte, die aus dem Vorwissen geschöpft werden, können sich ja als falsch herausstellen und erfordern daher zusätzliche Arbeit zu ihrer Korrektur, sofern die Diskrepanz überhaupt erst einmal erkannt wurde. Systeme gleicher Struktur sollten somit zu Beginn am schwierigsten zu handhaben sein, wenn die Variablenbeziehungen und deren Stärken kontraintuitiv sind (mit kontraintuitiv ist dabei gemeint: konträr zu den Annahmen, die aus dem Vorwissen des Pbn folgen), weniger schwierig die abstrakt dargebotenen Varianten und noch leichter die Varianten, die dem Vorwissen sehr nahekommen. Dies ist bislang noch nicht systematisch untersucht worden.

In einer neueren Arbeit untersuchen HESSE & GERRARDS (1985) den Zusammenhang zwischen der "Zielstruktur" eines Problemlösers und dem Abruf problemrelevanter Vorwissens. Aus Daten lauten Denkens ermitteln die Autoren, wie häufig formulierte Ziele auf Vorwissen bezogen werden; dies ist bei gut 80% der verbalisierten Ziele der Fall. HESSE & GERRARDS zeigen, daß ein derartiger Bezug hauptsächlich bei konkreten Zielen vorkommt. Ihre Interpretation besagt, daß es zwar eine von den Zielen ausgehende Wissensaktivierung gebe, andererseits dieses Vorwissen wenig "gebrauchsfertige Elemente wie Subziele oder Handlungsmöglichkeiten" enthalte. Hier könnte m.E. eine genauere Untersuchung der subjektiven Kausalmodelle weitere Erkenntnis über den Einfluß von Vorwissen liefern.

Probleme gibt es bezüglich der Diagnostik des Vorwissens, hier verstanden als die zu Versuchsbeginn bereits vorhandenen impliziten Kausalmodelle einer Person. Traditionell kann man einmal einen reichsspezifischen "Wissenstest" vornehmen, der den deklarativen Anteil des Vorwissens zugänglich macht, zum anderen könnten auch prozedurale Wissensanteile erfragt werden. Problematisch daran ist die Tatsache, daß es unterschiedliche Grade des Verstehens gibt und die eben geschilderte Diagnostik wohl eher die höheren Grade anspricht. GREENO & RILEY (1984) unterscheiden auf der Ebene theoretischen Verstehens (der Kenntnis allgemeiner Prinzipien) neben dem expliziten auch ein implizites Verstehen sowie eine noch niedrigere Stufe, Konformität genannt, auf der ein Verfahren "bloß gekannt", nicht aber verstanden wird. Eine Vorwissensdiagnostik müßte solchen unterschiedlichen Verstehensebenen Rechnung tragen.

Auch das Problem der möglichen Reaktivität einer Vorwissensdiagnostik ist zu beachten: Je präziser Fragen über ein implizites Kausalmodell gestellt werden, umso weniger läßt sich vermeiden, daß erst in der aktuellen diagnostischen Situation ein Modell von der Vp konstruiert bzw. modifiziert wird, das in dieser Form vorher nicht existierte.

Das Problem der Vorwissensdiagnostik ist letztlich ein Problem der Wissensdiagnostik ganz allgemein: Welche Möglichkeiten bestehen, die kognitive Struktur eines Individuums hinsichtlich eines ausgewählten Realitätsbereichs adäquat zu erfassen? OLDENBÜRGER (1981) spricht in diesem Zusammenhang davon, daß schlußendlich immer nur eine kognitive "Strukturierung" als konkrete Realisation einer prinzipiell unzugänglichen kognitiven Struktur erwartet werden darf. Sicher kann eine Wissensdiagnostik nur sinnvoll in Anlehnung an ein spezifisches Modell der Wissensrepräsentation erfolgen.

Annahme 7: Im Sinne gedächtnispsychologischer Theorien und Befunde muß auch für den Prozeß der Kausalmodellkonstruktion ein "Auffrischung-" bzw. "fading"-Effekt postuliert werden (Gedächtnisannahme).

Mit dieser Annahme soll berücksichtigt werden, daß menschliche Modellkonstrukteure Parameter vergessen können bzw. daß diese einem "fading"-Prozeß unterliegen, den man seinerseits mathematisch modellieren kann. ROUSE (1977) spricht von "estimators that forget" und zeigt, wie eine stärkere Gewichtung neuerer gegenüber älterer Systeminformation mathematisch formuliert werden kann (vgl. ROUSE 1977, p.277f). Darauf wollen wir hier nicht ausführlicher eingehen; dennoch ist festzustellen, daß gedächtnispsychologische Einteilungen wie die in Kurz- und Langzeitgedächtnis voll kompatibel mit den hier vorgestellten Annahmen sind, ja sogar im Modell von ROUSE (1977) einen zentralen Stellenwert erhalten: Von den drei Parametern β_1 , β_2 und α seines Modells stehen β_1 für den "short-term memory"-Parameter, β_2 für den "long-term memory"-Parameter und α für den Austausch zwischen beiden Gedächtnisformen. Während die Gedächtnisparameter Beschränkungen der menschlichen Fähigkeiten entsprechen, reflektiert der Austausch-Parameter vermutlich den aktuellen Entscheidungsprozeß bei der Prognose künftiger Systemzustände. Annahmen über den "guten" oder "schlechten" Systemmodellierer lassen sich durch unterschiedliche Parameterkonstellationen im Modell von ROUSE (1977) treffen.

5.6 Abschließende Bemerkungen zum vorgestellten Rahmenkonzept

Die in Kap. 5.5 vorgestellten sieben Annahmen eines Rahmenkonzepts zur Konstruktion von Kausalmodellen über die von Pbn zu steuernden dynamischen Systeme werfen einige Fragen (z.B. nach den anthropologischen Voraussetzungen) auf, die noch einmal kurz erläutert werden sollen. Außerdem wird der Versuch unternommen, eine Bewertung des vorgeschlagenen Ansatzes dahingehend vorzunehmen, inwiefern er einer-

seits kritisierte Schwächen bisheriger Ansätze überwindet und inwiefern sich einige der in dieser Arbeit geschilderten Alternativ-Konzeptionen über das Handeln in komplexen Situationen mit den von mir getroffenen Annahmen vereinbaren lassen.

Zunächst ein paar Worte zu den offenen Fragen der Annahmen über die (Re-)Konstruktion von Kausalmodellen. Es mag der Eindruck entstehen, Menschen müßten -wenn sie sich gemäß Modellannahmen verhielten- umfangreiche Rechenoperationen zur Konstruktion und Evaluation von Modellen durchführen, die weit über die begrenzte Kapazität ihrer Verarbeitungsstrukturen hinausgingen. Dies ist -wie in Annahme 5 zu zeigen versucht wurde- nicht notwendig der Fall. Unter Rekurs auf die Vorstellungen über "fuzzy sets" wird behauptet, daß komplexe Strukturen nicht notwendig präzise repräsentiert werden müssen. Der sich daraus ergebende Widerspruch zu Annahme 1, nach der das operative Abbild des Modellkonstruktors in Form von Parametern eines multivariaten AR_k -Prozesses darstellbar sein müsse, ist nur ein scheinbarer: Es wird weder postuliert, daß die subjektiven Strukturgleichungen mit den objektiv vorliegenden in allen Punkten übereinstimmen müssen, noch darf man davon ausgehen, daß überhaupt alle relevanten Strukturgleichungen identifiziert werden. Schließlich muß man einwenden, daß die Repräsentation komplexer Systeme in hohem Maß den Langzeitspeicher beansprucht, also dessen Kapazitätsvorteil nutzt.

Bezüglich weiterer anthropologischer Voraussetzungen möchte ich mich an dieser Stelle dem funktionalistischen Standpunkt von JOHNSON-LAIRD (1983) anschließen. Dieser Autor, der Bewußtsein (consciousness) als "the most puzzling of all the phenomena of mental life" (JOHNSON-LAIRD 1983, p.448) bezeichnet, läßt Bewußtsein prinzipiell einer der vier folgenden Kategorien angehören:

- (1) Es handelt sich bei Bewußtsein um ein übernatürliches Phänomen ("the ghost in the machine").
- (2) Es ist nur unter Rekurs auf bisher unbekannte und nicht berechenbare Prozeduren erklärbar.
- (3) Bewußtsein läßt sich -ähnlich wie Wetter- durch Theorien erklären, die in Computerprogrammen realisiert werden können, aber genausowenig, wie man einen Anti-Zyklon in einem Rechner entstehen lassen kann, läßt sich einem Computer Bewußtsein einhauchen (nur Organismen mit Gehirn können also Bewußtsein besitzen).
- (4) Auch Computer können so programmiert werden, als ob sie Bewußtsein besäßen.

Nur die beiden letztgenannten Möglichkeiten hält JOHNSON-LAIRD überhaupt zugänglich für und entscheidbar durch wissenschaftliche Untersuchungen. Nach seinem Standpunkt hängt Bewußtsein von "computations of the nervous systems" (p.450) ab; dieses Nervensystem muß über ein Gehirn bestimmter Größe und Komplexität verfügen. Da unser Gehirn als ein "real time processor" betrachtet werden muß (d.h. es ist ständig präsent und arbeitet fortlaufend mit neu eintreffenden bzw. abgerufenen Informationen), braucht es zum einen "computational power", die vom Gedächtnis abhängig ist, zum anderen "computational

speed", die die Größe der Einheiten bedingt, die simultan verarbeitet werden können.

Wenn man Fähigkeiten wie Spracherwerb und Sprachgebrauch oder schlußfolgerndes Denken betrachtet, die ein menschliches Gehirn scheinbar "mit der linken Hand" erledigt (und dabei trotz einiger falscher Schlußfolgerungen Unangemessenheiten aussagenlogischer Formalisierungen überwindet), dürften die Annahmen des vorigen Abschnitts weniger kurios erscheinen als sie es beim ersten Lesen vielleicht taten. Da oftmals die "Schwächen" menschlichen Schlußfolgerns hervorgehoben werden (z.B. Atmosphären- und Figur-Effekte; vgl. JOHNSON-LAIRD 1983, p.69ff), sollen hier einige "Stärken" genannt sein, die Zweifel daran aufkommen lassen, ob die Schwächen wirklich Schwächen der menschlichen Vernunft oder nicht vielmehr mangelnde Angemessenheit der zugrundegelegten normativen Modelle darstellen. Dies soll an zwei Beispielen von JOHNSON-LAIRD (1983) verdeutlicht werden. Nimmt man etwa folgende Prämissen:

- (a) mehr als die Hälfte aller x ist y ,
- (b) mehr als die Hälfte aller x ist z ,

so folgt nach aussagenlogischen Regeln (vgl. VON KUTSCHERA & BREITKOPF 1971) daraus nichts, da ein Operator "mehr als die Hälfte" dort nicht bekannt ist. Dennoch ist klar, daß man aus obigen Prämissen den - in bezug auf die Mengentheorie korrekten - Schluß ziehen würde:

- (c) mindestens ein y ist z .

Ein zweites Beispiel für die Stärken menschlichen Schlußfolgerns zeigt sich an der Unterschiedlichkeit der folgenden Schlüsse:

- | | |
|----------------------------|--|
| (a) John ist Schotte | (a') John ist Schotte |
| (b) Schotten sind Menschen | (b') Schotten sind über die Erde verstreut |
| ----- | |
| (c) John ist Mensch | (c') ? |

Die verwendeten Attribute "Mensch" bzw. "über die Erde verstreut" beziehen sich einmal auf jedes Mitglied der Klasse, zum anderen nur auf die Klasse insgesamt. Diesen Unterschied erkennt wiederum ein vernünftiger Bearbeiter dieser Prämissen auf Anhieb.

Überlegungen dazu, wie sehr die vorgeschlagenen Bausteine einer Theorie der Konstruktion kausaler Modelle über dynamische Prozesse mit Befunden vereinbar sind, nach denen das Verhalten eines menschlichen Individuums nicht im Einklang zu bestimmten normativen Modellen steht, müssen auch die Frage klären, inwieweit die dargelegten Annahmen die Theorie multivariater AR_k -Prozesse als normatives Modell verwenden. Die deskriptive Verwendung dieser Theorie zur Beschreibung eines Untersuchungsgegenstandes dürfte unbestritten sein; eine normative Verwendung liegt insofern vor, als behauptet wird, die optimale Systemmanipulation sei ohne Einsatz des vollständigen Instrumentariums multivariater AR_k -Prozesse nicht möglich. Da jedoch keineswegs postuliert wird, daß der naive Konstrukteur eines Kausalmodells optimal arbeiten

müsse, sondern im Gegenteil eine Reihe von Abweichungen gegenüber einem derartigen Optimum toleriert werden (und deren Ursachen erklärt werden), besteht m.E. nicht die Gefahr, zur langen Liste von Beispielen, wie sich Menschen nicht bestimmten formalen Modellen gemäß verhalten, ein weiteres Beispiel hinzuzufügen. Dies wird weiter unten durch die Diskussion des Konzepts "Lösungsgüte" noch ausgeführt.

Doch nun soll auf Parallelen des eigenen Ansatzes zu anderen Ansätzen hingewiesen werden. Auf die Ähnlichkeiten zum "operativen Abbildsystem" von HACKER (1978²) habe ich bereits hingewiesen; ergänzend kann nunmehr gesagt werden, daß meine Annahmen wesentlich stärker auf den Konstruktionsprozeß und die ihn begünstigenden bzw. erschwerenden Bedingungen eingehen sowie Aussagen über Kontexteffekte machen, die bei HACKER fehlen.

Auch zu OESTERREICHs (1981) Überlegungen passen die Annahmen über Kausalmodelle bzw. tragen zu deren Präzisierung bei, sofern es um die Situationsklasse "dynamisches System" geht. Im Rahmen der sieben Annahmen ist klar, daß der Problemlöser sowohl die Effizienz als auch die Divergenz einer Variablenkonstellation prüfen muß; die Frage, die sich stellt, ist, ob es sich dabei nicht um ein und denselben Feststellungsprozeß handelt, nämlich das Herausfinden von Parametern eines autoregressiven Prozesses. Ein wichtiger Unterschied zwischen OESTERREICHs Vorstellungen eines Handlungsfelds und meinen Annahmen liegt in der Bedeutung der Gewichtungen, die Handlungen mit Konsequenzen bzw. x- mit y-Variablen verbinden: Bei OESTERREICH sind es Wirkwahrscheinlichkeiten, die sich für jede Handlung zu 1 aufaddieren, während es im Kausalmodell-Ansatz um Wirkstärken geht, die keinen Restriktionen genügen müssen. Während bei OESTERREICH die verschiedenen Handlungsmöglichkeiten jeweils kategorial verstanden werden, also mit der Wahl einer Alternative nicht die Stärke der Maßnahme variieren kann, ist im Kausalmodell-Ansatz natürlich ein abgestuftes Mehrfachhandeln (vgl. FUHRER 1982) möglich. Man kann nicht nur jeweils mehr als eine Alternative pro Zeitpunkt wählen, sondern deren Stärke darf darüberhinaus beliebig ausfallen. Damit wird klar: Die OESTERREICH'sche Konzeption eines Handlungsfeldes kann als ein Spezialfall eines Kausalmodells aufgefaßt werden, es handelt sich also nicht um konkurrierende Vorstellungen.

HAYES-ROTH & HAYES-ROTH (1979) machen im Rahmen ihres Planungsmodells Annahmen über hierarchische Plan-Strukturen. Über diese geben die Annahmen zu Kausalmodellen nur am Rande Auskunft; genauso wenig werden Intention und Emotion des Systemkonstruktors berücksichtigt, Komponenten, die in DÖRNERs (1983c) Absichtsregulationsmodell zentralen Stellenwert einnehmen. Während damit ein eher generelles Rahmenmodell über die zeitliche Abfolge von Handlungsentscheidungen vorliegt, geht es im Kausalmodell-Ansatz wesentlich konkreter um eine bestimmte Problemklasse, nämlich die subjektive (Re-)Konstruktion dynamischer Systeme. Daher kann man hier nicht von konkurrierenden Vorstellungen sprechen.

Nach diesem Vergleich der sieben Annahmen zur Konstruktion subjektiver Kausalmodelle mit einigen handlungsorientierten Ansätzen soll zum Schluß noch einmal der Bogen zur kritischen Bestandsaufnahme gespannt werden. Insbesondere ist zu fragen, inwiefern erkannte Defizite bisheriger Forschung mit diesem Ansatz überwunden werden können.

Bezüglich der in Kap. 3 formulierten These 1 -konstatierte Theoriearmut- hoffe ich auf die Fruchtbarkeit des hier dargelegten Ansatzes wie auch auf Weiterentwicklungen anderer geschilderter Ansätze. Die mangelnde Integration anderer psychologischer Teildisziplinen (These 2) hat der Kausalmodell-Ansatz bislang nicht überwunden, wohl dagegen die mangelnde fachübergreifende Kooperation (These 3). Der systemtheoretische Ansatz, mit dem sich These 4 beschäftigt, wird genutzt, wenn man die mit Annahme 5 (Komplexitätsannahme) verbundene Möglichkeit von komplexitätsreduzierenden Maßnahmen untersucht: Das zu Beginn der Systemkonstruktion notwendige "Erschließen" des "Gerippes" impliziert eine Konnektivitätsanalyse auf höherem Niveau, das je nach Bedarf präzisiert werden kann. Das mit These 5 festgestellte Fehlen einer Taxonomie von Systemtypen scheint überwunden, blickt man auf die eingangs dieses Kapitels genannte Möglichkeit, komplexe Systeme in Form ihrer AR_k-Prozesse zu beschreiben. Die eine Typologie begründenden charakteristischen Merkmale solcher Prozesse wurden genannt und könnten in Nachfolgearbeiten als Grundlage einer Systematisierung komplexer Systeme dienen. Die Beschreibung von dynamischen Systemen mittels der ihnen zugrundeliegenden AR_k-Prozesse ist erheblich präziser als die bloße Angabe solcher Charakteristiken wie "komplex", "träge" etc. Die Thesen 6 und 7 (Meßfehler und Versuchsplanung) sind erst im Lichte konkreter Experimente zu beurteilen und können daher nicht zur Beurteilung des vorgelegten Ansatzes herangezogen werden. Auch die Bestimmung der Lösungsgüte (These 8) hängt von der Fragestellung eines konkreten Experiments ab, obwohl die sieben Annahmen zumindest erste Hinweise auf dessen Bestimmbarkeit geben (vgl. auch Kap.6). Hierbei zählt neben der konkreten Systemmanipulation auch das implizite wie explizite Wissen des Problemlösers über die Struktur des Systems. Anstatt den -für diese Art von Problemstellungen unglücklich gewählten- Begriff der "Lösungsgüte" zu benutzen, sollte man eher vom Ausmaß des Systemwissens und der Systemkontrolle bzw. vom Grad der Annäherung des naiven Kausalmodells an das tatsächlich vorliegende Modell sprechen.

Wie eine "Diagnostik des theoretischen wie praktischen Systemwissens" im Detail aussehen mag, kann momentan noch nicht angegeben werden. Daß eine solche Diagnostik notwendig ist, steht dagegen fest. Der im vergangenen Jahrzehnt wiederentdeckte Zugang über Daten lauten Denkens sollte dabei durchaus skeptisch betrachtet werden. Nimmt man etwa mit JOHNSON-LAIRD (1983, p.468f) ein paralleles Prozessieren von Informationen auf mehreren hierarchischen Ebenen an, darf die durch lautes Denken bedingte Störung solcher Parallelverarbeitung ebenso wenig in Kauf genommen werden wie die absolut unzureichende Menge der so produzierten Information. Die bereits eben vorgelegte Bewußtseinskonzeption von JOHNSON-LAIRD (1983) erzwingt geradezu notwendig die Existenz unbewußt ablaufender Prozesse auf niedrigeren Verarbeitungs-

stufen, die bei einem derartigen methodischen Zugang unzugänglich blieben. Hier muß eine Theorie der Wissensrepräsentation -eine Gedächtnistheorie also- Abhilfe schaffen, in der subjektive Kausalmodelle einen entsprechenden Stellenwert erhalten werden.

6 Abschluß: Ein Blick nach vorne

In den vorangegangenen Kapiteln hat der Leser aktuelle Arbeiten zum Umgang von Personen mit komplexen, computersimulierten Realitätsbereichen näher kennengelernt. Daß dabei nicht nur Erfreuliches, sondern auch Problematisches zum Vorschein kam, verwundert nicht angesichts der Materie, um die es hier geht. Daher kann auch kein Schlußstrich gezogen werden, der etwa im Sinne gesicherter Erkenntnisse zu interpretieren wäre. Dafür ist dieses Forschungsgebiet noch zu jung und die Aktivitäten verschiedener Forschungsgruppen zu zahlreich, um nicht neue Erkenntnisse erwarten zu lassen. So sollen abschließend -anstelle einer Zusammenfassung- drei m.E. zentrale Aspekte behandelt werden:

- a) eine Diskussion des Rahmenmodells, das den hier beschriebenen Arbeiten in bezug auf die Untersuchungssituation zugrundegelegt werden kann,
- b) eine Heuristik für das Verständnis von Vp-Verhalten beim Umgang mit einem computersimulierten Szenario,
- c) eine Systematisierung der Forschungsaufgaben, die sich in diesem Kontext stellen.

6.1 Rahmenmodell

In einem neueren Bericht schlagen OPWIS, SPADA & SCHWIERSCH (1985, p.4) einen konzeptuellen Rahmen für psychologische Untersuchungen zum Umgang von Personen mit Simulationssystemen vor. Über einen ausgewählten Realitätsbereich (RB) konstruiert der Forscher (Fo) ein Simulationsmodell 1 (SM1Fo), mit dem die Vp umgeht und über das sie ein mentales Modell (mMVp) konstruiert. Dieses mMVp ist der eigentliche Untersuchungsgegenstand, für den ein SM2Fo (= das psychologische Modell des Forschers über das mentale Modell der Vp) konstruiert wird. Dies scheinen mir jedoch nicht alle Modelle zu sein, die hier eine Rolle spielen. Gibt man in Klammern jeweils an, worauf sich das Modell bezieht, wird die Möglichkeit endloser Schleifen deutlich (vgl. Abb. 6.1).

Aus Abb. 6.1 geht hervor, daß mindestens sechs verschiedene Modelle im Spiel sind: Auf seiten der Vp ihr mentales Modell über den Realitätsbereich wie auch ihr mentales Modell über das ihr vorgelegte Simulationsmodell des Forschers, auf seiten des Forschers zusätzlich zu den zwei von OPWIS et al. (1985) angenommenen Modellen noch sein eigenes mentales Modell über den Realitätsbereich (sein Vorwissen) sowie sein Modell des mentalen Modells der Vp über den Realitätsbereich, also sein Modell des Vorwissens der Vp.

Das Problem besteht nicht nur darin, daß die Verschachtelungstiefe der Modelle unbekannt ist (gibt es etwa ein mentales Modell der Vp über das mentale Modell des Forschers über das mentale Modell der Vp über den Realitätsbereich, usw.), sondern daß auch unklar ist, zwischen welchen Modellen welche Art von Beziehung besteht (wie verträg-

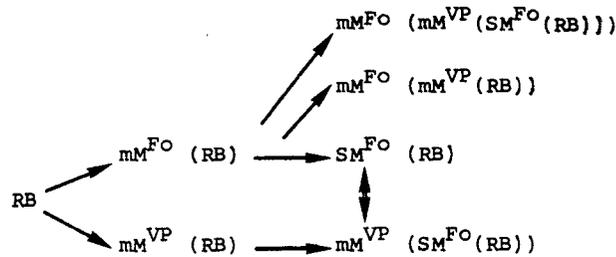


Abb. 6.1: Verschiedene "Modelle" bei der Untersuchung des Umgangs von Personen mit Simulationssystemen (RB=Realitätsbereich, mM=mentales Modell, SM=Simulationsmodell, Fo=Forscher, Vp=Versuchsperson).

lich ist etwa das mentale Modell einer Vp über einen Realitätsbereich mit dem mentalen Modell der Vp über die Simulation des Realitätsbereichs, usw.?). Faßt man die jeweiligen Modelle als strukturierte Mengen von Konzepten auf, stellt sich somit die Frage nach der Art von Morphismus, die zwischen je zwei Mengen besteht. Dies entspricht z.B. der Validitätsfrage, wo es um die elementare Beziehung zwischen Realitätsbereich und Simulationsmodell geht (vgl. PAGE 1983). Es dürfte jedoch klar geworden sein, daß es sich dabei nicht um die einzige Beziehung handelt, die diskussionsbedürftig ist.

Um diese Validitätsproblematik jedenfalls partiell zu entschärfen, gibt es also zwei Möglichkeiten: (a) Man versucht unter Einbezug fachwissenschaftlicher Befunde eine möglichst naturalistische Simulation des Gegenstandsbereichs (vgl. etwa die Berichte über den neuen Fahr Simulator von Daimler-Benz), oder (b) man verzichtet auf den Anspruch der Realitätsnähe und weist die Pbn explizit auf die "Fiktion" der Simulation hin (etwa durch Verlegen der Szenarien auf einen anderen Planeten mit eigenen Gesetzmäßigkeiten). Da der Aufwand für die Alternative (a) ungleich höher ist und darüberhinaus die psychologischen Fragestellungen Vorrang haben sollten, ist eine Entscheidung für Alternative (b) häufig naheliegender.

6.2 Eine Heuristik zum Verständnis des Vp-Verhaltens

Die Aufgaben und Probleme, die sich einer Vp beim Umgang mit einem computersimulierten System stellen, sind vergleichbar denen eines Forschers, der Hypothesen über einen ihn interessierenden Realitätsbereich aufstellt (mit dem winzigen, aber entscheidenden Unterschied, daß die tatsächlich zugrundeliegende Systemstruktur nie definitiv offengelegt sein wird, während sie uns bei einem simulierten Szenario bekannt ist). Von daher scheint es angebracht, den Prozeß der Modellbildung im wissenschaftlichen Bereich als Analogie für die Modellbildung der Vp in einer Simulationsstudie zu verwenden. BOX & JENKINS

(1976) haben einige wichtige Stufen genannt, die durchlaufen werden müssen (siehe Abb. 6.2).

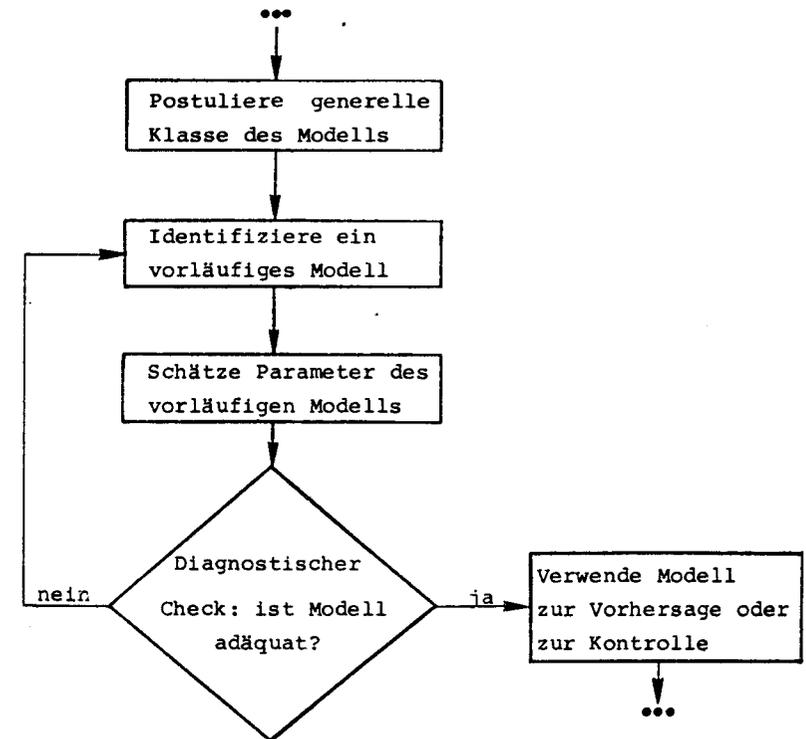


Abb. 6.2: Der iterative Zugang zur Modellbildung (nach BOX & JENKINS 1976, p.19).

Zunächst ist eine generelle Klasse von Modellen zu postulieren. Auf unsere Situation übertragen mag dies z.B. bedeuten, daß sich ein Pb allgemeine Vorstellungen über die Art der zu erwartenden Zusammenhänge macht. Die Identifikation eines vorläufigen Modells könnte darin bestehen, die vermuteten wichtigen Variablen in ein Kausalverhältnis zu setzen (Aufstellen "subjektiver Strukturgleichungen"). Für die jeweiligen Prozesse erfolgen dann Parameterschätzungen im Sinne "subjektiver Regressionsgewichte". Im Umgang mit dem Simulationssystem erfolgt schließlich der diagnostische Check auf Adäquatheit des Modells sowie die Verwendung eines akzeptierten Modells zur Vorhersage bzw. Kontrolle der ablaufenden Prozesse.

6.3 Systematisierung der Forschungsaufgaben

DÖRNER et al. (1983, p.114) haben seinerzeit vier "Generalfragen" gestellt, um die "Art und Weise, wie intelligente Individuen mit Unbestimmtheit und Komplexität umgehen", aufzuklären:

1. Welche Effekte erzeugen verschiedene Vpn im Umgang mit einem komplexen und zunächst teilweise unbekanntem System? ('Effektfrage')
2. Auf welche Merkmale der Denk-, Planungs- und Entscheidungsprozesse der Vpn sind die erzielten Effekte zurückführbar? ('Prozeßfrage')
3. Welches sind die situativen und in der Persönlichkeit der Vpn liegenden Bedingungen für die Art des Denkens, Planens und Entscheidens? ('Bedingungsfrage')
4. Von welcher Beschaffenheit muß ein theoretisches System sein, damit es in die Lage versetzt wird, das Handeln von Menschen in Unbestimmtheit und Komplexität in Abhängigkeit von Situations- und Persönlichkeitsmerkmalen prognostizieren zu können? ('Theoriefrage')" (DÖRNER et al. 1983, p.115).

Für eine umfassende Betrachtung des Problemlösens in komplexen, computersimulierten Realitätsbereichen sollte man diese Fragen neu formulieren, um die entscheidenden Einflußgrößen zu systematisieren. Analog zur Situation der Psychotherapie-Forschung, die durch KIESLERS (1971) Warnung vor dem Mythos von der Homogenität der Patienten und von der Uniformität der Therapeuten einen Weg zur feldexperimentellen Untersuchung einer sehr komplexen Materie gefunden hat, könnte man an dieser Stelle vor dem Mythos der "Homogenität des Vorwissens" und der "Uniformität der Systeme" warnen und die differentielle Sichtweise der gegenwärtigen Kognitionsforschung offenlegen. Aktuelle Überlegungen zur Wissensdiagnostik (SCHEELE & GROEBEN 1984, TERGAN 1984) belegen die Bedeutsamkeit des erstgenannten, systemtheoretische Analysen (LÜER & HUBNER 1985) die des letztgenannten Aspekts.

Zu den wesentlichen Bestimmungsstücken, die den Umgang mit komplexen, computersimulierten Realitätsbereichen determinieren, sind die in Tab. 6.1 aufgeführten Merkmale zu rechnen (vgl. dazu die Überlegungen von HUSSY 1984b zur Klärung des Konzepts "Problemschwierigkeit").

Im ersten Bereich -Personenmerkmale- kommen die je individuellen Voraussetzungen für den Umgang mit einem dynamischen System zum Vorschein. Die in Tab. 6.1 nur stichwortartig angegebenen Bereiche müssen adäquat operationalisiert werden, wobei unter Umständen eine einmalige Messung nicht ausreichen kann. Gerade im Bereich der Zustandsvariablen ist mit erheblichen anforderungsabhängigen Schwankungen zu rechnen. Diese "on-line" zu erfassen und zur aktuellen Bearbeitung eines komplexen Problems in Bezug zu setzen, dürfte erhebliche Schwierigkeiten mit sich bringen. Gerade die leicht zu erhebenden physiologischen Belastungsparameter leiden unter mangelnder Eindeutigkeit und können nur dann fruchtbar verwendet werden, wenn aus der zugrundegelegten Theorie klare Prognosen über den Verlauf bzw. die Veränderung solcher Größen folgen.

Tab. 6.1: Determinanten des Umgangs mit komplexen, computersimulierten Realitätsbereichen.

-
1. PERSONENMERKMALE
 - 1.1 Traits: Gedächtnis, Wissen, Intelligenz, kognitive Kompetenzen (z.B. heuristische, epistemische), metakognitive Kompetenzen (z.B. Selbstreflexion), Selbstsicherheit
 - 1.2 States: Emotion, Konzentration, Kompetenzeinschätzungen
 2. SYSTEMMERKMALE
 - 2.1 Statische Merkmale: Variablenzahl, Vernetztheit
 - 2.2 Dynamische Merkmale: Stabilitätsverhalten, Sensitivität, Kontrollierbarkeit
 3. MERKMALE DER UNTERSUCHUNGSSITUATION
 - 3.1 Systempräsentation: Art (grafisch, numerisch, sprachlich), Umfang (ein Simulationstakt vs. Zeitreihe)
 - 3.2 Transparenz: Zugänglichkeit zu statischen/dynamischen Systemdaten, Bekanntheit der Bewertungskriterien, Art der Zielvorgabe
 - 3.3 Weitere äußere Merkmale: Individual- vs. Gruppenbearbeitung, VI-Einfluß
-

Hinsichtlich des zweiten in Tab. 6.1 angesprochenen Bereichs von Einflußgrößen -Systemmerkmale- muß bereits vor der Datenerhebung klar sein, welche Anforderungen das Simulationssystem im Detail stellt und wie es sich unter verschiedensten Eingriffsvarianten verhält. Hierüber sollte der Experimentator größtmögliche Kenntnis und damit auch Kontrolle besitzen. Aus diesem Grund ist auch eine wesentlich sorgfältigere Dokumentation und Publikation der verwendeten Untersuchungsinstrumente zu fordern, damit eine möglichst klare Attribution von Versuchsergebnissen möglich wird. Gerade die Tatsache, daß gewählte Indikatoren der Leistungsgüte vielfach einfachsten teststatistischen Kriterien nicht genügen, sollte voreiligen Interpretationen vorbeugen und den Untersucher dazu auffordern, die Möglichkeit alternativer Deutungen in Betracht zu ziehen.

Die Merkmale der Untersuchungssituation -der dritte Bereich von Einflußgrößen- sind vielfältig variierbar und je nach Untersuchungsabsicht werden noch andere Aspekte hinzukommen. Man sollte sich jedoch darüber im Klaren sein, daß die Wirkungen einer ganzen Reihe derartiger Variablen noch nicht hinlänglich bekannt sind. Auch dies läßt eine sehr genaue Beschreibung der Untersuchungssituation als unerlässlich erscheinen.

Ein Forschungsprogramm, das diesen Gesichtspunkten Rechnung tragen will, muß notwendig aus einer Vielzahl von Teilprojekten bestehen, die jeweils einzelne dieser Aspekte herausgreifen und deren Auswirkungen systematisch untersuchen. Neben den Personenmerkmalen, zu denen inzwi-

schen eine Reihe von Informationen vorliegen (ich beziehe mich hier auf die Befunde, die aus dem Vergleich der post facto in "Erfolgreiche" und "Erfolgslose" eingeteilten Vpn gewonnen wurden - wobei das Erfolgskriterium ja durchaus problematisiert werden muß), sind auch Systemmerkmale neuerdings in ihrer Wirkung untersucht worden. Auch zu einzelnen Aspekten der Untersuchungssituation liegen Ergebnisse vor. Jedoch scheinen die bisher vorgelegten Arbeiten noch weit von einer systematischen Vorgehensweise entfernt. Stellt man sich einen aus den drei genannten Dimensionen gebildeten Würfel vor (dessen einzelne Zellen wiederum unterteilbar sind), zeigen sich gegenwärtig noch viele "leere" Teil-Würfel, zumal mit erheblichen Wechselwirkungen gerechnet werden darf. Die Bedeutung der genannten Wirkgrößen und deren gegenseitige Abhängigkeiten werden sich in zukünftigen Arbeiten erhellen müssen, will man eines Tages eine befriedigende Antwort auf die differentiellen Fragen der Denkforschung geben können.

Literatur

- AEBLI, H. 1980. Denken: Das Ordnen des Tuns. Band I: Kognitive Aspekte der Handlungstheorie. Stuttgart: Klett-Cotta.
- AEBLI, H. 1981. Denken: Das Ordnen des Tuns. Band II: Denkprozesse. Stuttgart: Klett-Cotta.
- ARIMOTO, S. & KIMURA, H. 1971. Optimum input test signals for system identification - an information-theoretical approach. *International Journal of Systems Science* 1, 279-290.
- ASHBY, W.R. 1958. Requisite variety and its implications for the control of complex systems. *Cybernetica* 1, 83-99.
- ASHBY, W.R. 1972. Systems and their informational measures. In: KLIR, G.J. (Ed.), *Trends in general systems theory*. New York: Wiley, 78-97.
- ASTROM, K.J. & EYKHOFF, P. 1971. System identification - a survey. *Automatica* 7, 123-162.
- ATKIN, R.H. 1974. *Mathematical structure in human affairs*. London: Heinemann.
- ATKIN, R.H. 1977. *Combinatorial connectivity in social systems. An application of simplicial complex structures to the study of large organizations*. Basel: Birkhäuser.
- ATKIN, R.H. 1981. *Multidimensional man*. Harmondsworth, Middlesex: Penguin.
- BATTMANN, W. 1983. Planen muß nicht hilfreich sein: Be- und Entlastung bei Planungsprozessen. Vortrag gehalten auf der 25. Tagung experimentell arbeitender Psychologen vom 27.-31.3.1985 in Hamburg.
- BEER, S. 1970. Managing modern complexity. *Futures* 2, 114-122 und 245-257.
- BENTLER, P.M. 1980. Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology* 31, 419-456.
- BERGIUS, R. 1964. Produktives Denken (Problemlösen). In: BERGIUS, R. (Ed.), *Allgemeine Psychologie. I. Der Aufbau des Erkennens*. 2. Halbband: Lernen und Denken. (=GOTTSCHALDT, K., LERSCH, P., SANDER, F. & THOMAE, H. (Eds.), *Handbuch der Psychologie in 12 Bänden*. Band 1). Göttingen: Hogrefe, 519-563.
- BISCHOF, N. 1981. Aristoteles, Galilei, Kurt Lewin - und die Folgen. In: MICHAELIS, W. (Ed.), *Bericht über den 32. Kongreß der DGfPs in Zürich 1980*. Band 1. Göttingen: Hogrefe, 17-39.
- BOSSEL, H., GULZOW, H., HAKES, W., KANKA, M., MATHES, K., METZLER, W., SCHÄFER, H., SCHNEIDER, G., SIMON, K.-H. & WISCHNIEWSKY, W. 1982. *Dynamik von Waldökosystemen. Mathematisches Modell und Computersimulation*. Kassel: Berichte der Arbeitsgruppe Mathematisierung an der Gesamthochschule Kassel, Sonderheft 1.
- BOURNE, L.E., EKSTRAND, B.R. & DOMINOWSKI, R.L. 1971. *The psychology of thinking*. Englewood Cliffs, N.J.: Prentice-Hall.
- BOX, G.E.P. & JENKINS, G.M. 1976. *Time series analysis*. San Francisco: Holden-Day.
- BRANDTSTÄDTER, J. 1982. Apriorische Elemente in psychologischen Forschungsprogrammen. *Zeitschrift für Sozialpsychologie* 13, 267-277.
- BREDENKAMP, J. 1969. Über die Anwendung von Signifikanztests bei theorie-testenden Experimenten. *Psychologische Beiträge* 11, 275-285.
- BREDENKAMP, J. 1972. *Der Signifikanztest in der psychologischen Forschung*. Frankfurt: Akademische Verlagsgesellschaft.
- BREDENKAMP, J. 1980. *Theorie und Planung psychologischer Experimente*. Darmstadt: Steinkopff.
- BRICKENKAMP, R. 1978⁴. Test d2. Aufmerksamkeits-Belastungs-Test. Göttingen: Hogrefe.
- BROEKSTRA, G. 1976. Some comments on the application of informational measures to the processing of activity arrays. *International Journal of General Systems* 3, 43-51.

- BROWN, A.L. 1978. Knowing when, where and how to remember: A problem of metacognition. In: GLASER, R. (Ed.), *Advances in instructional psychology*. Volume 1. Hillsdale, N.J.: Erlbaum, 77-165.
- CAPRA, F. 1983². *Wendezeit. Bausteine für ein neues Weltbild*. München: Scherz.
- CASTI, J. 1979. Connectivity, complexity, and catastrophe in large-scale systems. New York: Wiley.
- CASTI, J. 1982. Topological methods for social and behavioral systems. *International Journal of General Systems* 8, 187-210.
- CHASE, W.G. & SIMON, H.A. 1973. Perception in chess. *Cognitive Psychology* 4, 55-81.
- COHEN, J. 1977². *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- COHEN, J. & COHEN, P. 1975. *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, N.J.: Erlbaum.
- CORNACCHIO, J.V. 1977a. Maximum-entropy complexity measures. *International Journal of General Systems* 3, 215-225.
- CORNACCHIO, J.V. 1977b. System complexity - a bibliography. *International Journal of General Systems* 3, 267-271.
- CZAYKA, L. 1974. *Systemwissenschaft*. Pullach: Verlag Dokumentation.
- DASTYCH, P. 1973. A measure of the complexity of a subsystem. *Cybernetica* 16, 58-72.
- DAVIS, G.A. 1966. Current status of research and theory in human problem solving. *Psychological Bulletin* 66, 36-54.
- DÖRNER, D. 1974. Die kognitive Organisation beim Problemlösen. Versuche zu einer kybernetischen Theorie der elementaren Informationsverarbeitungsprozesse beim Denken. Bern: Huber.
- DÖRNER, D. 1975. Wie Menschen eine Welt verbessern wollten... Bild der Wissenschaft 12, Heft 2, 48-53.
- DÖRNER, D. 1976. *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.
- DÖRNER, D. 1979. Kognitive Merkmale erfolgreicher und erfolgloser Problemlöser beim Umgang mit sehr komplexen Systemen. In: UECKERT, H. & RHENIUS, D. (Eds.), *Komplexe menschliche Informationsverarbeitung*. Bern: Huber, 185-195.
- DÖRNER, D. 1981. Über die Schwierigkeiten menschlichen Umgangs mit Komplexität. *Psychologische Rundschau* 32, 163-179.
- DÖRNER, D. 1982. Wie man viele Probleme zugleich löst - oder auch nicht! *Sprache & Kognition* 1, 55-66.
- DÖRNER, D. 1983a. Denken, Problemlösen und Intelligenz. In: LÜER, G. (Ed.), *Bericht über den 33. Kongreß der DGfPs in Mainz 1982*. Band 1. Göttingen: Hogrefe, 354-367.
- DÖRNER, D. 1983b. Empirische Psychologie und Alltagsrelevanz. In: JUTTEMANN, G. (Ed.), *Psychologie in der Veränderung. Perspektiven für eine gegenstandsangemessene Forschungspraxis*. Weinheim: Beltz, 13-29.
- DÖRNER, D. 1983c. The organization of action in time. Bamberg: Memorandum Nr. 19 am Lehrstuhl Psychologie II der Universität Bamberg.
- DÖRNER, D. 1984. Einige Anmerkungen zu Funkes "Diagnose der westdeutschen Problemlöseforschung in Form einiger Thesen". *Sprache & Kognition* 3, 171-172.
- DÖRNER, D. & KREUZIG, H.W. 1983. Problemlösefähigkeit und Intelligenz. *Psychologische Rundschau* 34, 185-192.
- DÖRNER, D. & REITHER, F. 1978. Über das Problemlösen in sehr komplexen Realitätsbereichen. *Zeitschrift für Experimentelle und Angewandte Psychologie* 25, 527-551.
- DÖRNER, D. & STAUDEL, T. 1979. Planen und Entscheiden in sehr komplexen Systemen. In: ECKENSBERGER, L. (Ed.), *Bericht über den 31. Kongreß der DGfPs in Mannheim 1978*. Band 1. Göttingen: Hogrefe, 440-442.

- DÖRNER, D., DREWES, U. & REITHER, F. 1975. Über das Problemlösen in sehr komplexen Realitätsbereichen. In: TACK, W.H. (Ed.), *Bericht über den 29. Kongreß der DGfPs in Salzburg 1974*. Band 1. Göttingen: Hogrefe, 339-340.
- DÖRNER, D., REITHER, F. & STAUDEL, T. 1983. Emotion und problemlösendes Denken. In: MANDL, H. & HUBER, G.L. (Eds.), *Emotion und Kognition*. München: Urban & Schwarzenberg, 61-81.
- DÖRNER, D., KREUZIG, H.W., REITHER, F. & STAUDEL, T. (Eds.) *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Huber.
- DUNCAN, O.D. 1975. *Introduction to structural equation models*. New York: Academic Press.
- DUNCKER, K. 1935. *Zur Psychologie des produktiven Denkens*. Berlin: Julius Springer.
- DUSSAUCHOY, A.E. 1982. Generalized information theory and decomposability of systems. *International Journal of General Systems* 9, 13-36.
- ELGOOD, C. 1982². *Handbook of management games*. Aldershot, Hampshire: Gower Publishing.
- ERDFELDER, E. 1984. Zur Bedeutung und Kontrolle des Beta-Fehlers bei der inferenzstatistischen Prüfung log-linearer Modelle. *Zeitschrift für Sozialpsychologie* 15, 18-32.
- ERICSSON, K.A. & SIMON, H.A. 1980. Verbal reports as data. *Psychological Review* 87, 215-251.
- EYFERTH, K., HOFFMANN-PLATO, I., MUCHOWSKI, L., OTREMBIA, H., ROSSBACH, H., SPIESS, M. & WIDOWSKI, D. 1982. *Studienprojekt Handlungsorganisation*. Berlin: Institut für Psychologie der TU Berlin (Forschungsbericht Nr. 82-4, korrigierter Nachdruck).
- FERDINAND, A.E. 1974. A theory of system complexity. *International Journal of General Systems* 1, 19-33.
- FLAMMER, A. 1981. Towards a theory of question asking. *Psychological Research* 43, 407-420.
- FLAVELL, J.H. 1979. Metacognition and cognitive monitoring. A new area of cognitive-developmental inquiry. *American Psychologist* 34, 906-911.
- FREEDLE, R. 1977. Psychology, Thomanian topologies, deviant logics, and human development. In: DATAN, N. & REESE, H.W. (Eds.), *Life-span developmental psychology. Dialectical perspectives on experimental research*. New York: Academic Press, 317-341.
- FUHRER, U. 1982. Defiziente und sub-optimale Strategien des Handelns im Umgang mit hoher Aktivationskomplexität. *Psychologische Beiträge* 24, 583-600.
- FUNKE, J. 1981. Mondlandung - ein neuer Aufgabentyp zur Erforschung komplexen Problemlösens. *Trierer Psychologische Berichte* 8, Heft 9.
- FUNKE, J. 1983a. Bereichsspezifische Effekte beim Problemlösen und der Einfluß von Vorerfahrung. *Trierer Psychologische Berichte* 10, Heft 6.
- FUNKE, J. 1983b. Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Prädiktor? *Diagnostica* 29, 283-302.
- FUNKE, J. 1984a. Alles bestätigt? Anmerkungen zum Kommentar von Wiebke Putz-Osterloh (1983). *Diagnostica* 30, 104-110.
- FUNKE, J. 1984b. Diagnose der westdeutschen Problemlöseforschung in Form einiger Thesen. *Sprache & Kognition* 3, 159-171.
- FUNKE, J. 1984c. *Komplexes Problemlösen: kritische Bestandsaufnahme und weiterführende Perspektiven*. Trier: Fachbereich I der Universität Trier (Dissertation, Photodruck).
- FUNKE, J. 1985. Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle. *Zeitschrift für Psychologie* 193, 435-457.
- FUNKE, J. & HUSSY, W. 1984. Komplexes Problemlösen: Beiträge zu seiner Erfassung sowie zur Frage der Bereichs- und Erfahrungsabhängigkeit. *Zeitschrift für Experimentelle und Angewandte Psychologie* 31, 19-38.

- FUNKE, J. & STEYER, R. 1985. Komplexes Problemlösen als Konstruktion und Anwendung von Kausalmodellen. In: ALBERT, D. (Ed.), Bericht über den 34. Kongreß der DGfPs in Wien 1984. Band 1. Göttingen: Hogrefe, 264-267.
- GAINES, B.R. 1977. System identification, approximation and complexity. *International Journal of General Systems* 3, 145-174.
- GALLOPIN, G.C. 1972. Structural properties of food webs. In: PATEN, B.C. (Ed.), 241-282.
- GEDIGA, G. 1983. Ein mathematisches Modell für den Herrscher von Summaria. Osnabrück: Arbeitsberichte Psychologische Methoden Nr. 11 aus dem Fachbereich Psychologie der Universität Osnabrück.
- GEDIGA, G., SCHÖTTKE, H. & TUCKE, M. 1982. Problemlösen in einer komplexen Situation. Ein experimenteller Beitrag im Umgang mit einem schwer durchschaubaren Problem. Osnabrück: Psychologische Forschungsberichte Nr. 27 aus dem Fachbereich 8 der Universität Osnabrück.
- GEDIGA, G., SCHÖTTKE, H. & TUCKE, M. 1983. Problemlösen in einer komplexen Situation. *Archiv für Psychologie* 135, 325-339.
- GEDIGA, G., SCHÖTTKE, H. & TUCKE-BRESSLER, M. 1984. Problemlösen und Intelligenz. Osnabrück: Psychologische Forschungsberichte Nr. 34 aus dem Fachbereich 8 der Universität Osnabrück.
- GEORGE, L. 1977. Tests for system complexity. *International Journal of General Systems* 3, 253-258.
- GLASER, R., DAMRIN, D.E. & GARDNER, R.M. 1954. The tab item: A technique for the measurement of proficiency in diagnostic problem-solving tasks. *Educational and Psychological Measurement* 14, 283-294.
- GOLDBERGER, A.S. 1964. *Econometric theory*. New York: Wiley.
- GOLDBERGER, A.S. 1971. Econometrics and psychometrics: A survey of communalities. *Psychometrika* 36, 83-107.
- GOLDBERGER, A. & DUNCAN, O. (Eds.) 1973. *Structural equation models in the social sciences*. New York: Seminar Press.
- GRAUMANN, C.F. (Ed.) 1965. *Denken*. Köln: Kiepenheuer & Witsch.
- GREENO, J.G. & RILEY, M.S. 1984. Prozesse des Verstehens und ihre Entwicklung. In: WEINERT, F.E. & KLUWE, R.H. (Eds.), *Metakognition, Motivation und Lernen*. Stuttgart: Kohlhammer, 252-274.
- GROEBEN, N. & WESTMEYER, H. 1975. *Kriterien psychologischer Forschung*. München: Juventa.
- GUSTAFSON, D.E. & SPEYER, J.L. 1975. Design of linear regulators for nonlinear stochastic systems. *Journal of Spacecrafts and Rockets* 12, 351-358.
- HACH, J.-P. & HELDT, P.H. 1984. Das Cockpit des Airbus A 310. *Spektrum der Wissenschaft*, März 1984, 38-52.
- HACKER, W. 1978². *Allgemeine Arbeits- und Ingenieurpsychologie*. Psychische Struktur und Regulation von Arbeitstätigkeiten. Bern: Huber.
- HAGER, W.R. & WESTERMANN, R.W. 1982. Die Elle - 10 Jahre danach. *Zeitschrift für Sozialpsychologie* 13, 250-252.
- HAMERLE, A. 1982. Latent-Trait-Modelle: Die methodische und multivariate statistische Analyse kategorialer Reaktionsmodelle. Weinheim: Beltz.
- HAND, H. & SIMS, H. 1975. Statistical evaluation of complex gaming performance. *Management Science* 21, 708-717.
- HAYES-ROTH, B. & HAYES-ROTH, F. 1979. A cognitive model of planning. *Cognitive Science* 3, 275-310.
- HAYES-ROTH, F., KLAHR, P. & MOSTOW, D.J. 1981. Advice taking and knowledge refinement: An iterative view of skill acquisition. In: ANDERSON, J.R. (Ed.), *Cognitive skills and their acquisition*. Hillsdale, N.J.: Erlbaum, 231-253.
- HERRMANN, T. 1982. Über begriffliche Schwächen kognitivistischer Kognitionstheorien: Begriffsinflation und Akteur-System-Kontamination. *Sprache & Kognition* 1, 3-14.

- HERRMANN, T. 1983. Nützliche Fiktionen. Anmerkungen zur Funktion kognitionspsychologischer Theoriebildungen. *Sprache & Kognition* 2, 88-99.
- HESSE, F.W. 1979. Alternative Ansätze zur Entwicklung heuristischer Strategien für den Bereich des schlußfolgernden Denkens. In: UECKERT, H. & RHENIUS, D. (Eds.), *Komplexe menschliche Informationsverarbeitung*. Bern: Huber, 153-161.
- HESSE, F.W. 1982. Effekte des semantischen Kontexts auf die Bearbeitung komplexer Probleme. *Zeitschrift für Experimentelle und Angewandte Psychologie* 29, 62-91.
- HESSE, F.W. 1983. Aufbau der Wissensrepräsentation in Problemlösesituationen und Mechanismen der Problembearbeitung - Entwicklung einer Perspektive. Göttingen: Arbeitsbericht Nr. 6 der Abteilung für Allgemeine Psychologie und Methodenlehre des Instituts für Psychologie der Georg-August-Universität Göttingen.
- HESSE, F.W. & GERRARDS, A. 1985. Einfluß der internen Zielstruktur auf den Abruf von Vorwissen beim Problemlösen. In: ALBERT, D. (Ed.), Bericht über den 34. Kongreß der DGfPs in Wien 1984. Band 1. Göttingen: Hogrefe, 267-270.
- HESSE, F.W., SPIES, K. & LÜER, G. 1983. Einfluß motivationaler Faktoren auf das Problemlöseverhalten im Umgang mit komplexen Problemen. *Zeitschrift für Experimentelle und Angewandte Psychologie* 30, 400-424.
- HEUSER, J. 1978. Zur differentiellen Wirkung von Streß auf das Problemlösen. *Zeitschrift für Experimentelle und Angewandte Psychologie* 25, 379-406.
- HIGASHI, M. & KLIR, G.J. 1982. Measures of uncertainty and information based on possibility distributions. *International Journal of General Systems* 9, 43-58.
- HORST, P. 1949. A generalized expression for the reliability of measures. *Psychometrika* 14, 21-35.
- HUBER, G.L. & MANDL, H. (Eds.) 1982. *Verbale Daten: Eine Einführung in die Grundlagen und Methoden der Erhebung und Auswertung*. Weinheim: Beltz.
- HUSSY, W. 1983. Komplexe menschliche Informationsverarbeitung: das SPIV-Modell. *Sprache & Kognition* 2, 47-62.
- HUSSY, W. 1984a. *Denkpsychologie*. Ein Lehrbuch. Band 1: Geschichte, Begriffs- und Problemlöseforschung, Intelligenz. Stuttgart: Kohlhammer.
- HUSSY, W. 1984b. Zum Begriff der Problemschwierigkeit beim komplexen Problemlösen. *Trierer Psychologische Berichte* 11, Heft 4.
- HUSSY, W. 1985. Komplexes Problemlösen - Eine Sackgasse? *Zeitschrift für Experimentelle und Angewandte Psychologie* 32, 55-74.
- HUSSY, W. 1986. *Denkpsychologie*. Ein Lehrbuch. Band 2. Stuttgart: Kohlhammer (in Vorbereitung).
- JÄGER, A.O. 1984. Intelligenzstrukturforschung: konkurrierende Modelle, neue Entwicklungen, Perspektiven. *Psychologische Rundschau* 35, 21-35.
- JÖRESKOG, K.G. 1978. Structural analysis of covariance and correlation matrices. *Psychometrika* 43, 443-477.
- JOHNSON, D.M. 1972. *Systematic introduction to the psychology of thinking*. New York: Harper & Row.
- JOHNSON-LAIRD, P.N. 1983. *Mental models. Towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.
- JULISCH, B. & KRAUSE, W. 1976. Semantischer Kontext und Problemlösungsprozesse. In: KLIX, F. (Ed.), *Psychologische Beiträge zur Analyse kognitiver Prozesse*. Berlin: Deutscher Verlag der Wissenschaften, 274-301.
- KARLINS, M., LEE, R.E. & SCHRODER, H.M. 1969. Creativity and information search in a problem solving context. *Psychonomic Science* 8, 165-166.

- KIESLER, D.J. 1971. Experimental designs in psychotherapy research. In: BERGIN, A.E. & GARFIELD, S.L. (Eds.), Handbook of psychotherapy and behavior change: an empirical analysis. New York: Wiley, 36-74.
- KLIR, G.J. 1972. The polyphonic general systems theory. In: KLIR, G.J. (Ed.), Trends in general systems theory. New York: Wiley, 1-18.
- KLIR, G.J. 1974. Taxonomy of systems. In: MURRAY, T.H. (Ed.), Interdisciplinary aspects of general systems theory. College Park, Maryland: Society for General Systems Research.
- KLIR, G.J. 1976. Identification of generative structures in empirical data. International Journal of General Systems 3, 89-104.
- KLIX, F. 1971. Information und Verhalten. Bern: Huber.
- KLIX, F. 1979. Kognitive Psychologie als integrativer Bestandteil psychologischer Grundlagenforschung: Einführung. In: UECKERT, H. & RHENIUS, D. (Eds.), Komplexe menschliche Informationsverarbeitung. Bern: Huber, 16-19.
- KLIX, F. 1980. Die Allgemeine Psychologie und die Erforschung kognitiver Prozesse. Zeitschrift für Psychologie 188, 117-139.
- KLUWE, R. 1979. Wissen und Denken. Stuttgart: Kohlhammer.
- KLUWE, R. 1983. Subjektive Ordnungsstrukturen. DFG-Projektantrag. Hamburg: Fachbereich Pädagogik der Hochschule der Bundeswehr (unveröffentlichtes Manuskript).
- KLUWE, R. & REIMANN, H. 1983. Problemlösen bei vernetzten, komplexen Problemen: Effekte des Verbalisierens auf die Problemlöseleistung. Hamburg: Bericht aus dem Fachbereich Pädagogik der Hochschule der Bundeswehr (vervielfältigtes Manuskript).
- KLUWE, R., MISIAK, C. & REIMANN, H. 1984. Lernvorgänge beim Umgang mit Systemen. Die Ausbildung subjektiver Ordnungsstrukturen durch Erfahrungen beim Umgang mit umfangreichen Systemen. Hamburg: Fachbereich Pädagogik der Hochschule der Bundeswehr (vervielfältigter DFG-Bericht).
- KLUWE, R., MISIAK, C. & SCHMIDLE, R. 1985. Wissenserwerb beim Umgang mit umfangreichen Systemen - Lernvorgänge als Ausbildung subjektiver Ordnungsstrukturen. In: ALBERT, D. (Ed.), Bericht über den 34. Kongreß der DGfPs in Wien 1984. Band 1. Göttingen: Hogrefe, 255-257.
- KRALLMANN, H. 1981. Ansätze und Problematik zur Optimierung dynamischer Systeme. In: ALBERTIN, L. & MÜLLER, N. (Eds.), Umfassende Modellierung regionaler Systeme - Probleme, Modelle, Praxisbezug. Köln: Verlag TÜV Rheinland, 123-135.
- KRAUSE, W. 1982a. Problemlösen - Stand und Perspektiven. Teil I. Zeitschrift für Psychologie 190, 17-36.
- KRAUSE, W. 1982b. Problemlösen - Stand und Perspektiven. Teil II. Zeitschrift für Psychologie 190, 141-169.
- KREUZIG, H.W. 1979a. Gütekriterien für die kognitiven Prozesse bei Entscheidungssituationen in sehr komplexen Realitätsbereichen und ihr Zusammenhang mit Persönlichkeitsmerkmalen. In: UECKERT, H. & RHENIUS, D. (Eds.), Komplexe menschliche Informationsverarbeitung. Bern: Huber, 196-209.
- KREUZIG, H.W. 1979b. Möglichkeiten zur Prognose der Güte komplexer geistiger Abläufe. In: ECKENSBERGER, L. (Ed.), Bericht über den 31. Kongreß der DGfPs in Mannheim 1978. Band 1. Göttingen: Hogrefe, 443-445.
- KREUZIG, H.W. 1981. Über den Zugang zu komplexem Problemlösen mittels prozeßorientierter kognitiver Persönlichkeitsmerkmale. Zeitschrift für Experimentelle und Angewandte Psychologie 28, 294-308.
- KREUZIG, H.W. 1983. Computer-Simulation als diagnostisches Instrument. In: LUER, G. (Ed.), Bericht über den 33. Kongreß der DGfPs in Mainz 1982. Band 1. Göttingen: Hogrefe, 147-151.
- KUHL, J. 1983. Emotion, Kognition und Motivation: II. Die funktionale Bedeutung der Emotionen für das problemlösende Denken und für das konkrete Handeln. Sprache & Kognition 2, 228-253.

- KUHLE, H.J. & BADKE, P. 1985. Die Güte von Entscheidungsvorschlägen für komplexe Situationen als Prädiktor für Problemlösefähigkeit. In: ALBERT, D. (Ed.), Bericht über den 34. Kongreß der DGfPs in Wien 1984. Band 1. Göttingen: Hogrefe, 273-275.
- LANGHEINE, R. 1980. Log-lineare Modelle zur multivariaten Analyse qualitativer Daten: eine Einführung. München: Oldenbourg.
- LANGHEINE, R. 1982. Kausalanalyse qualitativer Daten mit manifesten und latenten Variablen. Zeitschrift für Sozialpsychologie 13, 163-176.
- LANGFELDT-NAGEL, M. 1982. Untersuchungen zur Konstruktvalidität der Grundintelligenztests (CFT) von Cattell und Weiss. Diagnostica 28, 65-79.
- LAZARFELD, P.F. 1959. Latent structure analysis. In: KOCH, S. (Ed.), Psychology: A study of a science. Volume III. New York: McGraw-Hill, 476-543.
- LEWIS, M.S. 1982. Topological relationships among crisis variables. Psychotherapy: Theory, Research, Practice 19, 289-296.
- LIENERT, G.A. 1973². Verteilungsfreie Methoden in der Biostatistik. Band 1. Meisenheim/Glan: Hain.
- LÖFGREN, L. 1977. Complexity of descriptors of systems: a foundational study. International Journal of General Systems 3, 197-214.
- LUENBERGER, D.L. 1979. Introduction to dynamic systems: Theory, models and applications. New York: Wiley.
- LUER, G. 1973. Gesetzmäßige Denkabläufe beim Problemlösen. Ein empirischer Beitrag für eine psychologische Theorie der Entwicklung des Denkens. Weinheim: Beltz.
- LUER, G. & HÜBNER, R. 1985. Systemtheoretische Analyse der Steuerung eines komplexen Systems. In: ALBERT, D. (Ed.), Bericht über den 34. Kongreß der DGfPs in Wien 1984. Band 1. Göttingen: Hogrefe, 261-264.
- LUER, G. & PUTZ-OSTERLOH, W. 1978. Problem solving. The German Journal of Psychology 2, 240-258.
- LUER, G., HÜBNER, R. & LASS, U. 1984. Studies in problem solving processes. Göttingen: Arbeitsbericht aus dem Institut für Psychologie der Georg-August-Universität Göttingen.
- LUHMANN, N. 1968. Die Programmierung von Entscheidungen und das Problem der Flexibilität. In: MAYNTZ, R. (Ed.), Bürokratische Organisation. Köln: Kiepenheuer & Witsch, 324-341.
- MACKINNON, A.J. & WEARING, A.J. 1980. Complexity and decision making. Behavioral Science 25, 285-296.
- MANDL, H. & HUBER, G.L. (Eds.) 1983. Emotion und Kognition. München: Urban & Schwarzenberg.
- MCGUIRE, C.H. & BABBOTT, D. 1967. Simulation technique in the measurement of problem solving skills. Journal of Educational Measurement 4, 1-10.
- MEADOWS, D., MEADOWS, D., ZAHN, E. & MILLING, P. 1972. Die Grenzen des Wachstums. Bericht des Club of Rome zur Lage der Menschheit. Stuttgart: Deutsche Verlags-Anstalt.
- MEHRA, R.K. 1974a. Optimal inputs for linear system identification. IEEE Transactions on Automatic Control 19, 192-200.
- MEHRA, R.K. 1974b. Optimal input signals for parameter estimation in dynamic systems - Survey and new results. IEEE Transactions on Automatic Control 19, 753-768.
- METRAUX, A. (Ed.) 1981. Wissenschaftstheorie I. (=Band 1 der Kurt-Lewin-Werkausgabe, herausgegeben von C.F. GRAUMANN). Bern/Stuttgart: Huber/Klett-Cotta.
- MILLER, G.A. 1956. The magical number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review 63, 81-97.
- MORAAL, J. & KRAISS, K.-F. (Eds.) 1981. Manned systems design. Methods, equipment, and application. New York: Plenum Press.

- MORAY, N. (Ed.) Mental workload. Its theory and measurement. New York: Plenum Press.
- MORIN, E. 1974. Complexity. *International Social Science Journal* 26, 555-582.
- MUTHIG, K.P. & SCHÖNPFUG, W. 1981. Externe Speicher und rekonstruktives Verhalten. In: MICHAELIS, W. (Ed.), Bericht über den 32. Kongreß der DGfPs in Zürich 1980. Band 1. Göttingen: Hogrefe, 225-229.
- NEWELL, A., SHAW, J.C. & SIMON, H.A. 1960. Report on a general problem solving program. Proceedings of the International Conference on Information Processing. Paris: Unesco, 256-264.
- NISBETT, R.E. & WILSON, T.D. 1977. Telling more than we can know: verbal reports on mental processes. *Psychological Review* 84, 231-259.
- NORRIS, D.R. & SNYDER, C.A. 1982. External validation of simulation games. *Simulation & Games* 13, 73-85.
- OERTER, R. 1971. *Psychologie des Denkens*. Donauwörth: Auer.
- OESTERREICH, R. 1981. Handlungsregulation und Kontrolle. München: Urban & Schwarzenberg.
- OLDENBÜRGER, H.-A. 1981. Methodenheuristische Überlegungen und Untersuchungen zur "Erhebung" und Repräsentation kognitiver Strukturen. Göttingen: Mathematisch-Naturwissenschaftliche Fachbereiche der Georg-August-Universität Göttingen (Dissertation, Photodruck).
- OPP, K.-D. 1970. Methodologie der Sozialwissenschaften. Einführung in Probleme ihrer Theorienbildung. Reinbek: Rowohlt.
- OPP, K.-D. & SCHMIDT, P. 1976. Einführung in die Mehrvariablenanalyse. Grundlagen der Formulierung und Prüfung komplexer sozialwissenschaftlicher Aussagen. Reinbek: Rowohlt.
- OPWIS, K. & SPADA, H. 1983. Informationsverarbeitung beim Wissenserwerb über idealtypische biologische Populationsmodelle. Vortrag gehalten auf der 25. Tagung experimentell arbeitender Psychologen vom 27.-31.3.1983 in Hamburg.
- OPWIS, K. & SPADA, H. 1985. Erwerb und Anwendung von Wissen über ökologische Systeme. In: ALBERT, D. (Ed.), Bericht über den 34. Kongreß der DGfPs in Wien 1984. Band 1. Göttingen: Hogrefe, 258-260.
- OPWIS, K., SPADA, H. & SCHWIERSCH, M. 1985. Erwerb und Anwendung von Wissen über ein ökologisches System. Freiburg: Forschungsbericht Nr. 23 des Psychologischen Instituts der Universität Freiburg.
- OSTERLOH, M. 1983. Handlungsspielräume und Informationsverarbeitung. Bern: Huber.
- PAGE, B. 1983. Der Gültigkeitsnachweis von komplexen Simulationsmodellen. *Angewandte Informatik* 25, 149-157.
- PATTEN, B.C. (Ed.) 1971. Systems analysis and simulation in ecology. Volume 1. New York: Academic Press.
- PATTEN, B.C. (Ed.) 1972. Systems analysis and simulation in ecology. Volume 2. New York: Academic Press.
- PICHLER, F. 1975. *Mathematische Systemtheorie*. Berlin: de Gruyter.
- PICHLER, F. 1981. Dekomposition komplexer Systeme aus der Sicht der Allgemeinen Systemlehre. In: ALBERTIN, L. & MÜLLER, N. (Eds.), *Umfassende Modellierung regionaler Systeme - Probleme, Modelle, Praxisbezug*. Köln: Verlag TÜV Rheinland, 89-94.
- PITT, R.B. 1983. Development of a general problem-solving schema in adolescence and early adulthood. *Journal of Experimental Psychology: General* 112, 547-584.
- PREUSSLER, W. 1985. Über die Bedingungen der Prognose eines bivariaten ökologischen Systems. Bamberg: Memorandum No. 31. am Lehrstuhl Psychologie II der Universität Bamberg.
- PREUSSLER, W. & DÖRNER, D. 1985. Prognose der Entwicklung ökologischer Systeme. Vortrag gehalten auf der 27. Tagung experimentell arbeitender Psychologen vom 31.3.-4.4.1985 in Wuppertal.

- PUTZ-OSTERLOH, W. 1974. Über die Effektivität von Problemlösetraining. *Zeitschrift für Psychologie* 182, 253-276.
- PUTZ-OSTERLOH, W. 1981. Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg. *Zeitschrift für Psychologie* 189, 79-100.
- PUTZ-OSTERLOH, W. 1983a. Kommentare zu dem Aufsatz von J. Funke: Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Prädiktor? *Diagnostica* 29, 303-309.
- PUTZ-OSTERLOH, W. 1983b. Über Determinanten komplexer Problemlöseleistungen und Möglichkeiten zu ihrer Erfassung. *Sprache & Kognition* 2, 100-116.
- PUTZ-OSTERLOH, W. 1985a. Die Veränderbarkeit des Verhaltens im Umgang mit komplexen Problemen. In: ALBERT, D. (Ed.), Bericht über den 34. Kongreß der DGfPs in Wien 1984. Band 1. Göttingen: Hogrefe, 270-272.
- PUTZ-OSTERLOH, W. 1985b. Selbstreflexion, Testintelligenz und interindividuelle Unterschiede bei der Bewältigung komplexer Probleme. Vortrag gehalten auf der 27. Tagung experimentell arbeitender Psychologen vom 31.3.-4.4.1985 in Wuppertal.
- PUTZ-OSTERLOH, W. & LÜER, G. 1979. Wann produzieren Probanden räumliche Vorstellungen beim Lösen von Raumvorstellungsaufgaben? *Zeitschrift für Experimentelle und Angewandte Psychologie* 26, 138-156.
- PUTZ-OSTERLOH, W. & LÜER, G. 1981. Über die Vorhersagbarkeit komplexer Problemlöseleistungen durch Ergebnisse in einem Intelligenztest. *Zeitschrift für Experimentelle und Angewandte Psychologie* 28, 309-334.
- RASMUSSEN, J. & ROUSE, W.B. (Eds.) 1981. *Human detection and diagnosis of system failures*. New York: Plenum Press.
- RAVEN, J.C. 1965. *Advanced Progressive Matrices. Sets I and II. Plan and use of the scale with a report of experimental work*. London: Lewis.
- RAY, W.S. 1955. Complex tasks for use in human problem-solving research. *Psychological Bulletin* 52, 134-149.
- REIMANN, H. & KLUWE, R. 1983. Effekte des Verbalisierens auf die Problemlöseleistung beim Umgang mit komplexen Systemen. Vortrag gehalten auf der 25. Tagung experimentell arbeitender Psychologen vom 27.-31.3.1983 in Hamburg.
- REINHARDT, F. & SOEDER, H. 1982⁵. *dtv-Atlas zur Mathematik. Tafeln und Texte*. München: Deutscher Taschenbuch Verlag.
- REITHER, F. 1979. Über die kognitive Organisation bei der Bewältigung von Krisensituationen. In: UECKERT, H. & RHENIUS, D. (Eds.), *Komplexe menschliche Informationsverarbeitung*. Bern: Huber, 210-222.
- REITHER, F. 1981. About thinking and acting of experts in complex situations. *Simulation & Games* 12, 125-140.
- RICHTER, O. 1985. *Simulation des Verhaltens ökologischer Systeme. Mathematische Methoden und Modelle*. Weinheim: VCH Verlagsgesellschaft.
- RIMOLDI, H.J. 1960. A technique for the study of problem-solving. *Educational and Psychological Measurement* 20, 449-460.
- ROSEN, R. 1975. Complexity and error in social dynamics. *International Journal of General Systems* 2, 145-148.
- ROSEN, R. 1977. Complexity as a system property. *International Journal of General Systems* 3, 227-232.
- ROST, J. & VENT, U. 1985. Der Einfluß zweier Arten der Wissenspräsentation und Verhaltensrückmeldung auf die Problemlöseleistung in komplexen Situationen. Vortrag gehalten auf der 27. Tagung experimentell arbeitender Psychologen vom 31.3.-4.4.1985 in Wuppertal.
- ROTH, T. 1985. Sprachstatistisch objektivierbare Denkstilunterschiede zwischen "guten" und "schlechten" Bearbeitern komplexer Probleme. Göttingen: Arbeitsbericht aus dem Institut für Psychologie der Georg-August-Universität Göttingen.
- ROUSE, W.B. 1973. A model of the human in a cognitive prediction task. *IEEE Transactions on Systems, Man, and Cybernetics* 3, 473-477.

- ROUSE, W.B. 1977. A theory of human decision making in stochastic estimation tasks. *IEEE Transactions on Systems, Man, and Cybernetics* 7, 274-283.
- RUBEN, B.D. & LEDERMAN, L.C. 1982. Instructional simulation gaming: validity, reliability, and utility. *Simulation & Games* 13, 233-244.
- RÜPPELL, H. & RÜSCHSTROER, H. 1984. GIN & CHIPS. Ein praxisorientiertes Curriculum zur Ausbildung der produktiven Intelligenz. Bonn: Berichte aus dem Psychologischen Institut der Universität Bonn 1984, Heft 4.
- SAARI, D.G. 1977. A qualitative model for the dynamics of cognitive processes. *Journal of Mathematical Psychology* 15, 145-168.
- SAHAL, D. 1976. Elements of an emerging theory of complexity per se. *Cybernetica* 19, 5-38.
- SAMUELSON, P.A. 1939. Interaction between the multiplier analysis and the principle of acceleration. *Review of Economic Statistics* 21, 75-78.
- SAVORY, S.E. (Ed.) 1985. Künstliche Intelligenz und Expertensysteme. München: Oldenbourg.
- SCHEELE, B. & GROEBEN, N. 1984. Die Heidelberger Struktur-lege-Technik (SLT). Eine Dialog-Konsens-Methode zur Erhebung Subjektiver Theorien mittlerer Reichweite. Weinheim: Beltz.
- SCHICK, K. 1981³. Lineares Optimieren. Einführung in die mathematische Behandlung moderner Probleme in den Wirtschaftswissenschaften. Frankfurt: Diesterweg.
- SCHOTT, F., NEEB, K.-E. & WIEBERG, H.-J.W. 1984. Kleine Systeme als Problem der internen Repräsentation und der kognitiven Entwicklung. Eine Pilotstudie. Gießen: Fachbereich 06 Psychologie der Justus-Liebig-Universität Gießen.
- SCHÖTTKE, H. & GEDIGA, G. 1982. Psychologische Experimente auf Kleincomputern II. "HAMURABI der König von Summaria". Arbeitsberichte Psychologische Methoden Nr. 9 aus dem Fachbereich Psychologie der Universität Osnabrück.
- SCHULZ, P. 1979. Regulation und Fehlregulation im Verhalten. II. Streß durch Fehlregulation. *Psychologische Beiträge* 21, 597-621.
- SCHULZ, P. 1980. Regulation und Fehlregulation im Verhalten. V. Die wechselseitige Beeinflussung von mentaler und emotionaler Beanspruchung. *Psychologische Beiträge* 22, 633-656.
- SCHULZ, T., MUTHIG, K.-P. & KOEPLER, K. 1981. Theorie, Experiment und Versuchsplanung in der Psychologie. Stuttgart: Kohlhammer.
- SEIDEL, R. 1977. Objektive Beschreibung von Problemen und Beurteilung von Problemlösungsprozessen an Hand exakter Bewertungen der Problemzustände - dargestellt am Schachspiel. *Zeitschrift für Psychologie* 185, 434-454.
- SELZ, O. 1913. Über die Gesetze des geordneten Denkverlaufs. Stuttgart: Spemann.
- SELZ, O. 1922. Zur Psychologie des produktiven Denkens und des Irrtums. Bonn: F. Cohen.
- SILJAK, D.D. 1978. Large-scale dynamic systems. Stability and structure. New York: North-Holland.
- SIMON, D.P. & SIMON, H.A. 1978. Individual differences in solving physics problems. In: SIEGLER, R. (Ed.), *Children's thinking: What develops?* Hillsdale, N.J.: Erlbaum.
- SIMON, H.A. 1962. The architecture of complexity. *Proceedings of the American Philosophical Society* 106, 467-482.
- SIMON, H. & WEDEKIND, J. 1980. Das Computer-unterstützte Planspiel TANALÄND als Test- und Trainingsinstrument zum Problemlösen in komplexen Systemen. In: SIMON, H. (Ed.), *Computer-Simulation und Modellbildung im Unterricht*. München: Oldenbourg, 273-288.
- SIXTL, F. 1982². Meßmethoden der Psychologie. Theoretische Grundlagen und Probleme. Weinheim: Beltz.
- SPADA, H., MAY, R. & OPWIS, K. 1983. Wissensaufbau und Handlungsbewertung bei ökologischen Problemen. Arbeitsbericht zum DFG-Projekt. Freiburg: Psychologisches Institut der Universität Freiburg.

- SPADA, H., REIMANN, P. & HAUSLER, B. 1983. Hypothesenerarbeitung und Wissensaufbau beim Schüler. In: KÖTTER, L. & MANDL, H. (Eds.), *Kognitive Prozesse und Unterricht*. Jahrbuch für Empirische Erziehungswissenschaft 1983. Düsseldorf: Schwann, 139-167.
- SPEEDIE, S.M., TREFFINGER, D.J. & HOUTZ, J.C. 1976. Classification and evaluation of problem solving tasks. *Contemporary Educational Psychology* 1, 52-75.
- SPIES, K. & HESSE, F.W. 1983. Analyse von Motivationseffekten in komplexen Problemlösesituationen aufgrund von Daten des lauten Denkens. In: LUER, G. (Ed.), *Bericht über den 33. Kongreß der DGfPs in Mainz 1982*. Band 1. Göttingen: Hogrefe, 389-393.
- STERNBERG, R. 1982. Reasoning, problem solving, and intelligence. In: STERNBERG, R.J. (Ed.), *Handbook of human intelligence*. Cambridge: Cambridge University Press, 225-307.
- STEWART, I.N. & PEREGOY, P.L. 1983. Catastrophe theory modeling in psychology. *Psychological Bulletin* 94, 336-362.
- STEYER, R. 1982a. Modelle zur kausalen Erklärung statistischer Zusammenhänge. Frankfurt: Institut für Psychologie der Universität Frankfurt (unveröffentlichtes Manuskript).
- STEYER, R. 1982b. Structural equations, stability, and equilibrium points in multivariate autoregressive processes. Paper read at the 13th European Mathematical Psychology Group Meeting in Bielefeld, September 1982.
- STEYER, R. 1983a. Causal linear stochastic dependencies: An introduction. *Trierer Psychologische Berichte* 10, Heft 10.
- STEYER, R. 1983b. Causal linear stochastic dependencies: The formal theory. *Trierer Psychologische Berichte* 10, Heft 11.
- STEYER, R. 1983c. Modelle zur kausalen Erklärung statistischer Zusammenhänge. In: BREDEKAMP, J. & FEGER, H. (Eds.), *Strukturierung und Reduzierung von Daten*. (=Enzyklopädie der Psychologie, Themenbereich B: Methodologie und Methoden, Serie I: Forschungsmethoden der Psychologie, Band 4). Göttingen: Hogrefe, 59-153.
- STEYER, R. 1984. Conditional expectations: An introduction to the concept and its applications in empirical sciences. *Trierer Psychologische Berichte* 11, Heft 3.
- STROHSCHNEIDER, S. 1985. Eine Untersuchung zur Stabilität und Validität von Handeln in komplexen Realitätsbereichen. Vortrag gehalten auf der 27. Tagung experimentell arbeitender Psychologen vom 31.3.-4.4.1985 in Wuppertal.
- SÜLLWOLD, F. 1960. Bedingungen und Gesetzmäßigkeiten des Problemlösungsverhaltens. In: THOMAE, H. (Ed.), *Bericht über den 22. Kongreß der DGfPs in Bonn 1960*. Göttingen: Hogrefe, 96-115.
- SUSSMAN, H. & ZAHLER, R. 1978. Catastrophe theory as applied to the social and biological sciences: a critique. *Synthese* 37, 117-216.
- SYDOW, A. 1981. Hierarchisierung und Dekomposition komplexer Modelle - Anwendung auf volkswirtschaftliche Planungssysteme. In: ALBERTIN, L. & MÜLLER, N. (Eds.), *Umfassende Modellierung regionaler Systeme - Probleme, Modelle, Praxisbezug*. Köln: Verlag TÜV Rheinland, 95-107.
- SYDOW, H. 1980. Mathematische Modellierung der Strukturrepräsentation und der Strukturzeugung in Denkprozessen. *Zeitschrift für Psychologie* 188, 166-197.
- TACK, W.H. 1980. Zur Theorie psychometrischer Verfahren. Formalisierung der Erfassung von Situationsabhängigkeit und Veränderung. *Zeitschrift für Differentielle und Diagnostische Psychologie* 1, 87-106.
- TCHON, K. 1983. Towards a global analysis of systems. *International Journal of General Systems* 9, 171-175.
- TCHON, K. & WOJCIECHOWSKA, J. 1981. Towards a formal theory of description. *International Journal of General Systems* 6, 217-224.
- TENT, L. 1984. Intelligenz und Problemlösefähigkeit. Kommentar zu Dörner & Kreuzig. *Psychologische Rundschau* 35, 152-153.

- TERGAN, S.-O. 1984. Diagnose von Wissensstrukturen. Tübingen: Forschungsbericht Nr. 30 aus dem Deutschen Institut für Fernstudien an der Universität Tübingen.
- THALMAIER, A. 1979. Zur kognitiven Bewältigung der optimalen Steuerung eines dynamischen Systems. Zeitschrift für Experimentelle und Angewandte Psychologie 26, 388-421.
- THOM, R. 1975. Structural stability and morphogenesis. Reading, Mass.: Addison-Wesley.
- THOMPSON, K.R. & KEON, T.L. 1982. The creation of certain and uncertain environments. A study using a management simulation. Simulation & Games 13, 437-450.
- UECKERT, H. & RHENIUS, D. (Eds.) 1979. Komplexe menschliche Informationsverarbeitung. Beiträge zur Tagung "Kognitive Psychologie" in Hamburg 1978. Bern: Huber.
- VAN HEUSDEN, A. 1980. Human prediction of third-order autoregressive time-series. IEEE Transactions on Systems, Man, and Cybernetics 10, 38-43.
- VEMURI, V. 1978. Modeling of complex systems. An introduction. New York: Academic Press.
- VON KUTSCHERA, F. & BREITKOPF, A. 1971. Einführung in die moderne Logik. Freiburg: Alber.
- WALKER, C.C. 1971. Behavior of a class of complex systems: The effect of system size on properties of terminal cycles. Journal of Cybernetics 1, 55-67.
- WALKER, C.C. & AADRYAN, A.A. 1971. Amount of computation preceding externally detectable steady-state behavior in a class of complex systems. Bio-Medical Computing 2, 85-94.
- WALKER, C.C. & ASHBY, W.R. 1966. On temporal characteristics of behavior in certain complex systems. Kybernetik 3, 100-108.
- WALLER, R.J. 1982. Complexity and the boundaries of human policy making. International Journal of General Systems 9, 1-11.
- WEINBERG, G.M. 1972. A computer approach to general systems theory. In: KLIR, G.J. (Ed.), Trends in general systems theory. New York: Wiley, 98-141.
- WEINERT, F.E. & KLUWE, R.H. (Eds.) 1984. Metakognition, Motivation und Lernen. Stuttgart: Kohlhammer.
- WESTERMANN, R. & HAGER, W. 1982. Entscheidung über statistische und wissenschaftliche Hypothesen: Zur Differenzierung und Systematisierung der Beziehungen. Zeitschrift für Sozialpsychologie 13, 13-21.
- WINER, B.J. 1970. Statistical principles in experimental design. London: McGraw-Hill.
- WOLFE, J. 1978. The effects of game complexity on the acquisition of business policy knowledge. Decision Sciences 9, 143-155.
- ZADEH, L.A. 1965. Fuzzy sets. Information and Control 8, 338-353.
- ZADEH, L.A. 1973. Outline of a new approach to the analysis of complex systems and decision processes. IEEE Transactions on Systems, Man, and Cybernetics 3, 28-44.
- ZEEMAN, E.C. 1976. Catastrophe theory. Scientific American 234, No. 4, 65-83.
- ZEEMAN, E.C. 1977. Selected papers, 1972-1977. Reading, Mass.: Addison-Wesley.
- ZIMMERMAN, D.W. & WILLIAMS, R.H. 1977. The theory of test validity and correlated errors of measurement. Journal of Mathematical Psychology 16, 135-152.
- ZIMMERMANN, H.-J. 1976. Description and optimization of fuzzy systems. International Journal of General Systems 2, 209-215.
- ZUCKERMAN, D.W. & HORN, R.E. 1973. The guide to simulations/games for education and training. Lexington, Mass.: Informational Resources.
- ZWICKER, E. 1981. Simulation und Analyse dynamischer Systeme in Wirtschafts- und Sozialwissenschaften. Berlin: de Gruyter.

Sachregister

- Absichtsbehandlung 8, 22, 23
AR-Prozeß 85ff, 119, 121
Eigenschaften 87f
Formen 91
- bang-bang-Steuerung 16
Betriebswirtschaft 25
Betroffenheit 13
Bewußtsein 118, 121
Blickbewegung 12
- Computerprogramm 14, 49ff, 118, 123
- Differenzgleichung 30f, 113
Dynamik 30f
- Eigendynamik 16, 96, 101, 114
Einzelfallanalyse 18
Experten vs. Novizen 42, 51, 78, 109, 112
Expertensystem 68
- fuzzy systems 27, 115
- Gestaltpsychologie 3
Gütekriterium 8, 15, 48, 51, 57f, 75
- Hysteresis 40
- Informatik 25
Inzidenzmatrix 33ff
- Käferzucht 93
Katastrophen 37ff
-theorie 39ff, 52
Kausaldiagramm 31
Kausalmodell 24, 84ff, 98ff, 107ff, 110f, 117f, 119f
subjektives 109, 120f
Gedächtnisannahme 117
Kontextannahme 115f
Ökonomieannahme 113f
Repräsentationsannahme 108ff
Sequenzannahme 111f
Unvollständigkeitsannahme 114f
Zeitverzögerungsannahme 112f
- Komplexität 3, 6, 26f, 95, 115, 125
Kontroll- 28f, 95
statische 28, 95
dynamische 30, 95
- Konnektivität
siehe Vernetztheit
- lautes Denken 24
law of requisite variety 29
Lösungsgüte 8, 12ff, 16, 43, 47f, 78, 121
Indikator 47, 57, 59, 67
Reliabilität 48ff, 78f
- Validität 48ff, 78f
- Mehrfachhandeln 120
Meßfehler 44
Metakognition 24
Methodik 20f, 24f, 45, 52
Modell 24, 85, 124ff
der Untersuchungssituation 123f
dynamisches 85ff, 90
Modellbildung 125
Multiplikator-Akzeler.-Modell 30f
- Notfallreaktion 9, 24
- Ökologie 25
operatives Abbildsystem 120
Operatoren 91, 109
Anwendungsbereich 92f
Kosten 94
Reversibilität 92, 112
Wirkungsbreite 91
Wirkungssicherheit 93
Optimierung 43, 47, 51, 78
- Personenmerkmale 12ff, 82, 126f
Polytelie 98
Problemlösen 2, 3, 23, 125
einfaches 1
komplexes 1, 3, 4ff, 15, 20, 25, 45ff, 69ff, 76f, 98
und Intelligenz 8ff, 12, 44, 46, 69ff, 73, 76f, 125
Taxonomien 41ff
Diagnostik 44
Problemschwierigkeit 11, 81f, 101, 116, 126
- q-Analyse 33ff
- Reliabilität 13, 15, 48, 76
Replikation 70f
Repräsentativität 62f
- Signifikanz 45, 73, 76,
Simplex 33
Simulationssysteme 5, 7, 10f, 25, 124ff
Übersicht 11
Realitätsnähe 7, 9, 59, 123f
Simulation der 63ff, 68
DAGU 12
DORI 12
ENERGIE 14
EPIDEMIE 13
HAMURABI 13, 57, 105f, 113
LOHHAUSEN 7ff, 22, 26, 47, 78
MINISEE 14f
MONDLANDUNG 15f, 47, 78, 103ff
MORO 13
ÖKOSYSTEM 17, 96f, 100
PORAEU 17
SIM002/SIM003 17f

TAILORSHOP 9f, 45, 47, 51,
 53ff, 62f, 69, 73, 76
 TANALAND 5ff, 78
 URGOS 73
 WELT 13, 99
 Situationsmerkmale 14-19, 127
 Sprachpsychologie 24, 119
 Sprachstatistik 24
 Stabilität 13, 26, 37ff, 41
 von Güteindikatoren 15, 76f
 von Systemen 37, 63
 -messung 37
 -arten 37, 38
 Stichprobenspezifität 46
 Streß 24
 Strukturgleichungsmodell 24, 85
 Strukturvektor 35f
 System 1; 15ff, 31f, 46, 85, 116
 Definition 85
 chaotisches 52
 Uniformität 126
 Systemtheorie 1, 14, 26

 Theorie 20ff, 82f, 84, 121
 Training 109
 Transparenz/Intransparenz 6, 10,
 12, 17, 70, 72, 77, 79ff, 94,
 98
 Trendwerte 49f

 Untersuchungssituation 15, 77, 127
 Modell der 15, 123f

 Validität 7, 13, 15, 41, 48, 62f,
 79, 124
 Verbaldaten 18, 24, 110, 121
 Vernetztheit 6, 10, 17, 26, 28,
 31ff, 55ff, 95f, 99ff, 121
 Versuchsleitereffekt 45f, 80
 Versuchsplanung 45f
 Vorwissen 17, 60f, 115ff, 123
 Einfluß 61, 110, 112
 Diagnostik 116f
 Homogenität 126

 WEIV-Paradigma 14
 Wissen 6, 14, 18
 -serwerb 17, 111
 -srepräsentation 19, 99, 122
 -sdiagnostik 108, 117, 121, 126

 Zufallsschwankungen 55, 57f

Namensregister

AADRYAN, A.A. 37
 ACH, N. 2
 AEBLI, H. 3, 42
 ARIMOTO, S. 112
 ASHBY, W.R. 29, 37
 ASTRÖM, K.J. 27
 ATKIN, R.H. 33, 36, 37

 BABBOTT, D. 80
 BADKE, P. 12
 BATTMANN, W. 24
 BEER, S. 27
 BENTLER, P.M. 85
 BERGIUS, R. 3
 BISCHOF, N. 1
 BOSSEL, H. 62
 BOURNE, L.E. 42
 BOX, G.E.P. 124, 125
 BRANDTSTÄDTER, J. 21, 69
 BREDEKAMP, J. 45, 46, 71, 80
 BRICKENKAMP, R. 47
 BROEKSTRA, G. 37
 BROWN, A.L. 24

 CAPRA, F. 1
 CASTI, J. 26-28, 30, 32-34, 37, 39
 CHASE, W.G. 109
 COHEN, J. 45, 46, 111
 COHEN, P. 45
 CORNACCHIO, J.V. 28, 37
 CZAYKA, L. 63

 DAMRIN, D.E. 3, 80
 DASTYCH, P. 37
 DAVIS, G.A. 42
 DOMINOWSKI, R.L. 42
 DÖRNER, D. 3-11, 16, 17, 20-24, 26,
 27, 33, 41-44, 46, 48, 51-53,
 61, 62, 81, 82, 91, 93, 94, 109,
 110, 112, 120, 125, 126
 DREWES, U. 3, 5-7
 DUNCAN, O.D. 24, 85
 DUNCKER, K. 2, 3
 DUSSAUCHOY, A.E. 37

 EKSTRAND, B.R. 42
 ELGOOD, C. 25
 ERDFELDER, E. 25
 ERICSSON, K.A. 24
 EYFERTH, K. 11, 13, 99
 EYKHOFF, P. 27

 FERDINAND, A.E. 27
 FISHER 41
 FLAMMER, A. 46
 FLAVELL, J.H. 24
 FREEDLE, R. 41
 FUHRER, U. 46
 FUNKE, J. 10, 11, 15-17, 20, 43,
 45, 47, 53, 58, 63, 67, 69, 76-
 78, 81, 103, 104

 GAINES, B.R. 27
 GARDNER, R.M. 3, 80
 GEDIGA, G. 11, 13, 57, 105, 106,
 113
 GERRARDS, A. 11, 116
 GEORGE, L. 27
 GLASER, R. 3, 80
 GOLDBERGER, A.S. 24, 85
 GRAUMANN, C.F. 3
 GREENO, J.G. 116
 GROEBEN, N. 111, 126
 GUSTAFSON, D.E. 112

 HACH, J.-P. 46
 HACKER, W. 120
 HAGER, W.R. 46, 70, 72
 HAMERLE, A. 85
 HAND, H. 25, 79
 HAUSLER, B. 14
 HAYES-ROTH, B. 120
 HAYES-ROTH, F. 107, 120
 HELDT, P.H. 46
 HERRMANN, T. 82, 83
 HESSE, F.W. 11-13, 24, 42, 45, 99,
 116
 HEUSER, J. 24
 HIGASHI, M. 37
 HORN, R.E. 25
 HORST, P. 48, 50, 78
 HOUTZ, J.C. 42
 HUBER, G.L. 2, 23, 24
 HÜBNER, R. 11, 12, 126
 HUSSY, W. 3, 10, 11, 15, 17, 43,
 45, 47, 81, 82, 95, 126

 JÄGER, A.O. 10
 JENKINS, G.M. 124, 125
 JOHNSON, D.M. 42
 JOHNSON-LAIRD, P.N. 115, 118, 119,
 121
 JÖRESKOG, K.G. 85
 JÜLISCH, B. 1

 KARLINS, M. 80
 KEON, T.L. 25
 KIESLER, D.J. 126
 KIMURA, H. 112
 KLAHR, P. 107
 KLIR, G.J. 27, 32, 37, 44
 KLIX, F. 2, 23
 KLUWE, R.H. 11, 17, 18, 24, 46, 99
 KOEPLER, K. 111
 KRAISS, K.-F. 3
 KRALLMANN, H. 47
 KRAUSE, W. 1, 4
 KREUZIG, H.W. 4, 7, 8, 10, 20, 24,
 46, 62, 63
 KUHL, J. 2, 23
 KUHLE, H.-P. 12
 KULPE, O. 2

LANGEHEINE, R. 25
 LANGFELDT-NAGEL, M. 44
 LASS, U. 11, 12
 LAZARSELD, P.F. 85
 LEDERMAN, L.C. 25, 79
 LEWIN, K. 108
 LEWIS, M.S. 41
 LIENERT, G.A. 72
 LÖFGREN, L. 27
 LUENBERGER, D.L. 26
 LUER, G. 3, 4, 10-13, 17, 44, 45,
 53, 55, 69, 71-73, 77, 79, 81,
 126
 LUHMANN, N. 43

 MANDL, H. 2, 23, 24
 MAY, R. 11, 14
 MAYER, R. 3
 McGUIRE, C.H. 80
 MEADOWS, D. 1
 MEHRA, R.K. 112
 METRAUX, A. 109
 MILLER, G.A. 114
 MILLING, P. 1
 MISIAK, C. 11, 18
 MORAAI, J. 3
 MORAY, N. 24
 MORIN, E. 27
 MOSTOW, D.J. 107
 MUTHIG, K.-P. 12, 111

 NEEB, K.-E. 18
 NEWELL, A. 3
 NISBETT, R.E. 24, 110
 NORRIS, D.R. 25

 OERTER, R. 3
 OESTERREICH, R. 110, 120
 OLDENBÜRGER, H. 117
 OPP, K.-D. 22
 OPWIS, K. 11, 14, 15, 25, 123
 OSTERLOH, M. 43

 PAGE, B. 124
 PATTEN, B.C. 1
 PEREGOY, P.L. 39-41
 PICHLER, F. 26, 28, 37
 PITT, R.B. 112
 PREUSSLER, W. 11, 17
 PUTZ-OSTERLOH, W. 3, 4, 9-11, 13,
 17, 24, 43-45, 48, 50, 53-55,
 58, 63, 67, 69-75, 77-82, 93,
 112

 RASMUSSEN, J. 3, 80
 RAVEN, J.C. 70, 71
 RAY, W.S. 3
 REIMANN, H. 11, 17, 18, 46
 REIMANN, P. 14
 REINHARDT, F. 33, 35
 REITHER, F. 3-7, 9, 11, 12, 20, 24
 RHENIUS, D. 27
 RICHTER, O. 25
 RIMOLDI, H.J. 80

RILEY, M.S. 116
 ROSEN, R. 27
 ROST, J. 11, 14
 ROTH, T. 11, 12, 24
 ROUSE, W.B. 3, 80, 112, 113, 117
 RUBEN, B.D. 25, 79
 RUPPELL, H. 109
 RUSCHSTROER, H. 109

 SAARI, D.G. 41
 SAHAL, D. 27
 SAMUELSON, P.A. 30
 SAVORY, S.E. 68
 SCHEELE, B. 126
 SCHICK, K. 61
 SCHMIDLE, R. 11, 18
 SCHOTT, F. 18
 SCHÖNPFUG, W. 12
 SCHÖTTKE, H. 11, 13, 57, 105, 106,
 113
 SCHULZ, P. 24
 SCHULZ, T. 111
 SCHWIERSCH, M. 11, 14, 15, 25, 123
 SEIDEL, R. 1
 SELZ, O. 2, 3
 SHAW, J.C. 3
 SILJAK, D.D. 26
 SIMON, D.P. 109
 SIMON, H.A. 3, 24, 28, 109
 SIMON, H. 11
 SIMS, H. 25, 79
 SIXTL, F. 50, 78
 SNYDER, C.A. 25
 SOEDER, H. 33, 35
 SPADA, H. 11, 14, 15, 25, 123
 SPEEDIE, S.M. 42
 SPEYER, J.L. 112
 SPIES, K. 11, 13, 45
 STAUDEL, T. 4, 7, 9, 20
 STERNBERG, R.J. 79
 STEWART, I.N. 39, 40, 41
 STEYER, R. 16, 24, 84-89, 91, 93
 STREITZ, N. 53
 STROHSCHNEIDER, S. 11, 13, 52
 SULLWOLD, F. 3
 SUSSMAN, H. 41
 SYDOW, A. 28, 37
 SYDOW, H. 1
 SZEKELY 3

 TACK, W.H. 44
 TCHON, K. 27
 TENT, L. 10
 TERGAN, S.O. 126
 TINBERGEN 31
 THALMAIER, A. 11, 15, 16, 47, 78
 THOM, R. 39
 THOMPSON, K.R. 25
 TREFFINGER, D.J. 42
 TUCKE, M. 11, 13, 57, 105, 106, 113
 TUCKE-BRESSLER, M. 11, 13, 57

 UECKERT, H. 27

VAN HEUSDEN, A. 112, 113
 VEMURI, V. 26, 44
 VENT, U. 11, 14

 WALKER, C.C. 37
 WALLER, R.J. 27
 WEDEKIND, J. 11
 WEINBERG, G.M. 39
 WEINERT, F.E. 24
 WERTHEIMER, M. 2, 3
 WESTERMANN, R.W. 46, 70, 72
 WESTMEYER, H. 111
 WIEBERG, H.J.-W. 18
 WILLIAMS, R.H. 44, 79
 WILSON, T.D. 24, 110
 WINER, B.J. 71
 WOJCIECHOWSKA, J. 27
 WOLFE, J. 25

 ZADEH, L.A. 115
 ZAHLER, R. 41
 ZAHN, E. 1
 ZREMAN, E.C. 39
 ZIMMERMAN, D.W. 44, 79
 ZIMMERMANN, H.-J. 27
 ZUCKERMAN, D.W. 25
 ZWICKER, E. 26, 30-32, 37-39, 85,
 89 128, 131