

## Diagnose der westdeutschen Problemlöseforschung in Form einiger Thesen<sup>1</sup>

Mit einigen Anmerkungen von D. Dörner

Joachim Funke

Fachbereich I - Psychologie - der Universität Trier

*Zusammenfassung.* Der gegenwärtige bundesrepublikanische Forschungsstand im Bereich komplexen Problemlösens wird einer kritischen Diagnose unterzogen. Acht Thesen machen auf Probleme aufmerksam, die sich gegenwärtig stellen, und geben Impulse für weitere Arbeiten. Die Thesen behaupten im einzelnen (1) eine Theoriearmut gegenwärtiger Problemlöseforschung, (2) den geringen Einbezug von Erträgen einschlägiger psychologischer Teilfächer, (3) eine mangelnde fächerübergreifende Kooperation, (4) das noch unvollständige Ausschöpfen des systemtheoretischen Ansatzes, (5) das Fehlen einer Taxonomie für komplexe Problemstellungen, (6) die ungenügende Berücksichtigung des Meßfehlers bei Maßen der Problemlösefähigkeit, (7) eine Tendenz, bevorzugt «signifikante» Ergebnisse zu berichten, sowie (8) ein suboptimales versuchsplanerisches und auswertungstechnisches Vorgehen. Von der Diskussion dieser Aspekte wird eine konstruktive Wirkung erwartet.

*Diagnosis of problem solving research in the FRG in form of some theses.*

*Summary.* The present state of West-German research on problem solving in complex situations is analyzed critically. Eight theses refer to problems which are instantly seen and which also possibly give hints for further research. These theses attend to (1) a lack of theory in present problem solving research, (2) the little use of findings from other psychological disciplines, (3) a deficient cooperation with other relevant branches of science, (4) the imperfect realisation of the system-analytic approach, (5) the lack of a taxonomy for complex problem situations, (6) the disregard of measurement error concerning indicators of problem solving quality, (7) a tendency to favor «significant» results and to suppress insignificant ones, and (8) a suboptimal procedure with regard to experimental design and statistical analyses. The theses are intended to initiate a constructive discussion.

### Vorbemerkung

Komplexes Problemlösen ist seit Mitte der siebziger Jahre zu einem bevorzugten Forschungsthema verschiedener Arbeitsgruppen in der Bundesrepublik Deutschland geworden, wohl nicht zuletzt angeregt durch politische und gesellschaftliche Entwicklungen, in denen die Fähigkeit von Menschen bei der Lösung globaler Probleme (Umweltschäden, Energieprobleme, Ressourcenknappheit, Armut, Hunger, Kriegsgefährdung) offenkundig gefordert war und ist. Die vielfältigen Forschungstätigkeiten, in denen der menschliche Umgang mit komplizierten Systemen näher analysiert wurde, bedürfen jedoch einiger Anregungen, um dieses wichtige und herausfordernde Thema fruchtbar weiterzubehandeln. Im folgenden möchte ich daher einige Thesen formulieren,

die sich kritisch mit der – in der neueren Psychologie bisher noch kurzen – Forschungs«tradition» dieses Teilgebiets beschäftigen. Dabei sollen keineswegs die Verdienste in Zweifel gezogen werden, die sich schon allein durch die Eröffnung dieses Sujets für die psychologische Forschung ergeben haben. Die thesenartig formulierten Bemerkungen dienen hauptsächlich dazu, weitere Impulse zu geben.

### These 1

*Forschung zum komplexen Problemlösen wird momentan theoriearm betrieben. Die Phase des explorierenden Vorgehens muß abgelöst werden durch eine Phase der Modellbildung und -testung.*

Die «Lohhausen»-Studie von Dörner, Kreuzig, Reither & Stäudel (1983), die als hervorstechende Arbeit dieses Forschungszweigs herangezogen werden kann, verdeutlicht unsere

<sup>1</sup> Für kritische Anmerkungen und Diskussionen über diese Arbeit danke ich Prof. Jürgen Bredenkamp, Prof. Dietrich Dörner, Dipl.-Psych. Edgar Erdfelder, Prof. Walter Hussy sowie einem anonymen Gutachter.

These in eindrucksvoller Weise: an den insgesamt 48 Probanden, die das computersimulierte System der fiktiven Stadt «Lohhausen» als Bürgermeister lenken durften, fielen jeweils zirka 100 000 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 bieten eine «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 «Prüfbasis») der Stichprobe;
- (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 voneinander sind. Diese Einschränkung bezieht sich auf die klassische Einteilung von apriorischen und aposteriorischen Erkenntnissen nach Kant sowie dessen Unterscheidung zwischen analytischen und synthetischen Urteilen. Brandstädter (1982, p. 268) formuliert

den 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. Brandstädter 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., diese Aussagen kann man lediglich als «eine Art theoretischer Vorform» (Dörner) akzeptieren, aus der eine Phänomenologie der «guten» bzw. «schlechten» Problemlöser resultiert. Die anspruchsvolle 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 entschieden werden, ob die Theorie zu widersprüchlichen

Hypothesen führt oder ob eine bestimmte Untersuchung mit der Theorie 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 der «Löser» komplexer Probleme).

Nimmt man einmal das präziseste «Theorie»segment heraus, nämlich die «Theorie der Absichtsbehandlung» (Dörner, 1982), die bereits eine wesentliche Überarbeitung der Kapitel 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 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:

$$(1) \quad a_i = w_i \times d_i \times e_i \times l_i + ag_i.$$

Eine Konsequenz von (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 (vgl. Dörner, 1983b, der die Unangemessenheit des Modells bestätigt); auch das Projektionsmodell (Dörner, 1982) findet – so die Mitteilung – inzwischen keine Verwendung mehr.

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.

## These 2

*Der Einbezug 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 Integrations-*

*kraft des «information processing approach», auf die Klix (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. Kuhl, 1983; Mandl & Huber, 1983) oder «Gedächtnis» (vgl. Dörner, 1983a), auf die sich das Interesse der Problemlöseforschung bereits gerichtet hat, zeigen exemplarisch den Vorteil solcher Bereichsüberschreitungen, deren Kehrseite in der Unklarheit darüber besteht, wie ein derartig weitreichender Anspruch einer «Totalpsychologie» (Dörner) einzulösen ist.

– *Metakognition.* Neben dem Einfluß von «Selbstreflexion» auf das Problemlöseverhalten (vgl. Hesse, 1979, 1982; Putz-Osterloh, 1981, 1983; Reither, 1979) wurde auch in der «Lohhausen»-Studie eine ganze Reihe der von Metakognitionsforschern für wichtig erachteten Eigenschaften wie «predicting», «checking», «monitoring», «reality testing», «coordination and control» (nach Brown, 1978; vgl. auch Kluwe & Friedrichsen, 1984) 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; Weibert & 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. Auch dazu liegen interessante Befunde vor.

– *Sprachpsychologie.* Angesichts der intensiven Analyse von Verbalprotokollen (vgl. Huber & Mandl, 1982) und der umfangreichen verbalen Interaktion zwischen Versuchsleiter und Proband verwundert das fast völlige Fehlen von Überlegungen zu diesem Punkt. Während die Intelligenzforschung aus eben diesen Erwägungen ihr Testmaterial weitgehend sprachfrei gestaltet hat, führt die Problemlöseforschung nunmehr extensiv sprachliche Äußerungen in den Forschungsrahmen ein, ohne die daraus resultierenden Überlagerungseffekte kritisch zu reflektieren.

- *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 Variablenetzen 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. Selbst in nicht-experimenteller Forschung kann unter bestimmten Prämissen (vgl. Steyer, 1983) ein kausales Modell 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, 1982; zum Problem der Modellprüfung: Erdfelder, 1984), finden in den einschlägigen Arbeiten keinen Niederschlag; dort findet man selten mehr als bivariate Korrelationen oder einfache Varianzanalysen, obwohl die Zusammenhangshypothesen multiple Ursachen und multiple Wirkungen postulieren.

### These 3

*Trotz gegenteiliger Erwartungen findet keine fachübergreifende Kooperation statt, die vom Gegenstand her jedoch gefordert wäre. Gedacht ist dabei an 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- und 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.

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, 1982), viele Untersuchungen mit derartigen Simulationsprogrammen sind in den einschlägigen Zeitschriften wie «Management Science» (z.B. Hand & Sims, 1975), «Simulation and Games» (z.B. Norris & Snyder, 1982; Ruben & Lederman, 1982; Thompson & Keon, 1982) usw. publiziert; neben (seltenen) Beiträgen über psychologische Aspekte des Handelns in Simulationsspielen finden sich dort zahlreiche Diskussionen über die Bestimmbarkeit der Lösungsqualität oder die Trainierbarkeit bestimmter Eingriffsstrategien. Gerade die Frage nach Kriterien der Lösungsgüte hat sich ja als durchaus problematisch herausgestellt (vgl. Funke, 1983).

### These 4

*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 erlauben wir uns 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 «viele 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. einführig Casti, 1979; Luenberger, 1979; Pichler, 1975; Siljak, 1978; Vemuri, 1978; Zwicker, 1981), auch wenn sich deren 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 (In-)Stabilität (im Englischen spricht man von den drei C's: «complexity», «connectivity» und «catastrophe»).

(1) *Das Konzept der Komplexität.* «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.). Aus der vielfältigen Literatur zur Komplexität (vgl. die Bibliographie von Cornacchio, 1977) sollen hier einige Aspekte des Konzepts näher erläutert werden. – Neben statischer Komplexität, die sich mit der Komplexität der systembildenden Subsysteme beschäftigt, versteht man unter dynamischer Komplexität die «computational length», die die Subsystem-Verbindungen benötigen, um den Prozeß zu realisieren. Kontrollkomplexität schließlich befaßt sich mit den erforderlichen Maßnahmen und Berechnungen zum Betreiben eines Systems in gewünschter Weise. Wichtig ist die Feststellung, daß Dimensionalität und Komplexität eines Systems nicht gleichgesetzt werden dürfen: ein hochdimensionales System kann simpel sein, wenn z.B. nur die Diagonale der Matrix aus Eingangs- und Ausgangsvariablen besetzt ist. Kennzeichen statischer Komplexität finden sich in der hierarchischen Struktur von Subsystemen, der Art ihrer Verbindungsmuster, der Verschiedenheit der Komponenten sowie der Stärke von Interaktionen.

Zur Erfassung der Kontrollkomplexität kann man das Verhältnis zwischen Störungs- und Kontrollvariation bestimmen. Nehmen wir an, es gebe einen Satz an Kontrollvariablen  $K = \{\alpha, \beta, \gamma\}$ , die der System«kontrolleur» beeinflussen kann; weiterhin gebe es einen Satz an Störvariablen  $S = \{1, 2, 3\}$ , und das System selbst möge drei Zustände  $Z = \{a, b, c\}$  annehmen. Tabelle 1 veranschaulicht zwei mögliche Kombinationen der Zuordnung von Zuständen zu diesen Kontroll- und Störvariablen.

Der linke und rechte Teil von Tabelle 1 zeigt zwei verschiedene Systeme  $\Sigma_1$  und  $\Sigma_2$  mit unterschiedlich hoher Kontrollkomplexität. Während in  $\Sigma_1$  der Kontrolleur durch geeignete Wahl einer Kontrollvariable  $k \in K$  bei Eintritt eines beliebigen Störungstyps  $s \in S$  jeden gewünschten Zustand  $z \in Z$  herbeiführen kann (= maximale Kontrolle), bietet  $\Sigma_2$  nur noch beschränkte Manipulations- bzw. Steuerungsmöglichkeiten: lediglich bei Störungstyp 3 kann der Kontrolleur zwischen Zuständen b

Tabelle 1. Veranschaulichung 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  $\Sigma_1$  (links) und  $\Sigma_2$  (rechts).

Störvariable	Kontrollvariable			Kontrollvariable		
	$\alpha$	$\beta$	$\gamma$	$\alpha$	$\beta$	$\gamma$
1	b	a	c	a	a	a
2	a	c	b	a	a	a
3	c	b	a	b	b	c

und c wählen, in den beiden restlichen Störfällen bleiben alle Kontrollvariablen wirkungslos (= geringe Kontrolle). Aus der Kybernetik weiß man, daß die totale Variation (tV) im Verhalten eines Systems mindestens so groß ist wie das Verhältnis von Störungsvariation (SV) zu Kontrollvariation (KV), also («law of requisite variety»; vgl. Ashby, 1958):

$$(2) \quad tV \geq \frac{SV}{KV}$$

Dies besagt, daß tV mit zunehmendem Ausmaß an KV bei gleicher SV abnehmen kann bzw. umgekehrt, daß tV bei sinkender KV anwachsen kann (vgl. Casti, 1979, p. 101). Die Kenntnis von KV und SV dürfte bei der Beurteilung von Leistungen in computersimulierten Systemen daher sehr wichtig sein, da im Falle hoher SV und geringer KV schon einfache Systeme kaum mehr beherrschbar sind, menschliches «Versagen» somit keine psychologisch relevante Ursache mehr besitzt, sondern systembedingt auftritt.

(2) *Das Konzept der Vernetztheit.* Vernetztheit ist oftmals implizit definierendes Merkmal vom Begriff «System»: Klir (1972a, p. 1) nennt z.B. ein System «... an arrangement of certain components so arranged as to form a whole». Man kann es auch umgekehrt 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  $\Sigma$  are related to each other». Eine Veranschaulichung der Konnektivität erlaubt beispielsweise die graphentheore-

tische Darstellung oder die – äquivalente – «adjacency matrix» (siehe Abb. 1a und 1b).

Bei den in Abbildung 1 gewählten 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 (vgl. Atkin, 1974, 1977). Die in Abbildung 1 gezeigten Darstellungsmöglichkeiten geben übrigens auch Dörner et al. (1983, p. 42 und 46) für ein fiktives Beispiel an, bleiben aber auf diesem Niveau stehen. Dabei bestehen z.B. Möglichkeiten, einen derartigen Komplex in sogenannte «Simplexe» aufzulösen und deren Dichte untereinander zu bestimmen (Ermittlung der Homotopie von Wegen; Reinhardt und Söder, 1982, 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. Mit dem Stichwort «Exzentrizität» (vgl. Casti, 1979) bezeichnet man schließlich die relative Bedeutsamkeit eines Simplex für den ganzen Komplex sowie seine relative Bedeutsamkeit als Agent der Verknüpfung. Faßt man einen Komplex als «Schweizer Käse» auf, so liefert die von Casti (1979, 1982) beschriebene Q-Analyse des Komplexes Aussagen über die Verbindung der Substanz; über seine Löcher dagegen gibt die topologische Homologietheorie (s.o.) Auskunft.

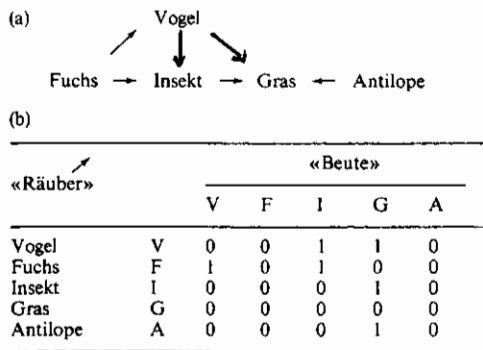


Abbildung 1. Darstellungsformen der Systemvernetzung am Beispiel einer Nahrungskette: (a) Graph, (b) «adjacency matrix» (nach Casti, 1979).

(3) *Das Konzept der Stabilität.* 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.

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, Einheitssprung) 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 (vgl. Zwicker, 1981, p. 75 ff.).

- (a) Asymptotische versus nicht-asymptotische (= Liapunov-) Stabilität: während bei ersterer das System nach Störungseinwirkung mit fortlaufender Zeit gegen den ungestörten Zustand konvergiert, ist bei letzterem nur eine «Annäherung» an den ursprünglichen Zustand erforderlich.
- (b) 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.
- (c) 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 erst abwechselnd und mit abnehmender Stärke in beide Richtungen ausschlägt.
- (d) Lokale versus globale Stabilität: man bezeichnet ein System dann als global stabil, wenn es auch nach star-

ken 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 Einsturzgefahr).

Abbildung 2 verdeutlicht einige Formen der Stabilität, die hier erwähnt wurden.

Wie Weinberg (1972, p. 124) ausführt, darf man auch keinesfalls den Fehler begehen, die Stabilität eines Systems als etwas Gutes zu deuten. Er führt als Beispiel den Schwelbrand 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

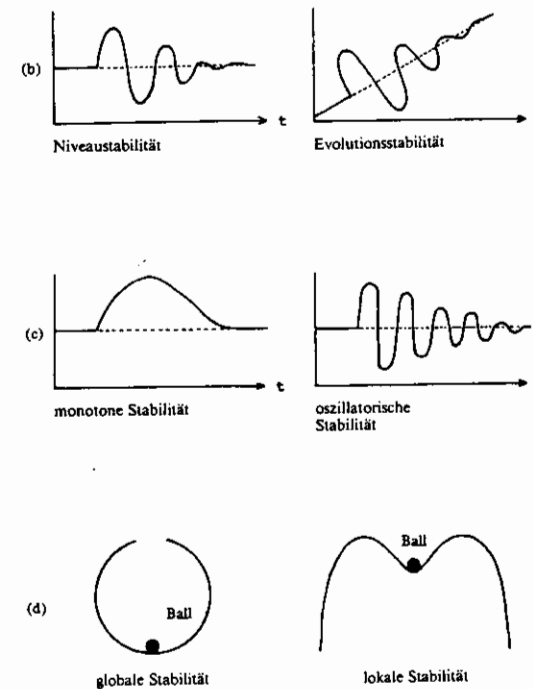


Abbildung 2. Veranschaulichung einiger Formen der Systemstabilität. Die Buchstaben (b), (c) und (d) entsprechen den Abschnitten im Text.

(«Sollwert») und der zweitens mit den Systemen in Widerspruch gerät, die nicht auf die Einhaltung bestimmter Sollwerte ausgerichtet sind (z.B. 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. Ein Spezialgebiet der Mathematik, die Topologie, liefert Mittel zur Beschreibung katastrophaler Verläufe. Die topologische Katastrophentheorie, entwickelt von Thom (1972) und eher allgemeinverständlich vorgetragen von Zeeman (1976), versucht eine oder mehrere Verhaltensvariablen durch eine oder mehrere Kontrollvariablen zu beschreiben, wobei unter bestimmten Bedingungen kleine Veränderungen der Kontrollvariable große Veränderungen der Verhaltensvariable nach sich ziehen können («catastrophe jumps»). Weitere Eigenschaften derartiger Modelle sind: Hystereseverhalten, Bimodalität, Divergenz und Unzugänglichkeit, Eigenschaften, die an anderer Stelle ausführlicher dargestellt werden (vgl. etwa Casti, 1982).

Die hier ausführlicher geschilderten Beschreibungsmöglichkeiten komplexer Systeme sind für die Problemlöseforschung in mehrerer Hinsicht wichtig: einerseits erlauben sie eine präzisere Charakterisierung des «Reizmaterials», andererseits ein vertieftes Verständnis menschlichen Umgangs mit Systemen, die bestimmte (und bestimmbare) Eigenschaften aufweisen. Erst aus dieser Kenntnis heraus sind präzise Aussagen über die Validität solcher Studien wie «Lohhausen» ableitbar.

**These 5**

Es fehlt eine differenzierte Taxonomie von Problem- bzw. System-Typen, um die an unterschiedlichen Problemen (Systemen) gewonnenen Ergebnisse systematisch vergleichen zu können. Prognosen, die aus bestimmten Modellvorstellungen resultierten, könnten 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; impliziter Bestandteil dieser Taxonomie sind aber auch Merkmale der Situation und Merkmale der Operatoren, die für die spezifischen Anforderungen eines 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.

Eine differenziertere Taxonomie stammt von Speedie, Treffinger & Houtz (1976): drei grobe Ebenen – Merkmale der Aufgabenumwelt (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 bisher um die differenzierteste zu handeln, wengleich 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) etwa

fragt, «ob Aufgaben- beziehungsweise Problemklassifikationen sowie Kriterien für die Lösung eines offenen Problems überhaupt intersubjektive Gültigkeit haben können» (p. 265). 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. Mit dem zweiten Aspekt, der Richtung des Suchprozesses, geht sie auf die Frage ein, welche Beschränkungen («constraints») Probanden als vorgegeben ansehen und welche nicht. 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.

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, «objektiveren» Kriterien zu verlangen (Funke, 1983), könnte man nunmehr dazu übergehen, subjektive Kriterien anzuwenden, die dann allerdings mit den indi-

viduellen Zielsetzungen jedes Probanden genau abgestimmt sein müssen. Jeder Proband sollte somit an seinen eigenen Kriterien gemessen werden, eine Forderung nach Einzelfallanalysen also, bei deren Umsetzung man nur noch auf globaler Ebene Aussagen über Gruppen von Individuen machen könnte.

### These 6

*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 den Arbeiten von Putz-Osterloh (1981; vgl. auch Putz-Osterloh & Lüer, 1981), Dörner & Kreuzig (1983) 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 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. Die Arbeiten der Intelligenzstrukturforscher (vgl. Jäger, 1983) reflektieren diese Fragen schon länger.

### These 7

*Wie in allen Teildisziplinen der Psychologie gilt auch für die Problemlöseforschung das Argument, nicht nur signifikante Ergebnisse zu berichten. Gerade im Rahmen eines Falsifikationsmodells über wissenschaftlichen Erkenntnisgewinn kommt dem Nichteintreffen postulierter Effekte größere Bedeutung zu als der Bestätigung von Hypothesen.*

Dieses Argument bedarf eigentlich keiner umfangreichen Begründung. Daß die im Publikationswesen häufig anzutreffende Tendenz, möglichst nur «signifikante» Befunde berichten zu wollen, unglücklich ist, braucht ebenso wenig ausgeführt werden wie die Problematik des Signifikanztests in bestimmten Kontexten (vgl. Bredenkamp, 1969, 1972).

### These 8

*Will man im Bereich komplexen Problemlösens experimentell arbeiten, sollten versuchsplanerische und auswertungstechnische Überlegungen verstärkt Gewicht erhalten. Auf postexperimentelle Aufteilungsprozeduren («gute» versus «schlechte» Problemlöser) kann verzichtet werden.*

Will man theoretisch experimentieren, spielt die Kontrolle des Beta-Fehlers eine ebenso wichtige Rolle wie die des Alpha-Fehlers; die apriori-Bestimmung des Stichprobenumfangs auf der Basis einer postulierten Effektstärke ist gut beschrieben (z.B. Cohen, 1977), wird aber bisher in der Problemlöseforschung genauso selten angetroffen wie eine – «post festum» immerhin noch mögliche – aposteriori-Poweranalyse (vgl. auch Hager & Westermann, 1982, zur Beurteilung der Qualität veröffentlichter empirischer Beiträge in der «Zeitschrift für Sozialpsychologie»). Wenn schon N und  $\alpha$  vom Experimentator (willkürlich) festgelegt werden, sollte wenigstens auch  $\beta$  (und damit die Power des Tests) oder die in der Population erwartete Mindesteffektstärke spezifiziert werden; liegen nämlich drei der vier Parameter fest, ist der verbleibende durch die vorangegangenen Entscheidungen determiniert. Idea-

erweise errechnet man aus vorher festgelegten Werten für  $\alpha$ ,  $\beta$  und Mindesteffekt die Stichprobengröße.

Vorliegende (quasi-)experimentelle Studien zum komplexen Problemlösen (Funke, 1983; Funke & Hussy, 1984; Hesse, 1982; Hesse, Spies & Lüer, 1983; Putz-Osterloh, 1981, 1983) gehen kaum über ein klassisches 2x2-Design hinaus, obwohl die zu testenden Hypothesen oftmals andere Versuchspläne verlangen. Es gibt sogar Studien mit einem einzigen zweistufigen Faktor (z. B. «Betroffenheit» bei Hesse, Spies & Lüer, 1983) oder einem einzigen dreistufigen Faktor (z. B. «Gruppenzugehörigkeit» bei Putz-Osterloh, 1983). Ob diese Designs der Komplexität des Untersuchungsgegenstands und der theoretisch postulierten vielfältigen Einflußgrößen gerecht werden, scheint fraglich. Oftmals ist einer der Faktoren (z. B. «Testintelligenz») quantitativ und wird künstlich dichotomisiert, 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. These 4) begründetes Maß der Lösungsgüte definieren zu können, erscheint es in derartigen Fällen gewinnbringender, die metrische Information regressionsanalytisch auszuschöpfen (vgl. Kapitel 8.2.2 in Cohen & Cohen, 1975: «When not to use the 2x2-design»), ein Argument übrigens, das auch für eine eigene Arbeit (Funke, 1983) 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, 1983, p. 112 links oben; siehe auch These 7). 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 aufseiten der abhängigen Variablen interessieren nicht isolierte Auswirkungen, sondern Auswirkungsgefüge.

Ein weiteres Argument bezüglich Versuchsplanung betrifft die sozialpsychologische Seite

des Experimentierens (vgl. Bredenkamp, 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 Vermittlerrolle mehr zukommt. Die Interaktion des Probanden mit komplexen Problemen sollte soweit wie möglich automatisiert ablaufen, da schon die bloße Anwesenheit von Versuchsleitern mit unterschiedlichem sozialen Status bedeutsame Effekte auf die Bearbeitung komplexer Probleme zeigt (siehe Kluwe & Reimann, 1983; Reimann & Kluwe, 1983). Außerdem entspricht es eher der Realität vieler, insbesondere technischer Problemsituationen, daß das Individuum Auskunft über den Systemzustand durch vielfältige «Instrumente» (darunter z. B. Anzeige- und Kontrolltafeln, Computer, Warnsignale) erfährt als durch einen «Informanten» (vgl. Rasmussen & Rouse, 1981).

Zuletzt sei auf das Problem der Stichprobenspezifität hingewiesen. Gerade angesichts der oft geringen Probandenzahlen muß sichergestellt sein, auf welche Population man seine Aussagen beziehen möchte. Daß nicht nur Studenten untersucht werden sollten (und vielleicht daraus auch noch eine Subpopulation wie z. B. Skiläufer; vgl. Fuhrer, 1982), muß selbstverständlich sein. Andernfalls dürften geringe Korrelationen zwischen «Problemlösefähigkeit und Intelligenz» (Dörner & Kreuzig, 1983) selbst dann nicht mehr überraschen, wenn reliable Gütemaße verwendet würden.

#### Nachbemerkung

Die acht vorstehenden Thesen sind in vielerlei Hinsicht unvollständig, teilweise vielleicht auch einseitig; einige Thesen stellen auch Spezifikationen anderer Thesen dar, sie sind also nicht unabhängig voneinander. Ihre Funktion sehe ich darin, aufmerksam zu machen auf einige kritische Aspekte des gegenwärtigen Forschungsstands im Bereich komplexen Problemlösens. Von einem komplexen Gegenstand darf nicht erwartet werden, daß sich seine Grundstrukturen in kurzer Zeit und mit einfachen Mitteln erhellen lassen. Umso notwendiger scheint es mir, offene Fragen anzusprechen

und Schwächen bisheriger Arbeit in konstruktiver Absicht zu diskutieren. Möglichst detaillierte Schilderungen der durchgeführten Studien und ein Höchstmaß an Skepsis und Vorsicht gegenüber den eigenen «Befunden» sind dazu erforderlich. Replikationsstudien sollten nicht nur von der eigenen Arbeitsgruppe, sondern auch von anderen Forschern durchgeführt werden (vgl. die berechtigten Forderungen von Sommer & Sommer, 1983, bezüglich der sog. Milwaukee-Studie!). Unter solchen Prämissen halte ich weitere Analysen von komplexem Problemlösen für sinnvoll. Die bisher vorgelegten Daten können dabei – sofern sie sich als stabil herausstellen – als Heuristik dienen.

#### Literatur

- Aebli, H.: *Denken: das Ordnen des Tuns. Band II: Denkprozesse*. Stuttgart: Klett-Cotta 1981.
- Ashby, W.R.: Requisite variety and its implications for the control of complex systems. *Cybernetica*, 1958, 1, 83-99.
- Atkin, R.H.: *Mathematical Structure in Human Affairs*. London: Heinemann 1974.
- Atkin, R.H.: *Combinatorial Connectivity in Social Systems. An Application of Simplicial Complex Structures to the Study of Large Organizations*. Basel: Birkhäuser 1977.
- Battmann, W.: Planen muß nicht hilfreich sein: Be- und Entlastung bei Planungsprozessen. Vortrag gehalten auf der 25. Tagung experimentell arbeitender Psychologen vom 27.-31.3.1983 in Hamburg.
- Bourne, L.E., Ekstrand, B.R. & Dominowski, R.L.: *The Psychology of Thinking*. Englewood Cliffs, N.J.: Prentice-Hall 1971.
- Brandtstädter, J.: Apriorische Elemente in psychologischen Forschungsprogrammen. *Zeitschrift für Sozialpsychologie*, 1982, 13, 267-277.
- Bredenkamp, J.: Über die Anwendung von Signifikanztests bei theorie-testenden Experimenten. *Psychologische Beiträge*, 1969, 11, 275-285.
- Bredenkamp, J.: *Der Signifikanztest in der psychologischen Forschung*. Frankfurt: Akademische Verlagsgesellschaft 1972.
- Bredenkamp, J.: *Theorie und Planung psychologischer Experimente*. Darmstadt: Steinkopff 1980.
- Brown, A.L.: Knowing when, where and how to remember: A problem of metacognition. In Glaser, R. (Ed.), *Advances in Instructional Psychology*. Vol. 1. Hillsdale, N.J.: Erlbaum 1978, 77-165.
- Casti, J.: *Connectivity, Complexity, and Catastrophe in Large-Scale Systems*. New York: Wiley 1979.
- Casti, J.: Topological methods for social and behavioral systems. *International Journal of General Systems*, 1982, 8, 187-210.
- Cohen, J.: *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press 1977<sup>2</sup>.
- Cohen, J. & Cohen, P.: *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Hillsdale, N.J.: Lawrence Erlbaum 1975.
- Cornacchio, J.V.: System complexity – a bibliography. *International Journal of General Systems*, 1977, 3, 267-271.
- Davis, G.A.: Current status of research and theory in human problem solving. *Psychological Bulletin*, 1966, 66, 36-54.
- Dörner, D.: *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer 1976.
- Dörner, D.: Über die Schwierigkeiten menschlichen Umgangs mit Komplexität. *Psychologische Rundschau*, 1981, 32, 163-179.
- Dörner, D.: Wie man viele Probleme zugleich löst – oder auch nicht! *Sprache & Kognition*, 1982, 1, 55-66.
- Dörner, D.: Denken, Problemlösen und Intelligenz. In Lüer, G. (Ed.), 1983a, a.a.O., 354-367.
- Dörner, D.: Schriftliche Mitteilung, 1983b.
- Dörner, D. & Kreuzig, H.W.: Problemlösefähigkeit und Intelligenz. *Psychologische Rundschau*, 1983, 34, 185-192.
- Dörner, D., Kreuzig, H.W., Reither, F. & Stäudel, T. (Eds.): *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Huber 1983.
- Elgood, C.: *Handbook of Management Games*. Aldershot, Hampshire: Gower Publishing 1982<sup>2</sup>.
- Erdfelder, E.: Zur Bedeutung und Kontrolle des Beta-Fehlers bei der inferenzstatistischen Prüfung log-linearer Modelle. *Zeitschrift für Sozialpsychologie*, 1984, 15, 18-32.
- Flavell, J.H.: Metacognition and cognitive monitoring. A new area of cognitive-developmental inquiry. *American Psychologist*, 1979, 34, 906-911.
- Fuhrer, U.: Defiziente und sub-optimale Strategien des Handelns im Umgang mit hoher Aktivationskomplexität. *Psychologische Beiträge*, 1982, 24, 583-600.
- Funke, J.: Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Prädiktor? *Diagnostica*, 1983, 29, 283-302.
- Funke, J. & Hussy, W.: Komplexes Problemlösen: Beiträge zu seiner Erfassung sowie zur Frage der Bereichs- und Erfahrungsabhängigkeit. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 1984, 31, 19-38.
- Goldberger, A. & Duncan, O. (Eds.): *Structural Equation Models in the Social Sciences*. New York: Seminar Press 1973.
- Hager, W. & Westermann, R.: Die Eile – 10 Jahre danach. *Zeitschrift für Sozialpsychologie*, 1982, 13, 250-252.
- Hand, H. & Sims, H.: Statistical evaluation of complex gaming performance. *Management Science*, 1975, 21, 708-717.
- Hesse, F.W.: Alternative Ansätze zur Entwicklung heuristischer Strategien für den Bereich des schlußfolgernden Denkens. In Ueckert, H. & Rhenius, D. (Eds.), 1979, a.a.O., 153-161.
- Hesse, F.W.: Effekte des semantischen Kontexts auf die Bearbeitung komplexer Probleme. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 1982, 29, 62-91.
- Hesse, F.W., Spies, K. & Lüer, G.: Einfluß motivationaler Faktoren auf das Problemlöseverhalten im Umgang mit komplexen Problemen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 1983, 30, 400-424.
- Heuser, J.: Zur differentiellen Wirkung von Streß auf das Problemlösen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 1978, 25, 379-406.
- Huber, G.L. & Mandl, H. (Eds.): *Verbale Daten: eine Einführung in die Grundlagen und Methoden der Erhebung und Auswertung*. Weinheim: Beltz 1982.
- Jäger, A.O.: Intelligenzstrukturforschung: konkurrierende Modelle, neue Entwicklungen, Perspektiven. In Lüer, G. (Hrsg.), 1983, a.a.O., 339-353.

- Johnson, D.M.: *Systematic Introduction to the Psychology of Thinking*. New York: Harper and Row 1972.
- Klir, G.J.: The polyphonic general systems theory. In: Klir, G.J. (Ed.), 1972a, a.a.O., 1-18.
- Klir, G.J. (Ed.): *Trends in General Systems Theory*. New York: Wiley 1972b.
- Klix, F.: Kognitive Psychologie als integrativer Bestandteil psychologischer Grundlagenforschung: Einführung. In Ueckert, H. & Rhenius, D. (Eds.), 1979, a.a.O., 16-19.
- Klix, F.: Die Allgemeine Psychologie und die Erforschung kognitiver Prozesse. *Zeitschrift für Psychologie*, 1980, 188, 117-139.
- Kluwe, R.: *Wissen und Denken*. Stuttgart: Kohlhammer 1979.
- Kluwe, R.H. and Friedrichsen, G.: Mechanisms of control and regulation in problem solving. In Beckmann, J. and Kuhl, J. (Eds.), *Action-Control: From Cognition to Behavior*. New York: Springer 1984 (in press).
- Kluwe, R. & Reimann, H.: Problemlösen bei vernetzten, komplexen Problemen: Effekte des Verbalisierens auf die Problemlöseleistung. Hamburg: Bericht aus dem Fachbereich Pädagogik der Hochschule der Bundeswehr (unveröffentlichtes Manuskript) 1983.
- Kreuzig, H.W.: Computer-Simulation als diagnostisches Instrument. In Lürer, G. (Ed.), 1983, a.a.O., 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*, 1983, 2, 228-253.
- Langeheine, R.: Kausalanalyse qualitativer Daten mit manifesten und latenten Variablen. *Zeitschrift für Sozialpsychologie*, 1982, 13, 163-176.
- Langfeldt-Nagel, M.: Untersuchungen zur Konstruktvalidität der Grundintelligenztests (CFT) von Cattell und Weiss. *Diagnostica*, 1982, 28, 65-79.
- Luenberger, D.L.: *Introduction to Dynamic Systems: Theory, Models and Applications*. New York: Wiley 1979.
- Lürer, G. (Ed.): *Bericht über den 33. Kongress der Deutschen Gesellschaft für Psychologie in Mainz 1982*. Band 1 und 2. Göttingen: Hogrefe 1983.
- Luhmann, N.: Die Programmierung von Entscheidungen und das Problem der Flexibilität. In Mayntz, R. (Ed.), *Bürokratische Organisation*. Köln: Kiepenheuer & Witsch, 1968, 324-341.
- Mandl, H. & Huber, G.L. (Eds.): *Emotion und Kognition*. München: Urban & Schwarzenberg 1983.
- Moray, N. (Ed.): *Mental Workload. Its Theory and Measurement*. New York: Plenum Press 1979.
- Norris, D.R. & Snyder, C.A.: External validation of simulation games. *Simulation & Games*, 1982, 13, 73-85.
- Opp, K.-D.: *Methodologie der Sozialwissenschaften. Einführung in Probleme ihrer Theoriebildung*. Reinbek: Rowohlt 1970.
- Osterloh, M.: *Handlungsspielräume und Informationsverarbeitung*. Bern: Huber 1983.
- Pichler, F.: *Mathematische Systemtheorie*. Berlin: de Gruyter 1975.
- Putz-Osterloh, W.: Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg. *Zeitschrift für Psychologie*, 1981, 189, 79-100.
- Putz-Osterloh, W.: Über Determinanten komplexer Problemlöseleistungen und Möglichkeiten zu ihrer Erfassung. *Sprache & Kognition*, 1983, 2, 100-116.
- Putz-Osterloh, W. & Lürer, G.: Wann produzieren Probanden räumliche Vorstellungen beim Lösen von Raumvorstellungsaufgaben? *Zeitschrift für Experimentelle und Angewandte Psychologie*, 1979, 26, 138-156.
- Putz-Osterloh, W. & Lürer, G.: Über die Vorhersagbarkeit komplexer Problemlöseleistungen durch Ergebnisse in einem Intelligenztest. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 1981, 28, 309-334.
- Rasmussen, J. & Rouse, W.B. (Eds.): *Human Detection and Diagnosis of System Failures*. New York: Plenum Press 1981.
- Reimann, H. & Kluwe, R.: Effekte des Verbalisierens auf die Problemlöseleistung beim Umgang mit komplexen Systemen. Vortrag gehalten auf dem 25. Tagung experimentell arbeitender Psychologen vom 27.-31.3.1983 in Hamburg.
- Reinhardt, F. & Soeder, H.: *dtv-Atlas zur Mathematik. Tafeln und Texte. Band I. Grundlagen, Algebra und Geometrie*. München: Deutscher Taschenbuch Verlag 1982.
- Reither, F.: Über die kognitive Organisation bei der Bewältigung von Krisensituationen. In Ueckert, H. & Rhenius, D. (Eds.), 1979, a.a.O., 210-222.
- Ruben, B.D. & Lederman, L.C.: Instructional simulation gaming: validity, reliability, and utility. *Simulation & Games*, 1982, 13, 233-244.
- Schulz, P.: Regulation und Fehlregulation im Verhalten. II. Streß durch Fehlregulation. *Psychologische Beiträge*, 1979, 21, 597-621.
- Schulz, P.: Regulation und Fehlregulation im Verhalten. V. Die wechselseitige Beeinflussung von mentaler und emotionaler Beanspruchung. *Psychologische Beiträge*, 1980, 22, 633-656.
- Siljak, D.D.: *Large-Scale Dynamic Systems. Stability and Structure*. New York: North Holland 1978.
- Sommer, R. & Sommer, B.A.: Mystery in Milwaukee: Early intervention, IQ, and psychology textbooks. *American Psychologist*, 1983, 38, 982-985.
- Speedie, S.M. Treffinger, D.J. & Houtz, J.C.: Classification and evaluation of problem solving tasks. *Contemporary Educational Psychology*, 1976, 1, 52-75.
- Steyer, R.: Modelle zur kausalen Erklärung statistischer Zusammenhänge. In Bredenkamp, J. & Feger, H. (Eds.), *Enzyklopädie der Psychologie: Forschungsmethoden der Psychologie*. Band 4. Göttingen: Hogrefe 1983, 59-153.
- Tack, W.H.: Zur Theorie psychometrischer Verfahren. Formalisierung der Erfassung von Situationsabhängigkeit und Veränderung. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 1980, 1, 87-106.
- Thom, R.: *Stabilität structurelle et morphogénèse*. Reading, Mass.: Benjamin 1972.
- Thompson, K.R. & Keon, T.L.: The creation of certain and uncertain environments. A study using a management simulation. *Simulation & Games*, 1982, 13, 437-450.
- Ueckert, H. & Rhenius, D. (Eds.): *Komplexe menschliche Informationsverarbeitung. Beiträge zur Tagung «Kognitive Psychologie» in Hamburg 1978*. Bern: Huber 1979.
- Vemuri, V.: *Modeling of Complex Systems. An Introduction*. New York: Academic Press 1978.
- Weinberg, G.M.: A computer approach to general systems theory. In Klir, G.J. (Ed.), 1972, a.a.O., 98-141.
- Weinert, F.E. & Kluwe, R. (Eds.): *Metakognition, Motivation und Lernen*. Stuttgart: Kohlhammer 1984.
- Wolfe, J.: The effects of game complexity on the acquisition of business policy knowledge. *Decision Sciences*, 1978, 9, 143-155.
- Zeeman, E.C.: Catastrophe theory. *Scientific American*, 1976, 234, No. 4, 65-83.

- Zimmerman, D.W. & Williams, R.H.: The theory of test validity and correlated errors of measurement. *Journal of Mathematical Psychology*, 1977, 16, 135-152.
- Zuckerman, D.W. & Horn, R.E.: *The Guide to Simulations/Games for Education and Training*. Lexington, Mass.: Informational Resources 1973.

Zwicker, E.: *Simulation und Analyse dynamischer Systeme in Wirtschafts- und Sozialwissenschaften*. Berlin: de Gruyter 1981.

Dipl.-Psych. Joachim Funke, Fachbereich I - Psychologie - der Universität Trier, Postfach 3825, D-5500 Trier

## Einige Anmerkungen zu Funkes «Diagnose der westdeutschen Problemlöseforschung in Form einiger Thesen»

Als einem der Hauptbetroffenen der Funkeschen Diagnose und Kritik seien mir einige Anmerkungen zu den Ausführungen von Funke erlaubt; ich möchte dabei nicht so sehr die kritischen Anmerkungen von Funke einer Kritik unterziehen, sondern einige Punkte klarstellen, andere herausarbeiten, vor allem um damit den weiterführenden Charakter der Funkeschen Diagnose zu unterstreichen.

Zunächst ist anzumerken, daß viele Punkte, die Funke kritisiert, ganz einfach auf die üblichen Zeitverzögerungen bei Publikationen, insbesondere bei Buchpublikationen, zurückzuführen sind. So existiert z.B. inzwischen eine (Funke bekannte) Theorie der Absichtsbehandlung (Vorg-Modell), die die von Funke zu Recht bemängelten formalen Unzulänglichkeiten der bisher publizierten Theorien zu diesem Gebiet vermeidet. Zu These 2 wäre in diesem Zusammenhang anzumerken, daß in der 1978 beendeten Lohhausen-Studie kaum Publikationen berücksichtigt werden konnten, die 1978 bzw. später erschienen sind. Bestimmte Aspekte der «Metakognition» z.B. finden in unseren Arbeiten seit jeher Beachtung; dies gilt beispielsweise für das Thema Selbstreflexion. Hierfür haben wir bereits 1978 (Dörner) ein Modell vorgelegt; viele Aspekte metakognitiver Prozesse werden in der Lohhausen-Studie explizit angesprochen, allerdings ohne Bezug auf Brown (1978), da uns diese Aspekte menschlicher kognitiver Prozesse auch vor der Lektüre der Arbeit von Brown bereits aufgefallen waren. Auch dem Thema «Sprache und Denken» haben wir uns bereits früh zugewandt

(Dörner, 1976, publiziert 1981), allerdings gehen wir gerne zu, daß in dieser Beziehung differenziertere Arbeit zu leisten bleibt. Aus den in der genannten Arbeit geschilderten Gründen halten wir es auch nicht für sehr glücklich, daß man in der Intelligenzforschung versucht, Testmaterial weitgehend sprachfrei zu gestalten.

In seiner These 4 schlägt Funke einen weitreichenderen Gebrauch des systemtheoretischen Ansatzes vor. Wir halten diesen Vorschlag für richtig, obwohl wir beispielsweise Zweifel daran haben, ob man bei komplizierteren Systemen z.B. das von Funke innerhalb der These 4 vorgeschlagene Verfahren zur Erfassung der Kontrollkomplexität bei umfangreicheren Systemen verwenden kann. Was wird z.B. aus der Tabelle 1 von Funke, wenn man nicht nur eine Störvariable mit drei verschiedenen Zuständen und eine Kontrollvariable mit drei verschiedenen Zuständen hat, sondern sagen wir zehn Kontrollvariablen mit jeweils 100 verschiedenen Zuständen und eine entsprechende Menge von Störvariablen? Und was wird aus der Tabelle, wenn man nicht nur drei Systemzustände hat, sondern einige Millionen? Systeme dieser Größenordnung sind aber keineswegs ungewöhnlich.

Sehr wichtig erscheint uns der Hinweis auf die Notwendigkeit, das Problem der Meßfehler zu diskutieren, welches Funke in seiner These Nr. 6 anspricht. Hier gibt es sicherlich generell bei der Erforschung des Denkens und Problemlösens ernsthafte Probleme. Man wird meistens nicht z.B. die Reliabilität eines Gütemaßes dadurch bestimmen können, daß man die glei-

chen Versuchspersonen noch einmal vor das gleiche Problem setzt. Dies verbietet sich aus denselben Gründen, aus denen es sich verbietet, die «Witzigkeit» eines Witzes durch Lautstärke und Dauer des Lachens zu erfassen, um dann die Reliabilität dieses Maßes dadurch zu erfassen, daß man dem gleichen Probanden denselben Witz noch einmal erzählt. Dasselbe Problem ist beim zweiten Mal für dieselbe Versuchsperson nicht mehr dasselbe Problem. Vernünftig und aussichtsreich aber erscheint es uns, die Reliabilität bestimmter elementarer Verhaltensmerkmale der Pbn beim Problemlösen festzustellen. Hier bieten sich solche Parameter wie «Anzahl der Fragen eines bestimmten Typs», «Anzahl der Entscheidungen», «Veränderung der Anzahl von Entscheidungen im Verlauf des Versuchs» an. Vielleicht gibt es auch noch elementarere Parameter, wie z. B. die «Stillzeiten» nach einer Informationsgabe als Maß für die Verarbeitungstiefe der Information oder ähnliches. Im ganzen wird man mehr versuchen müssen, Parameter des «erzeugenden Systems» verlässlich zu erfassen, als die Effekte (z. B. in Maßen der Problemlösegröße). (In

diesem Zusammenhang sei darauf hingewiesen, daß sich in der Lohhausen-Studie bemerkenswerte Konstanz hinsichtlich der Verhaltensunterschiede zwischen den «guten» und den «schlechten» Versuchspersonen zeigten. Siehe Dörner et al. 1983, p. 219, 222, 223, 224, 229, 236, 238f., 244, 270.) Die Konsistenz der Verhaltensunterschiede deutet u. E. auf eine hohe Reliabilität der erfaßten Merkmale hin.

D. Dörner

#### Literatur

- Brown, A.L.: Knowing when, where and how to remember: A problem of metacognition. In Glaser, R. (Ed.), *Advances in Instructional Psychology*. Hillsdale N.J.: Lawrence Erlbaum 1978.
- Dörner, D.: Self reflection and problem solving. In Klix, F. (Ed.), *Human and Artificial Intelligence*. Berlin: Deutscher Verlag der Wissenschaften 1978.
- Dörner, D.: Sprache und Denken. In Scharf, J. H. (Ed.), *Naturwissenschaftliche Linguistik*. Nova Acta Leopoldina. Halle (Saale): Leopoldina 1981.
- Dörner, D., Kreuzig, H.W., Reither, F. & Stäudel, Th. (Hrsg.): *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Huber 1983.