

Ruprecht-Karls-Universität Heidelberg
Psychologisches Institut

Die Reaktivität von Kausaldiagramm-Analysen beim komplexen Problemlösen

Eine experimentelle Untersuchung anhand des komplexen,
dynamischen Systems ÖKOSYSTEM

Diplomarbeit
Leimen, Februar 2006

vorgelegt von
Christine Blech
Sudetenweg 8
69181 Leimen

Erstbegutachter: Prof. Dr. J. Funke (Allgemeine und Theoretische Psychologie)

Zweitbegutachter: Dr. Jörg Zumbach (Pädagogische Psychologie)

Inhaltsverzeichnis

1	Einleitung	7
2	Komplexes Problemlösen	8
2.1	Gegenstandsbestimmung	9
2.2	Computersimulierte Szenarios als Methode	10
2.2.1	Realitätsnahe Szenarios	11
2.2.2	Szenarios auf der Basis formaler Modelle	13
2.3	Der DYNAMIS-Ansatz	14
2.3.1	Wissenserwerb	17
2.3.2	Wissensanwendung	22
2.3.3	Schwächen und Risiken des DYNAMIS-Ansatzes	23
3	Determinanten im Problemlöseprozess	24
3.1	Wissenserwerb und steuerndes Handeln	24
3.2	Systemfaktoren	26
3.3	Situationale Faktoren	28
3.3.1	Art der Informationsdarbietung	28
3.3.2	Erweiterte Aufgabenanforderungen	31
3.4	Personenfaktoren	36
3.4.1	Intelligenz	36
3.4.2	Spontaner Strategieeinsatz	37
3.4.3	Motivation und Emotion	38
3.5	Hypothesen	39
4	Methoden	40
4.1	Versuchsdesign	40
4.2	Materialien	42
4.2.1	Beschreibung der verwendeten Simulation	42
4.2.2	Abfrage von Strukturwissen	43
4.2.3	Abfrage von Oberflächenwissen	43

4.2.4	Abfrage emotionaler Zustände	44
4.2.5	Schulische Leistungen und soziodemografische Variablen	45
4.3	Teilnehmer	45
4.3.1	Vorüberlegungen zur Stichprobengröße	45
4.3.2	Stichprobenauswahl und -zusammensetzung	46
4.4	Versuchsdurchführung	48
4.5	Datenanalyseverfahren und -instrumente	51
5	Ergebnisse	52
5.1	Abhängige Variablen	52
5.1.1	Güte der Kausaldiagramme	52
5.1.2	Güte des Oberflächenwissens	52
5.1.3	Güte der Systemsteuerung	53
5.1.4	Güte der Strategieranwendung	53
5.1.5	Bearbeitungsdauer	54
5.2	Systemwissen und Systemsteuerung	54
5.2.1	Lerneffekte und Leistungen	54
5.2.2	Zusammenhänge zwischen Strukturwissen und Systemsteuerung	56
5.3	Vernetztheit	59
5.4	Kausaldiagramm-Analysen, Problemlösegüte und Problemlöseverhalten	60
5.4.1	Systemwissen	60
5.4.2	Steuerleistung	61
5.4.3	Strategieinsatz	62
5.4.4	Bearbeitungsdauer	63
5.5	Schulischer Erfolg, soziodemografische Variablen und Problemlöseerfolg	64
5.6	Emotionen und Problemlösegüte	67
6	Diskussion	69
	Literaturverzeichnis	79
	Anhang	84

Tabellenverzeichnis

1	Überblick über den Versuchsablauf	50
2	GdS, GdK _{sum} und GdOW im Verlauf der fünf Durchgänge	57
3	GdK _{sum} , GdS und Netto-Bearbeitungszeit im ANOVA-Design	61
4	Zwei Regressionsmodelle zur Vorhersage von GdK _{sum}	66
5	Zwei Regressionsmodelle zur Vorhersage von GdS	66
6	Geschlechtereffekte für GdK _{sum} und GdS.	66
7	Varianzaufklärung durch die Regressionsmodelle für GdK _{sum} und GdS . .	67

Abbildungsverzeichnis

1	Struktur des DYNAMIS-Systems ÖKOSYSTEM	16
2	Benutzeroberfläche der Simulation ÖKOSYSTEM	43
3	Benutzeroberfläche des computerbasierten Rekognitionstests	44
4	Pfadmodell zur Vorhersage von Steuerleistung aus Identifikationsleistung	58
5	GdK _{rel} , GdK _{vor} und GdK _{num} unter den zwei Vernetztheitsbedingungen .	59
6	Deskriptive Strategieberichte	63
7	Emotionsindizes im Verlauf des Versuchs	69

Zusammenfassung

Ein Experiment sollte den Einfluss von Kausaldiagramm-Analysen auf Problemlöseverhalten und -erfolg im Umgang mit komplexen, dynamischen Systemen untersuchen. Es wurde vermutet, dass die wissensdiagnostische Methode prozessbegleitender Kausaldiagramm-Analysen einen hypothesengeleiteten Problemlösestil sensu Klahr und Dunbar (1988) unterstützt und intensiviert. Als beobachtbare Konsequenz intermittierender Kausaldiagramm-Analysen wurde erhöhter Erwerb von Strukturwissen und gesteigerte Leistung im Steuern des Systems erwartet. 64 studentische Versuchspersonen nahmen an der Untersuchung teil. Sie bearbeiteten das DYNAMIS-Szenario ÖKOSYSTEM über fünf kombinierte Explorations- und Anwendungsdurchgänge.

Probanden, die in den Intervallen zwischen diesen Durchgängen Kausaldiagramme anfertigten, zeigten sich im abschließend erworbenem Strukturwissen Vergleichspersonen überlegen, die einen oberflächlichen Rekognitionstest, eine nicht szenariobezogene Aufgabe oder gar keine zusätzliche Aufgabe absolviert hatten. Dies spricht für eine Reaktivität von Kausaldiagramm-Analysen. Keine definitiven Hinweise auf Reaktivität hingegen erbrachte die Analyse der Bearbeitungszeiten und des Strategieeinsatzes im Explorieren des Systems: Kausal-instruierte Probanden benötigten weder mehr Zeit noch explorierten sie ÖKOSYSTEM geschickter. Ein Pfadmodell belegte wie vermutet einen hohen positiven, prädiktiven Zusammenhang des Strukturwissens auf die Steuerleistung; dennoch ging der Wissensvorteil bei Kausaldiagramm-Analysen nicht mit verbesserter Steuerleistung einher. Inwieweit spezifisches Eingriffswissen alternativ zum abstraktem Strukturwissen am Steuerprozess beteiligt ist, konnte im Rahmen des Experiments nicht geklärt werden.

Als weitere Determinanten auf die Problemlösequalität erwiesen sich der Vernetztheitsgrad des Systems (s. Funke, 1985) und das Geschlecht der Versuchspersonen. Allgemeine und spezifisch mathematische intellektuelle Leistungsfähigkeit hatten entgegen der Erwartung keinen nachweisbaren Einfluss auf den Problemlöseerfolg.

Abstract

An experiment was designed to investigate the influence which analyses of causal diagrams have on dealing with complex dynamic systems. It was hypothesised that the method of diagnosing structural knowledge by causal diagrams, applied during a dynamic task, would support and intensify problem solving processes guided by the formation of hypotheses sensu Klahr and Dunbar (1988). In consequence, enhanced structural knowledge and increased control performance should occur when causal diagrams are provided. 64 students volunteered to participate in the experiment. They worked on the DYNAMIS computer simulation ECOSYSTEM for five cycles of combined exploration and application.

Concerning the structural knowledge finally acquired, subjects confronted with causal diagrams during the intervals between the five cycles proved superior to control subjects who had been dealing with either a test of mere recognition, a task not related to the scenario or with no extra task at all. This finding has been interpreted in terms of causal diagrams being reactive measures. There was, however, no definite evidence of reactivity when analysing the time spent on the task or the strategies employed in exploration. Neither did subjects instructed with causal diagrams take more time in coping with ECOSYSTEM nor did they explore the system more skillfully. As expected, a path-analytical model hinted at a highly positive predictive effect of structural knowledge on control performance; yet increased knowledge associated with causal diagrams was not accompanied by improved control performance. Searching for alternative predictors of system control, the experiment could hardly elucidate the role of specific operational knowledge. The degree of connectivity implemented in the system (see Funke, 1985) as well as the subjects' sex turned out to be further determinants of the quality of complex problem solving. Against expectation, there was no such effect of general intellectual or specifically mathematical abilities.

1 Einleitung

Immer wieder, im Alltag, in der Wissenschaft, in der Politik, stoßen Menschen auf Gegebenheiten, die zu Veränderungen herausfordern. Ein funktionsuntüchtiges technisches Gerät soll repariert, ein Forschungsprojekt geplant und durchgeführt, eine politische Krise entschärft werden. Einfache Routinemaßnahmen genügen selten, um dies zu erreichen. Vielmehr sind verschiedene, aufeinander abgestimmte Operationen gefordert, die sich den jeweils aktuellsten Situationen in einer zeitlichen Entwicklung anpassen. Eine Person muss ganze Sequenzen von zielführenden Handlungsschritten vorausschauend überblicken, will sie sich nicht auf zufällige Erfolge durch planloses Ausprobieren verlassen.

Wie Menschen komplexe Probleme bewältigen, ist für den Psychologen mindestens so interessant wie für die problemlösende Person. Anstatt eine Lösung nur nach ihrem Endergebnis zu beurteilen (inwieweit wurde das angestrebte Ziel erreicht?), fragt die Problemlöseforschung dabei auch nach dem Lösungsweg bzw. dem Problemlöseprozess. Mit Hilfe von computersimulierten Szenarios, Nachbildungen realer komplexer Probleme in der virtuellen Laborwelt, lassen sich schrittweise Annäherungen an das Lösungsziel beobachten. Dem Paradigma des Kognitivismus folgend hat die Prozessforschung seit den siebziger Jahren ihre Perspektive erweitert: Sie erklärt den beobachtbaren Lösungsprozess auf der Grundlage von inneren mentalen Repräsentationen, die im Kopf des Problemlösers entstehen bzw. entstehen können. Zunächst ist die Art dieser Repräsentationen ein hypothetisches Konstrukt des Forschers.

Im DYNAMIS-Forschungsansatz, der komplexe Problemlösesituationen als Systeme vernetzter Variablen mit kausalen Abhängigkeiten simuliert, vermutet man, dass die subjektiven, internen Problemrepräsentationen einer Person Wissensnetzwerke (Kausalmodelle) sind, im formalen Aufbau dem Konstruktionsprinzip der DYNAMIS-Simulation vergleichbar (vgl. Funke, 1985). Einblick in diese angenommenen subjektiven Kausalmodelle erlaubt ein standardisiertes, grafisches Abfrageinstrument, die Methode der Kausaldiagramm-Analyse. Ein Kausaldiagramm erfasst sämtliches Wissen über kausale Zusammenhänge, die ein Problemlöser glaubt, über ein komplexes System herausgefunden zu haben.

In der Empirie gibt es keinen Zweifel darüber, dass Personen Kausaldiagramme relativ erfolgreich bearbeiten können. Umstritten ist, inwieweit dieses Können auf spontan generierte, kausale Repräsentationen zurückgeht. Entsteht ein subjektives Modell kausaler Systemabhängigkeiten, sobald eine Person begonnen hat, die Problemlösesituation frei zu erkunden? In diesem Fall würden Kausaldiagramm-Analysen nicht mehr als das bereits vorhandene und verfügbare Wissen erfassen. Oder beginnt eine Person erst dann über abstrakte Kausalzusammenhänge nachzudenken, wenn sie ausdrücklich dazu aufgefordert wird? In diesem Fall entstünde das interne subjektive Kausalmodell als Reaktion auf die Kausaldiagramm-Analyse.

Die Frage, ob Kausaldiagramm-Analysen ein reaktives Messverfahren darstellen, ist theoretisch bekannt (s. Kluwe, 1988), aber empirisch bislang nicht geprüft. Die vorliegende Arbeit beschreibt daher erstmals ein Experiment, das untersucht, inwieweit Kausaldiagramm-Analysen Einfluss auf Problemlösequalität und Problemlöseverhalten nehmen. Den Ausgangspunkt und Orientierungsrahmen bildet eine Studie von Funke (1985) zum Umgang mit dem komplexen, dynamischen Szenario ÖKOSYSTEM, die in ihren Grundzügen repliziert und um die neue Fragestellung erweitert wird. Eingebunden ist die empirische Arbeit in einen theoretischen Überblick über Gegenstand und Methodik der komplexen Problemlöseforschung mit besonderem Augenmerk auf den DYNAMIS-Ansatz und die Kausaldiagramm-Analyse.

2 Komplexes Problemlösen

Das Überblickskapitel zum komplexen Problemlösen stellt aufbauend auf Definitionen der Begriffe *Problem*, *komplexes Problem* und *komplexes Problemlösen* insbesondere die Methoden der Problemlöseforschung heraus. Mit realitätsnahen Szenarios einerseits und Szenarios auf der Basis formaler Modelle andererseits werden zwei große Klassen computersimulierter Szenarios, die zugrunde liegenden Forschungsziele und ihre praktische Umsetzung dargestellt. Ein Unterkapitel über Szenarios nach dem formalen DYNAMIS-Ansatz behandelt die wissensdiagnostische Methode der Kausaldiagramm-Analyse.

2.1 Gegenstandsbestimmung

So verbreitet der Begriff *Problem* in der Alltagsauffassung ist, so schwierig scheint es, eine verbindliche Definition für die Wissenschaft zu finden. Zahlreiche unterschiedliche Definitionsvorschläge existieren, deren einziges und zugleich wesentliches gemeinsames Merkmal es ist, ein Problem formal als eine zu überwindende Distanz oder Barriere zu umschreiben (s. Frensch & Funke, 1995). Eine bestehende Ausgangssituation, etwa ein defekter Fernseher, fordert zu einer Veränderung heraus, um einen angestrebten Zielzustand, etwa das Funktionieren des Fernsehers, zu erreichen. Bevor der Zielzustand eintritt, muss eine Barriere aufgehoben, muss das Problem gelöst werden. Im Gegensatz zu einer Aufgabe, die mit bekannten (Routine-)Mitteln zu bewerkstelligen ist, impliziert jedes Problem eine neuartige Anforderung. Das Mittel, das zum Zielzustand führt, muss erst gefunden, dann angewendet werden. Inwieweit zielführende Mittel neu oder bekannt und erprobt sind, hängt selbstverständlich nicht nur von der Problemsituation, sondern auch vom Können und den Vorerfahrungen des Problemlösers ab.

Unter *komplexen Problemen* verstehen Frensch und Funke (1995) Probleme, deren Bewältigung mehr als nur einen Lösungsschritt erfordert. Der Unterschied zwischen einem einfachen Problem und den multiplen Barrieren eines komplexen Problems liegt üblicherweise nicht nur in einem quantitativ erhöhten Lösungsaufwand. Vielmehr sind komplexe Probleme durch typische Eigenschaften gekennzeichnet, die sie auch qualitativ von einstufigen Problemen abheben: Komplexe Probleme verfügen neben ihrer formalen Komplexität über eine (Eigen-)Dynamik, die Veränderungen der Problemsituation während des Lösungsprozesses bedingt. Genaue Eigenschaften des Problems sind der problemlösenden Person anfangs nicht bekannt; man spricht von intransparenten Systemen. Auch der Zielzustand ist selten auf ein einzelnes Merkmal beschränkt, sondern kann sich aus vielen Zielen zusammensetzen (Polytelie), die im Verlauf des Problemlösens simultan zu berücksichtigen und eventuell gegeneinander abzuwägen sind (s. Dörner, Kreuzig et al., 1983).

Im Hinblick auf den aufwändigen Lösungsprozess komplexer Probleme – Informationsbeschaffung, Planen, Abwägen, Entscheiden u. a. m. – stimmen Theoretiker überein, dass kognitive Aktivität am Problemlösen notwendig beteiligt ist. Ob komplexes Problemlö-

sen mit dem Gegenstand der Denkpsychologie schlechthin gleichzusetzen ist (vgl. Dörner 1976), ist sicherlich fraglich. Wie Nickerson (1998) illustriert, denken wir hinreichend oft auch ohne ein offensichtliches, funktionales Ziel zu verfolgen. Umgekehrt benötigen wir zur Lösung eines Problems neben dem Denktakt fast immer eine ausführende Handlung. Beide Elemente, Verhalten und Kognition finden sich in der zusammenfassenden Definition von Frensch und Funke (1995, S. 18) wieder: „CPS [Complex Problem Solving] occurs to overcome barriers between a given state and a desired goal state by means of behavioral and/or cognitive multistep activities.“

In derselben Arbeit systematisieren die Autoren zentrale Komponenten einer Theorie zum komplexen Problemlösen: die Merkmale des Problems als solches, die Rahmenbedingungen bzw. die Problemlösesituation und Personenmerkmale des Problemlösers. Die Komponenten werden im Abschnitt „Determinanten im Problemlöseprozess“ näher erläutert und – z. T. in Verbindung mit spezielleren funktionalistischen Theorien (z. B. Dual Space Theory; Klahr & Dunbar, 1988) – auf die eigene Fragestellung bezogen.¹

2.2 Computersimulierte Szenarios als Methode

Es stellt sich zunächst die Frage, wie komplexe, dynamische, polytelische Problemstellungen in der Forschung operationalisierbar sind. Ein v. a. in den Vereinigten Staaten verbreiteter Ansatz besteht darin, Problemlöseverhalten in natürlichen, praktisch relevanten Wissensdomänen (z. B. physikalisches Wissen, Schreiben, Schachprobleme) zu untersuchen, bevorzugt im Experten-Novizen-Vergleich (s. Sternberg, 1995). In der europäischen Problemlöseforschung hingegen hat sich die Entwicklung angemessener Problemstellungen seit den siebziger Jahren an den Möglichkeiten der Computertechnologie orientiert. Computersimulierte „Mikrowelten“, die einen oft semantisch reichhaltigen Bereich der Alltagswelt nachmodellieren, werden – zumeist unerfahrenen – Problemlösern ähnlich einem Rollenspiel vorgegeben. Versuchspersonen können im jeweiligen Gegenstandsbereich Entwicklungen beobachten, Zusammenhänge eigenständig explorieren,

¹Für eine umfassende Rahmentheorie zum komplexen Problemlösen einschließlich kognitiver Teilprozesse, Emotion, Motivation und Problemlöseverhalten ist auf Dörners Handlungsregulationstheorie zu verweisen (s. z. B. Dörner et al., 1988). Da die Theorie keinen unmittelbaren Beitrag zur empirischen Prüfung in dieser Arbeit leistet, unterbleibt eine Darstellung an dieser Stelle.

Entscheidungen z. T. unter Echtzeit-Bedingungen treffen, Eingriffe vornehmen und die Konsequenzen ihrer durchgeführten oder unterlassenen Maßnahmen bewerten, sobald das System Rückmeldung über Erfolge oder Misserfolge gibt.

Die ökologische Validität, d. h. die Nähe zu realen Problemen, ist insbesondere von den Vertretern dieses Ansatzes immer wieder hervorgehoben worden. Nach Brehmer und Dörner (1993) verbinden computersimulierte Szenarios die Vorteile von Feldstudien mit den Vorteilen experimenteller Laboranordnungen, ohne ihre Nachteile mit sich zu bringen. Im Gegensatz zu Einzelfallstudien wollen „Mikrowelten“ am Computer überschaubar und systematisch manipulierbar sein; zugleich wollen sie den Gegenstandsbereich weniger einschränken als Problemstellungen in klassischen Laborexperimenten (vgl. auch Dörner & Wearing, 1995).

Innerhalb einer großen Vielfalt computersimulierter Szenarios sind der Anspruch der Realitätsnähe bzw. Maßstabstreue der Modellwelt einerseits und der Anspruch systematischer Manipulierbarkeit bzw. formaler Beschreibbarkeit andererseits jedoch unterschiedlich stark gewichtet, abhängig von den jeweiligen Forschungszielen. Eine grobe Orientierung, um die Methodik der komplexen Problemlöseforschung zu klassifizieren, liefern Frensch und Funke (1995) sowie Buchner (1995), die ausgehend von zwei verschiedenen innereuropäischen Forschungstraditionen sogenannte realitätsnahe Szenarios von Szenarios auf der Basis formaler Modelle abgrenzen.

2.2.1 Realitätsnahe Szenarios

Realitätsnahe Mikrowelten sind insbesondere durch ihre semantische Reichhaltigkeit verbunden mit hohem technischem Aufwand bekannt geworden. Ein typisches, beinahe klassisches Beispiel ist das Szenario LOHHAUSEN von Dörner, Kreuzig et al. (1983). Das LOHHAUSEN-Szenario versetzt Probanden in die Rolle eines mitteleuropäischen Kleinstadt-Bürgermeisters. Gewählt für einen simulierten Zeitraum von zehn Jahren ist es Aufgabe dieses Bürgermeisters, so gut wie möglich für das Wohlergehen seiner Stadt zu sorgen. Mehr als 2000 Variablen, teils direkt manipulierbar, teils verdeckt als Vermittlervariablen, bestimmen die Zusammenhänge in der simulierten Kommune. Ähnlich bekannt und beliebt sind die realitätsorientierten Szenarios SCHNEIDERWERKSTATT

(s. Putz-Osterloh, 1981, Putz-Osterloh & Lüer, 1981), ein simulierter Textilmanufakturbetrieb, das Entwicklungshilfe-Szenario MORO (s. z. B. Schaub & Strohschneider, 1992) oder das Szenario FEUER (Brehmer, 1995), in dem Problemlöser virtuell Maßnahmen zur Brandbekämpfung vornehmen.

Trotz Variationen in der Semantik und z. T. sehr unterschiedlichem Umfang steht hinter diesen und anderen realitätsnahen Szenarios eine gemeinsame Forschungstradition. Zurückgehend auf Dörner interessieren in erster Linie interindividuelle Unterschiede im Kontrollieren komplexer, dynamischer Systeme (Buchner, 1995). Dieser differenzialpsychologisch ausgerichtete Ansatz erfasst neben kognitiven Aspekten auch motivationale und soziale Einflüsse auf den Problemlöser (Frensch & Funke, 1995) und richtet sein besonderes Augenmerk auf typische „Fehler“ im Umgang mit den komplexen, dynamischen Systemen. Kognitive Fehlreaktionen, so argumentiert Dörner, seien diagnostisch oft aufschlussreicher als „reibungslöses“, erfolgreiches Problemlösen (Dörner, 1981, S. 165).

Ob man aus charakteristischen Irrtümern tatsächlich Einsichten über allgemein gültige, kognitive Prozesse gewinnen kann, zweifeln Kritiker dieses Ansatzes an. Allgemeine Aussagen erscheinen problematisch auch insofern, als ein eindeutiges Diagnosekriterium optimalen Problemlösens fehlt. Woran etwa ist das Wohlergehen der Bürger in der simulierten Kleinstadt LOHHAUSEN zu messen: Sollen wirtschaftliche (Kapital, Produktion), soziale (Zahl der Arbeitslosen und Wohnungssuchenden) oder psychologische Faktoren (Zufriedenheit der Bevölkerung) ausschlaggebend sein? Selbst wenn klare, den Probanden vorab mitgeteilte Zielvorgaben bestehen, beinhalten die meisten realitätsnahen Szenarios einen hoch komplexen, kaum überschaubaren Problemraum, der es unmöglich macht, die zu einem bestimmten Zeitpunkt effizientesten, zielführenden Interventionen zu bestimmen (vgl. Buchner, 1995). Die hochgradige, doch formal nicht definierte Vernetztheit der Szenarios erlaubt zudem keine isolierte Betrachtung einzelner Funktionsbereiche des Systems, die Voraussetzung wäre, um Befunde für unterschiedliche realitätsnahe Szenarios miteinander zu vergleichen oder kognitive Teilprozesse zu isolieren.

Wie Dörner argumentiert, ist eine derartige, bereits durch das Versuchsmaterial vorgegebene Zerlegung des Problemlöseprozesses gerade *nicht* erwünscht. Um für einen

inhaltlich relevanten Problembereich ein realitätsgetreues Abbild komplexer Problemlöseprozesse zu erstellen, wäre es nach Dörner sogar abträglich, einzelne Funktionen isoliert zu betrachten. Schließlich gehe es um das Zusammenwirken aller beteiligten Prozesse, etwa des Sammelns von Informationen, des Bildens von Hypothesen, Planungen und Entscheidungen (vgl. Dörner, 1992, S. 59). Dass Dörners Forschungsphilosophie nur eine mögliche Antwort auf die Kritik an den beschriebenen Computersimulationen ist, zeigen die Vertreter der folgenden Klassen von Szenarios.

2.2.2 Szenarios auf der Basis formaler Modelle

Genaue formale Beschreibbarkeit, eindeutige Zielkriterien und Bestimmbarkeit optimaler Interventionen für jeden beliebigen Zeitpunkt im Problemlöseprozess sind die zentralen Ansprüche, die Szenarios auf der Basis formaler Modelle ausschließlich für sich geltend machen (Buchner, 1995). Davon unberührt bleiben Dynamik und Komplexität, die diese Ansätze mit den realitätsorientierten Szenarios teilen. Nach dem Prinzip „Zuerst die Struktur, dann der Inhalt“ wird zunächst ein mathematisches Modell vollständig definiert, bevor man einen geeigneten semantischen Kontext auswählt (s. Funke, 2003, S. 155). Rein abstrakte Szenarios wie fiktive Maschinen bei Beckmann (1994) oder Kröner (2001) zeigen, dass eine semantische Einbettung bei manchen Fragestellungen sogar ganz unterbleiben kann.

Wie angedeutet erlaubt eine solche formale Bestimmung der Problemstellung, das Versuchsmaterial systematisch zu manipulieren und so den Einfluss der Systemeigenschaften auf kognitive Prozesse zu untersuchen. Anstatt bei interindividuellen Unterschieden und Problemlösedefiziten liegt der Untersuchungsschwerpunkt dieses vorwiegend allgemeinpsychologisch ausgerichteten Ansatzes bei Problemlösekompetenzen. Ausgehend von Broadbents Studien zum impliziten vs. expliziten Gedächtnis (z.B. Berry & Broadbent, 1984) interessieren Lernvorgänge, Wissensrepräsentationen, die ein Problemlöser im Umgang mit einem komplexen, dynamischen System aufbaut und die Anwendung des erworbenen Wissens beim Steuern des Systems (vgl. Buchner, 1995).

Ungefähr zur gleichen Zeit wie die realitätsnahen Szenarios haben sich zwei Typen formaler Modelle in der kognitiven Problemlöseforschung etabliert: Finite Automaten und

Szenarios nach dem DYNAMIS-Ansatz. In beiden Fällen erlaubt ein einziges abstraktes Grundmodell, eine ganze Klasse vergleichbarer konkreter Szenarios zu entwickeln. Auch die zentralen Aufgabenanforderungen sind dieselben: Wissenserwerb und Wissensanwendung bzw. Steuerung des Systems durch Anstreben eines definierten Zielzustands. Der Unterschied zwischen finiten Automaten und DYNAMIS-Szenarios liegt im Skalenniveau der Systemvariablen. Während finite Automaten diskrete Variablen auf Nominalskalenniveau einsetzen (was zu einer endlichen oder finiten Menge an Systemzuständen führt), verwendet DYNAMIS kontinuierliche Variablen. Für die Diagnose des Wissenserwerbs bedeutet dies, dass Versuchspersonen, die finite Automaten explorieren, primär nach Paar-Assoziationen von relevanten Systemzuständen gefragt werden.²

Bei DYNAMIS-Szenarios ist mit Kausaldiagramm-Analysen eine differenziertere Abfrage möglich. Sie soll Rückschlüsse auf subjektive Kausalmodelle oder Wissensnetze zulassen. Kausaldiagramm-Analysen, die zusammen mit der DYNAMIS-Variante ÖKOSYSTEM von Funke (1985) die Grundlage der vorliegenden Arbeit bilden, stellen also eine diagnostische Besonderheit des DYNAMIS-Ansatzes dar. Obwohl finite Automaten den DYNAMIS-Szenarios an ökologischer Validität mindestens ebenbürtig sind (vgl. Funke & Buchner, 1992; Buchner & Funke, 1993) und obwohl sie sich prinzipiell dazu eignen, das Verhältnis zwischen Wissenserwerb und Wissensanwendung zu untersuchen, wird der DYNAMIS-Ansatz daher im Folgenden schwerpunktmäßig behandelt.

2.3 Der DYNAMIS-Ansatz

Komplexe, dynamische Systeme für kontinuierliche Variablen wurden erstmals von Funke (1985; s. auch Funke, 1992b, 1993) auf der Basis linearer Gleichungssysteme formal beschrieben und entwickelt. Bei den simultanen Gleichungen handelt es sich um üblicherweise deterministische, multivariate autoregressive Prozesse (*AR*-Prozesse), d. h. Zeitreihenmodelle, in denen jeder gegenwärtige Zustand eines Variablensatzes über lineare Abhängigkeiten von früheren Zuständen desselben Variablensatzes definiert ist

²Typische Wissensdiagnostik bei finiten Automaten legt den Testpersonen zwei Elemente aus einem Tripel von Ausgangszustand, Intervention und interventionsabhängigem Folgezustand vor; das fehlende dritte Element muss der Proband ergänzen (vgl. Funke & Buchner, 1992, S. 33).

(s. z. B. Rinne & Specht, 2002). Die Anzahl der Variablen ist prinzipiell beliebig. Auch autoregressive Abhängigkeiten sind zu frei wählbaren Graden möglich und können mit Hilfe einer speziellen DYNAMIS-Software in konkrete Simulationen umgesetzt werden (s. Funke, 1993, 2003; Software im Internet unter <http://dynamis.uni-hd.de/>). Psychologische Relevanz und praktische Handhabbarkeit haben jedoch dazu geführt, dass die Mehrzahl der DYNAMIS-basierten Simulationen nur Abhängigkeiten vom unmittelbar vorangehenden Zustand berücksichtigt (*AR-1-Prozesse*), in selteneren Fällen Abhängigkeiten vom Vorgänger (*AR-2-Prozesse*), und dass die meisten DYNAMIS-Szenarios nicht mehr als sechs oder acht Variablen enthalten. Das Beispiel der DYNAMIS-Variante ÖKOSYSTEM (Funke, 1985) soll die Verwendung multivariater autoregressiver Prozesse verdeutlichen und zugleich in die Grundbegriffe der DYNAMIS-Terminologie einführen. Bei ÖKOSYSTEM (s. Abb. 1) handelt es sich um ein semantisch eingebettetes DYNAMIS-Szenario, das eine Kleingärtnersituation simuliert. Das Ökosystem wird durch lineare, kausale Zusammenhänge zwischen sechs Variablen bestimmt: Gift, Schädlingsfresser, Dünger, Käfer, Wasserverschmutzung, Blätter. Drei dieser Variablen, die Variablen auf der linken Seite, von denen Pfeile ausgehen (Gift, Schädlingsfresser, Dünger), können vom Problemlöser durch unmittelbare Eingriffe verändert werden. Zwischen diesen Input-Variablen oder *exogenen Variablen* sind per Definition keine systemimmanenten, kausalen Abhängigkeiten möglich. Exogene Variablen wirken stets nur auf Output-Variablen oder *endogene Variablen*, die in der Grafik auf der rechten Seite dargestellt sind (Käfer, Wasserverschmutzung, Blätter). Die endogenen Variablen sind also nur mittelbar zu beeinflussen. Hierbei können *parallele Effekte* auftreten: Eine exogene Variable (Dünger) kann auf mehrere endogene Variablen (Wasserverschmutzung, Blätterzahl) Einfluss nehmen oder mehrere exogene Variablen (Gift, Dünger) können dieselbe endogene Variable (Wasserverschmutzung) beeinflussen. Neben den direkten Effekten durch exogene Variablen können endogene Variablen auch von anderen endogenen Variablen abhängen: Die endogene Variable Blätterzahl hängt einerseits von der exogenen Variable Dünger ab, andererseits von der endogenen Variable Käfer. Man spricht von *Nebenwirkungen*. Der kreisförmige Pfeil für die Variable Käfer zeigt an, dass endogene Variablen auch einen Einfluss „auf sich selbst“ haben, technisch gesprochen einer *Eigendynamik*

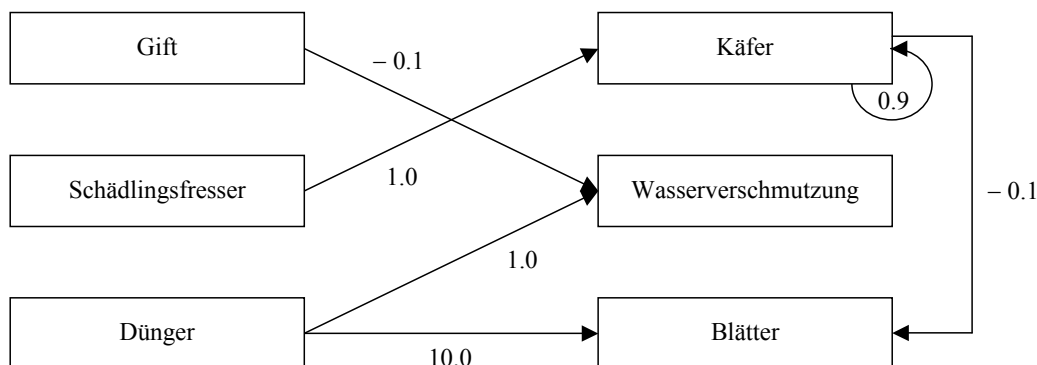


Abbildung 1: Struktur des DYNAMIS-Systems ÖKOSYSTEM. Exemplarisch dargestellt sind die Variablenbeziehungen für die Version mit sechs Teilprozessen für nicht zeitverzögerte Effekte.

unterliegen können. Ohne Einwirkung anderer Variablen kann sich der numerische Wert der Eigendynamik-Variable entweder erhöhen (Faktor größer als 1) oder erniedrigen (Faktor kleiner als 1). Keine eigendynamische Veränderung ist formal äquivalent zu einer Eigendynamik von 1 und wird normalerweise nicht dargestellt.

Ein Eigendynamik-Faktor von 0.9 wie im vorliegenden Beispiel bedeutet eine Multiplikation des numerischen Wertes mit dem Faktor 0.9, also eine Abnahme um 10 % auf 90 % des Ausgangswerts. Auch die anderen Zahlen an den Effektspfeilen geben multiplikative Gewichte an. Multipliziert wird jeweils der Wert einer beeinflussenden Variablen mit dem (entweder positiven oder negativen) Faktor, und dieses Produkt wird zum vorherigen Zustand der Output-Variablen addiert. Nehmen mehrere Variablen Einfluss auf eine endogene Variable, summieren sich die Produkte. Das folgende Gleichungssystem, eine alternative bzw. ergänzende Darstellungsform zum Diagramm, veranschaulicht dies. Jede endogene Variable wird durch eine Gleichung bestimmt (Wasserverschmutzung abgekürzt als „Wasser“):

$$\text{Käfer}_t = 1.0 \cdot \text{Schädlingsfresser}_{t-1} + 0.9 \cdot \text{Käfer}_{t-1} \quad (1)$$

$$\text{Wasser}_t = -0.1 \cdot \text{Gift}_{t-1} + 1.0 \cdot \text{Dünger}_{t-1} + 1.0 \cdot \text{Wasser}_{t-1} \quad (2)$$

$$\text{Blätter}_t = 10.0 \cdot \text{Dünger}_{t-1} - 0.1 \cdot \text{Käfer}_{t-1} + 1.0 \cdot \text{Blätter}_{t-1} \quad (3)$$

Die Indizes t und $t - 1$ geben an, auf welchen Zeitpunkt sich ein Variablenzustand bezieht. Kennzeichnet t den aktuellen Zustand einer Systemvariable, so ist $t - 1$ der direkt vorausgegangene Zeitpunkt auf einer diskreten Zeitachse. Abhängigkeiten von weiter zurückliegenden Systemzuständen $t - k$ sind in diesem Fall nur indirekt über $t - 1$ wirksam. Wie Buchner (1995) betont, wird der Zeitablauf des Systems in diesem wie auch in anderen zeitdiskreten DYNAMIS-Systemen durch die Eingriffe des Bearbeitenden gesteuert. Im Gegensatz zu Echtzeit-Systemen (s. z.B. das realitätsnahe Szenario FEUER bei Brehmer, 1995) ändert das System seinen Zustand nicht eigendynamisch nach einem festgelegten Zeitintervall. Ein neuer Takt beginnt erst, nachdem der Problemlöser Interventionen für alle gewünschten exogenen Variablen vorgenommen und diese dem System mitgeteilt hat.

Alle Eingriffe, die ein Problembearbeiter vornimmt, orientieren sich an zwei grundlegenden Aufgabenanforderungen: Wissenserwerb und Wissensanwendung. Beide Anforderungen sollen im Folgenden einschließlich diagnostischer Methoden erläutert werden.

2.3.1 Wissenserwerb

Wenn Versuchspersonen auch bei einem semantisch eingekleideten Szenario wie ÖKO-SYSTEM schon vor der Bearbeitung bestimmte Variablenzusammenhänge vermuten können, ist das System doch zunächst unbekannt. Erst die Aufforderung, es zu explorieren und Wissen zu erwerben, führt zu genaueren Kenntnissen. Nach Schoppek (2002) spielen drei Arten von Wissen eine Rolle: abstraktes Strukturwissen, spezifisches Eingriffswissen, auch Input-Output-Wissen genannt, und strategisches Wissen, das jedoch nicht zum Systemwissen im engeren Sinn gerechnet und erst in Abschnitt 3.3.2 erörtert werden soll. Die beiden anderen Wissenstypen, Strukturwissen und Eingriffswissen, gehen jeweils mit charakteristischen Diagnoseverfahren einher.

Strukturwissen und Kausaldiagramm-Analyse. Bei Kausaldiagramm-Analysen (Funke, 1985, 1992b) bildet das Strukturdiagramm eines DYNAMIS-Szenarios (s. Abb. 1)

den Bewertungsmaßstab für erworbenes Systemwissen. Nachdem Probanden erste Erfahrungen mit dem komplexen, dynamischen System gemacht haben, erhalten sie einen Vorlagebogen, der an Abb. 1 erinnert. Er enthält dieselbe Anordnung von Rechtecken mit einbeschriebenen Variablennamen, jedoch ohne die Pfeile und Zahlen. Gerade dies ist Aufgabe der Versuchsteilnehmer: entdeckte Kausalbeziehungen dem eigenen Kenntnisstand entsprechend durch formale Symbole einzutragen. Einfache Pfeile ohne zusätzliche Einträge geben an, dass eine vorhandene Relation erkannt wurde (*Relationswissen*). Ein Vorzeichen neben dem Pfeil steht für die Richtung des Zusammenhangs, die entweder gleichsinnig (+-Zeichen) oder gegensinnig (--Zeichen) ist (*Richtungswissen* oder *Vorzeichenwissen*).³ Wenn ein Proband darüber hinaus ein exaktes numerisches Gewicht herausfindet, trägt er auch die entsprechende Zahl im Kausaldiagramm ein (*Stärkewissen* oder *numerisches Wissen*). Aus Instruktionen sollte hervorgehen, dass Relations-, Richtungs- und Stärkeangaben selbstverständlich kombiniert werden können und dass jede Angabe idealerweise so präzise wie möglich ist.

Die Möglichkeit, erworbenes Systemwissen differenziert auf drei Genauigkeitsstufen zu untersuchen, ist ein großer Pluspunkt der Kausaldiagramm-Analyse. Er erklärt, warum die Methode, sei es im Paper-and-Pencil-Format oder in computer-unterstützten Modifikationen (s. Beckmann, 1994; Schulz, 2003), zu den üblichsten und beliebtesten wissensdiagnostischen Verfahren für DYNAMIS-Experimente zählt. Ein weiterer Vorteil ist, dass mit dem Maß *Güte der Kausaldiagramme* (GdK) von Funke (1992b) bzw. Müller (1993) ein hochreliables Wissensmaß entwickelt wurde (vgl. Müller, 1993). Berechnet werden können sowohl getrennte Strukturwerte für Relationswissen (GdK_{rel}), für Vorzeichenwissen (GdK_{vor}) und für numerisches Wissen (GdK_{num}) als auch ein Globalwert (GdK_{sum}), der sich additiv aus den gewichteten Komponenten zusammensetzt (Formeln nach Funke, 1992b, S. 83):

$$GdK_{rel} = \frac{(1 - p_{rel}) \cdot \text{Relationstreffer}}{\max(\text{Relationstreffer})} - \frac{p_{rel} \cdot \text{Relationsfehler}}{\max(\text{Relationsfehler})} \quad (4)$$

³Ein gleichsinniger Zusammenhang im Beispiel ÖKOSYSTEM ist: „Je mehr Dünger, desto mehr Blätter. Je weniger Dünger, desto weniger Blätter.“ Ein gegensinniger Zusammenhang ist: „Je mehr Gift, desto geringer die Wasserverschmutzung. Je weniger Gift, desto höher die Wasserverschmutzung.“

$$\text{GdK}_{\text{vor}} = \frac{(1 - p_{\text{vor}}) \cdot \text{Vorzeichentreffer}}{\max(\text{Vorzeichentreffer})} - \frac{p_{\text{vor}} \cdot \text{Vorzeichenfehler}}{\max(\text{Vorzeichenfehler})} \quad (5)$$

$$\text{GdK}_{\text{num}} = \frac{(1 - p_{\text{num}}) \cdot \text{Numeriktreffer}}{\max(\text{Numeriktreffer})} - \frac{p_{\text{rel}} \cdot \text{Numerikfehler}}{\max(\text{Numerikfehler})} \quad (6)$$

$$\text{GdK}_{\text{sum}} = g_1 \cdot \text{GdK}_{\text{rel}} + g_2 \cdot \text{GdK}_{\text{vor}} + g_3 \cdot \text{GdK}_{\text{num}} \quad (7)$$

wobei $g_1, g_2, g_3 = \text{Gewichte}$

Ein Treffer in den obigen Formeln bezeichnet ein richtig erkanntes Element (Relation, Vorzeichen oder numerisches Gewicht) im Strukturdiagramm. Ein Fehler entspricht einem falschen Alarm nach der signalentdeckungstheoretischen Auffassung, also einem Element, das eine Versuchsperson im subjektiven Kausaldiagramm eingezeichnet hat, ohne dass es mit dem objektiven Diagramm übereinstimmt. Auf diese Weise wird vermieden, dass GdK mit einer Ja-Sage-Antworttendenz konfundiert ist. Auch die Ratewahrscheinlichkeiten ($p_{\text{rel}}, p_{\text{vor}}, p_{\text{num}}$) für die jeweiligen Elemente werden berücksichtigt. Funke (1992b) schlägt vor, diese für Relations- und Vorzeichenwissen im Standardfall auf .5 festzusetzen, für Numerikwissen auf .0, da eine Person das tatsächliche Gewicht aus einer theoretisch unendlich großen Menge möglicher Antworten erraten müsste. Treffer- und Fehlerzahlen werden an den maximal möglichen Anzahlen relativiert, so dass GdK auch von Systemeigenschaften (Anzahl der Variablen, Grad der Vernetztheit) weitgehend unabhängig ist (vgl. Müller, 1993).

Die Gewichtung, die in der Summenformel vorgenommen wird, hängt vom verwendeten System und den Forschungszielen ab. Eine – oft sinnvolle – Gleichgewichtung aller drei Komponenten ergibt sich z. B. durch $g_1 = g_2 = g_3 = 1$.

Trotz seiner Gebräuchlichkeit und offensichtlicher Vorzüge weist die Kausaldiagramm-Analyse auch einige Schwächen auf. Wie Kluwe (1988, S. 370) anführt, ist es keineswegs selbstverständlich, dass Probanden auf Anhieb begreifen, wie ein Kausaldiagramm formal richtig zu zeichnen ist (unabhängig von inhaltlichen Schwierigkeiten, korrekte Zusammenhänge anzugeben). Kluwe führt das Verständnisproblem nicht allein auf den intellektuellen Anspruch der Aufgabe zurück, sondern vermutet, dass spontan aufgebaute Wissensrepräsentationen ein anderes, vom Kausaldiagramm abweichendes Format haben – in direktem Widerspruch zu Funke (1985), der eine Übereinstimmung von interner Repräsentation (subjektivem Kausalmodell) und Wissensdiagnose (objektivem

Kausalmodell) annimmt.

Kluwe (1988) wendet ein, dass „Fragen nach bestimmten Informationen den Versuchspersonen nahelegen könnten, bestimmte Komponenten der Aufgabenstellung oder Sachverhalte besonders zu beachten“ (S. 370/371). Kausaldiagramme parallel zur Bearbeitung eines DYNAMIS-Szenarios würden demnach die Aufmerksamkeit in erhöhtem Maß auf kausale Abhängigkeiten und auf die Vernetztheit des Systems lenken. Solche Reaktivitäts- oder Demand-Effekte können nach Kluwe z. B. erklären, warum Versuchspersonen in der ÖKOSYSTEM-Studie von Funke (1985) die lineare, im Kausaldiagramm erfragte Systemstruktur relativ gut identifizierten, eine zeitverzögerte, im Diagramm nicht enthaltene Komponente hingegen kaum. Zu vermuten ist, dass Reaktivitätseffekte umso stärker ins Gewicht fallen, je häufiger Probanden Kausaldiagramme bearbeiten.

Zwar betont Beckmann (1994, S. 73), dass „prozessbegleitende Wissensdiagnostik generell reaktiv“ sei, dennoch gibt es neben der Kausaldiagramm-Analyse Bemühungen um möglicherweise weniger reaktive Alternativen: die *Paar-Aufgabe*, bei der Probanden nacheinander für jeweils zwei Variablennamen entscheiden, ob zwischen den Variablen eine Relation vorliegt oder nicht (Preußler, 1996, 1997, 1998), und die *lexikalische Entscheidungsaufgabe* (Preußler, 1996). Letztere basiert auf den Prinzipien des assoziativen Primings (vgl. Goshen-Gottstein & Moscovitch, 1995). Den Probanden werden am Bildschirm sequentiell Variablennamen dargeboten, entweder in korrekter Schreibung (Targets) oder als ähnlich lautende Distraktornamen. Die lexikalische Entscheidungsaufgabe besteht darin, so rasch wie möglich anzugeben, ob ein präsentiertes Wort ein Target oder ein Distraktor ist. Dabei stellen Reaktionszeiten das wesentliche abhängige Maß dar: Reagiert eine Person wiederholt überdurchschnittlich schnell auf eine bestimmte Abfolge zweier Variablennamen, so wird ihr eine interne, assoziative Verknüpfung zwischen beiden Variablen unterstellt. Die Assoziation wiederum kann wie eine erkannte Relation in der Paar-Aufgabe oder im Kausaldiagramm aufgefasst werden. Ob sich Reaktionszeiten und z. T. oberflächliche Assoziationen mit reflektiertem, artikulierbarem Strukturwissen gleichsetzen lassen, bleibt freilich diskutabel (für Kritik an Preußlers Ansatz s. z. B. Funke, 2003, S. 166).

Beiden von Preußler vorgeschlagenen Maßen gemein ist, dass jeweils Relationen zwi-

schen Systemvariablen, nicht aber Richtung und Numerik diagnostiziert werden. Zumindest in dieser Hinsicht stellen die Maße deshalb keine vollständige Alternative zur Kausaldiagramm-Analyse dar.

Eingriffswissen, Konstruktions- und Rekognitionsmaße. Nicht alles Wissen, das eine Person über ein komplexes, dynamisches System erwirbt, muss in einem mentalen Modell durch abstrakte Variablenzusammenhänge repräsentiert sein. Vertreter der *Instance Theory* (Logan, 1988) gehen sogar davon aus, dass primär nur einzelne, konkrete Lernepisoden im Gedächtnis gespeichert werden. Aus diesen kann eine Person abstraktes Wissen bilden, sofern die Lernbedingungen es zu- oder auch veranlassen (vgl. Dienes & Fahey, 1995). Ein umfangreiches, dynamisches System mit vielen Variablen kann z. B. aus speicherökonomischen Gründen zur Bildung abstrakter Modelle herausfordern, während für kleine Systeme anwendungsnäheres Input-Output-Wissen ausreicht: Es mag hier genügen, sich spezifische Eingabewerte und die zu erwartenden Folgewerte zu merken (s. Schoppek, 2002). Auch die Salienz der Systemstruktur ist von Bedeutung: Systemzusammenhänge, deren abstraktes Prinzip leicht ersichtlich ist, sollten mit höherer Wahrscheinlichkeit auch beim Probanden als abstraktes Strukturwissen repräsentiert sein (s. Dienes & Fahey, 1995).⁴

Diagnose-Verfahren für spezifisches Eingriffswissen lassen sich aus der Wissensdiagnostik bei finiten Automaten herleiten. Konkrete Lernepisoden werden hier isoliert betrachtet als ein Tripel von einem Ausgangszustand, einer Intervention und einem interventionsabhängigem Folgezustand. Nach Funke und Buchner (1992) bieten sich auf dieser Grundlage zwei Möglichkeiten, das spezifische Wissen abzufragen.

Konstruktionsmaße. Im Fall von Konstruktionsmaßen legt man Versuchspersonen lediglich zwei Elemente aus dem Tripel vor und bittet sie, das dritte fehlende Element frei zu nennen. Eine prognostische Frage kann etwa lauten: Gegeben Ausgangszustand und

⁴Eine empirische Abgrenzung dieser Faktoren erscheint mir problematisch insofern, als die Salienz einer Systemstruktur meist mit der Systemgröße konfundiert ist. Systembeziehungen in kleinen Systemen sollten üblicherweise leichter zu erkennen sein als in großen Systemen, so dass abstraktes Wissen über kleine Systeme nach Schoppek (2002) zwar nicht notwendig wäre, nach Dienes und Fahey (1995) aber mit großer Wahrscheinlichkeit trotzdem vorliegt.

Intervention, welcher Folgezustand ist zu erwarten? Mit einer solchen Prädiktionsaufgabe diagnostizierten Funke und Müller (1988) bereits erfolgreich Systemwissen in einem DYNAMIS-Szenario. Das von ihnen verwendete Maß *Güte der Vorhersage* ermittelt sich aus dem Abstand der vorhergesagten Werte von den Werten, die im tatsächlichen Fall eintreten würden.

Rekognitionsmaße. Bei Rekognitionsitems haben Probanden die Aufgabe, für ein vollständiges Tripel zu entscheiden, ob es sich bei der Lernepisode um eine bekannte, „alte“ Lernsituation handelt oder um eine „neue“ Lernsituation, die den Teilnehmern unbekannt sein sollte, weil sie vom Experimentator als Distraktor konstruiert wurde. Ein Wissensindex berechnet sich aus den Anzahlen richtiger und falscher Antworten. Diese Methode, angewendet auf die DYNAMIS-Variante LINAS, findet sich bei Preußler (2001).

Im Unterschied zu Maßen für Strukturwissen kann schon eine oberflächliche Beschäftigung mit dem komplexen, dynamischen System ausreichen, um in einer Rekognitionsaufgabe recht gute Ergebnisse zu erzielen. Erfolgreiches Anfertigen von Kausaldiagrammen hingegen sollte in jedem Fall tiefere Verarbeitung erfordern. Die spezifischen Input-Output-Maße eignen sich daher meiner Meinung nach kaum als Ersatz, sondern als Ergänzung zu Kausaldiagrammen, wenn man den Aufbau abstrakter, mentaler Repräsentationen untersuchen will.

2.3.2 Wissensanwendung

Die zweite Anforderung, mit der Problemlöser im Umgang mit DYNAMIS-Systemen konfrontiert werden, ist die Steuerung des Systems. Definierte Zielzustände mehrerer endogener Variablen sollen durch Manipulation der exogenen Variablen erreicht und aufrecht erhalten werden. Da man erfolgreiche Eingriffe weitgehend auf das erworbene Systemwissen zurückführt, spricht man bei Steuerleistung oft synonym von Wissensanwendung. Genauer wird in Abschnitt 3.1 auf das Verhältnis von Wissenserwerb und Wissensanwendung eingegangen werden.

Auch um die Steuerleistung zu erfassen, gibt es – nach problembehafteten Vorläufern

(s. Funke, 1985) – ein gut erprobtes und validiertes Maß. *Güte der Systemsteuerung* (GdS) nach Müller (1993, S. 61) berechnet sich aus logarithmierten, durchschnittlichen Abweichungen der real erreichten Ist-Werte von den vorgegebenen Ziel- oder Sollwerten:

$$\text{GdS} = \frac{\sum_{i=1}^{n_y} \sum_{j=1}^{n_t} \ln|y_{i,j} - z_i|}{n_y \cdot n_t} \quad (8)$$

wobei n_y = Anzahl zu steuernder endogener Variablen, n_t = Anzahl zu steuernder Takte, z_i = Zielwert der endogenen Variable i , $y_{i,j}$ = Wert der Variable i in Takt j .

Abweichungen vom Zielwert werden nicht nur gesondert für endogene Variablen, sondern zunächst auch für jeden Takt separat ermittelt. Die logarithmische Transformation der Abweichungen hat sich als sinnvoll herausgestellt, v. a. wenn Ausreißerwerte auftreten. Solche hohen und wenig reliablen Abweichungswerte fallen nach dem Logarithmieren schwächer ins Gewicht als kleine, meist reliablere Abweichungen. Ebenso wie für ein einfaches Distanzmaß gilt für GdS: Je niedriger GdS, desto höher die Steuerleistung, je höher GdS, desto geringer die Steuerleistung.

2.3.3 Schwächen und Risiken des DYNAMIS-Ansatzes

In den bisherigen Ausführungen wurden die Vorteile formal begründeter Szenarios betont. Das schließt jedoch Kritik nicht aus. Ob DYNAMIS tatsächlich eine gelungene Verbindung aus realitätsnaher Komplexität und Dynamik einerseits und formaler Beschreibbarkeit andererseits darstellt, haben Kritiker angezweifelt. Insbesondere die ökologische Validität ausschließlich linearer Zusammenhänge erscheint ihnen fragwürdig (s. Buchner, 1995, S. 48). Eine weitere potenzielle Gefahr des Ansatzes liegt darin, den Einfluss formaler Systemeigenschaften auf das komplexe Problemlösen zu überbetonen, während man Personenfaktoren (kognitive, motivationale, emotionale Zustände und Dispositionen) oder Eigenschaften der Problemlösesituation kaum berücksichtigt (vgl. Buchner, 1995). Um dieser Gefahr vorzubeugen, orientiert sich das nächste Kapitel an den von Frensch und Funke (1995) vorgeschlagenen Prozesskomponenten, einer Taxonomie, die neben Systemeigenschaften auch Personen- und Situationseinflüsse nicht aus den Augen zu verlieren versucht.

3 Determinanten im Problemlöseprozess

Ziel des folgenden Abschnitts ist es, einen Überblick über Faktoren zu geben, die sich in der empirischen Forschung als bedeutsam für den Problemlöseprozess herausgestellt haben. Die Vorstellung der Einflussfaktoren richtet sich schwerpunktmäßig auf die Fragen: Welche Variablen fördern oder hemmen den Wissenserwerb im Umgang mit komplexen, dynamischen Systemen? Welche Variablen beeinflussen die Steuerleistung, und in welcher Weise erfolgt die Einflussnahme? Um beide Fragen sinnvoll miteinander zu verbinden, wird auch das Verhältnis zwischen Wissenserwerb bzw. Wissensrepräsentation und Steuerung diskutiert werden. Natürlich kann nur ein kleiner Ausschnitt empirischer Ergebnisse dargelegt und ein noch kleinerer Ausschnitt eingehend erörtert werden. Ausschlaggebend für die Auswahl sind die ÖKOSYSTEM-Studie von Funke (1985), die in dieser Arbeit teilweise repliziert wird, wie auch Überlegungen zum Einfluss der Kausaldiagramm-Analyse beim komplexen Problemlösen. Aus dieser Übersicht werden die inhaltlichen Hypothesen der eigenen Untersuchung hergeleitet und am Kapitelende zusammengefasst.

3.1 Wissenserwerb und steuerndes Handeln

Eigenschaften des Systems, situationale Rahmenbedingungen und Eigenschaften des Problemlösers können sowohl Einfluss auf die Güte des Wissenserwerbs als auch auf die Güte der Steuerung nehmen. So wird eine intelligente Person der allgemeinen Auffassung nach viel Wissen über ein dynamisches System erwerben und zugleich in der Lage sein, Steuerziele angemessen zu bewältigen. Lässt sich das Personenmerkmal Intelligenz dabei als die unmittelbare Ursache des Steuererfolgs interpretieren? Oder unterstützt die intellektuelle Befähigung primär den Wissenserwerb, der seinerseits die erfolgreiche Systemsteuerung bedingt? Zunächst ohne äußere Determinanten wie Intelligenz zu betrachten geht man üblicherweise davon aus, dass der Steuererfolg durch den Wissenserwerb vermittelt wird. Dies entspricht auch dem Ablauf eines typischen Problemlöseexperiments, in dem ein Proband während anfänglicher Explorationsphasen ausschließlich Systemwissen erwerben soll, während die Steuerung einer abschließenden Anwendungsphase vorbehalten ist.

In der Empirie finden sich nicht durchweg stützende Befunde. In der erwähnten ÖKO-SYSTEM-Studie von Funke (1985) beschreiben a posteriori ermittelte Korrelationskoeffizienten von ca. $r_s = -.60$ die Zusammenhänge zwischen Systemwissen und Steuergüte. Während in Funkes Untersuchung eine inhaltliche Interpretation unterbleibt, deuten andere Autoren negative statistische Zusammenhänge als Dissoziation der Prozesse Wissenserwerb und Steuerung: Auf Berry und Broadbent (1984) geht die Interpretation zurück, dass explizites, verbal erfassbares Systemwissen für die praktische Steueraufgabe möglicherweise irrelevant ist. Der Steuerprozess würde, wenn überhaupt, von anwendungsnäherem, implizitem, non-verbalem Wissen abhängen. Ähnlich argumentiert Putz-Osterloh (1993a, 1993b), die im experimentellen Vergleich zweier Problemlösesituationen im Mittel gleiche Steuerleistungen bei unterschiedlichem Wissensniveau feststellte.

Eine große Anzahl weiterer Studien hingegen befürwortet die Auffassung, Systemsteuerung sei die Anwendung von erworbenem, explizitem Strukturwissen. Eine signifikant positive a posteriori Korrelation zwischen Systemidentifikation und -steuerung fanden bereits Funke und Müller (1988; s. auch Funke, 1992b) in einer Untersuchung mit dem vorwissensneutralen DYNAMIS-System SINUS. Ebenfalls in einer Studie mit SINUS, im Rahmen eines umfangreicheren Pfadmodells, legten Beckmann und Guthke (1995) positive Zusammenhänge vor, ähnlich wie Vollmeyer und Rheinberg (1998) mit dem semantisch eingekleideten System BIOLOGYLAB und Kröner (2001) mit dem abstrakten Szenario MULTIFLUX. Die Korrelationskoeffizienten reichen von $r = .51$ bei Beckmann und Guthke (1995) bis $r = .65$ bei Kröner (2001). Das vielleicht eindrucksvollste Beispiel für die Abhängigkeit der Steuerleistung vom Systemwissen liefert Müllers (1993) LISREL-Analyse entsprechender latenter Konstruktvariablen: korrelative Zusammenhänge von $r = .83$ zu einem ersten, $r = .86$ zu einem zweiten Messzeitpunkt. Nach Müller (1993, S. 208) ist die Steuerleistung sogar vollständig auf das erworbene, abstrakt in Kausalbeziehungen darstellbare Strukturwissen zurückzuführen.

Preußler (1996), die Effekte von experimentell unterstütztem Wissenserwerb untersuchte, fand signifikant positive Zusammenhänge zwischen Wissen und steuerndem Handeln nur dann, wenn das Strukturwissen mit konkreten Beispielen anwendungsnah⁵ und zu-

⁵In einer Folgestudie (Preußler, 1998) stellt die Autorin heraus, dass es sich bei anwendungsnahem Wissen keineswegs um spezifisches Eingriffswissen handelt. Abstraktes Strukturwissen, sofern es im

dem früh genug vermittelt wurde, um nicht mit den kognitiven Anforderungen der Steueraufgabe zu interferieren. Offenbar kann Strukturwissen erst zur Anwendung kommen und Einfluss auf die Steuerleistung nehmen, sobald ein ausreichender Wissenserwerb abgeschlossen ist.

Der Zeitpunkt des Wissenserwerbs und die Anwendungsnahe stellen demnach Grundvoraussetzungen für einen effektiven Wissenserwerb und effektive Wissensanwendung dar. Weitere vermittelnde Faktoren werden in den nächsten drei Abschnitten angesprochen. Die Ausgangshypothese für die eigene empirische Untersuchung bildet zunächst die Annahme, dass Steuern eines komplexen Systems in erster Linie Wissensanwendung ist und daher positiv vom erworbenen Strukturwissen abhängt.

3.2 Systemfaktoren

Bei den Systemfaktoren handelt es sich um formale Systemeigenschaften, die – unabhängig von Semantik oder Darbietungsformat – Dynamik und Komplexität einer Simulation bestimmen. Für Szenarios vom DYNAMIS-Typ werden diese Faktoren vollständig durch die *AR*-Prozesse definiert (s. S. 14). Zu ihnen gehören die Anzahl der Variablen, die Anzahl der Relationen sowie einige spezielle Arten von Relationen: Eigendynamiken, Nebenwirkungen, parallele und zeitverzögerte Effekte (für experimentelle Befunde s. Funke, 1992b). Untersuchungen zum Einfluss formaler Systemeigenschaften orientieren sich an der Rahmentheorie, dass die technische Komplexität eines Systems die subjektiv empfundene bzw. psychologische Komplexität widerspiegelt. Formal aufwändige Systeme bedeuten für Problemlöser normalerweise einen erhöhten kognitiven Aufwand, erschwerten Wissenserwerb und eingeschränkte Wissensanwendung.

Aus der Vielzahl komplexitätsstiftender Systemeigenschaften variierte Funke (1985) im Szenario ÖKOSYSTEM zwei Faktoren, den Grad der zeitlichen Abhängigkeit und den Grad der Systemvernetztheit. Im 2×3 -Zwischengruppen-Design untersuchte er die getrennten und kombinierten Auswirkungen beider Faktoren auf das Problemlöseverhalten von 60 Probanden. Für den zweifach gestuften Faktor zeitliche Abhängigkeit wurden Systeme aus reinen *AR*-1-Prozessen (Bedingung ohne zeitverzögerte Rückmeldung) von

Anwendungskontext erläutert wird, sei effektiver als nicht übertragbares, reines Eingriffswissen.

Systemen abgegrenzt, in denen einige *AR*-1-Prozesse durch korrespondierende *AR*-2-Prozesse mit denselben numerischen Gewichten ersetzt wurden (Bedingung mit zeitverzögerter Rückmeldung). Vermutet wurde, dass ÖKOSYSTEM-Varianten mit zeitverzögerten Effekten schwieriger zu identifizieren und schwieriger zu steuern sind.

Der Faktor Vernetztheit sah drei Abstufungen vor, gekennzeichnet durch die Anzahl der Systemrelationen: Das Diagramm auf S. 16 entspricht dem mittleren Vernetztheitsgrad mit sechs Teilprozessen oder Relationen. Ein System mit geringer Vernetztheit bestand aus nur vier Teilprozessen,⁶ während hohe Vernetztheit über acht Teilprozesse operationalisiert wurde.⁷ Funke prüfte so die Annahme, dass die subjektive Komplexität eines Szenarios nicht allein von der (hier konstant gehaltenen) Variablenzahl oder Größe des Systems abhängt, sondern auch von der Anzahl der Relationen im System. Der Erwartung nach sollten die Güte der Systemidentifikation und die Steuerleistung mit wachsendem Vernetztheitsgrad monoton abnehmen.

Die Ergebnisse entsprechen weitgehend den Hypothesen. Ein leistungsmindernder Effekt der zeitverzögerten Rückmeldung auf den Steuererfolg konnte allerdings formal nicht nachgewiesen werden. Der erwartete monotone Abfall im Wissenserwerb zeigte sich zudem nur bei unverzögerter Rückmeldung. Es bleibt sicherlich unbestritten, dass Zeiteffekte wesentlich zur Komplexität und Dynamik eines Problems beitragen und von Probanden dennoch oft übersehen werden (vgl. Dörner, 1981); zumindest für Szenarios nach dem DYNAMIS-Ansatz hat sich der Einfluss zeitverzögerter Rückmeldungen aber als wenig systematisch herausgestellt.

In relativ systematischer Weise scheint hohe Vernetztheit mit ungenauem, unvollständigem Wissen und mit eingeschränkter Steuerleistung einherzugehen. In einer jüngeren, umfangreichen Untersuchung konnte Kluge (2003) Vernetztheit ebenfalls als validen Indikator für die Komplexität bzw. Schwierigkeit eines Simulationsproblems einsetzen. Kluge, die u. a. Wissens- und Anwendungsgüte für drei Vernetztheitsvarianten des DYNAMIS-Szenarios COLORSIM verglich, bestätigte Funkes (1985) Befunde. Darüber hinaus fand sie einen moderierenden Einfluss des Vernetztheitsgrades auf das Verhältnis von Wissen und Steuern. Gütemaße beider Problemanforderungen korrelierten durchweg po-

⁶fehlend: Nebenwirkung Käfer – Blätter, direkter Effekt Dünger – Wasserverschmutzung

⁷zusätzlich: Nebenwirkung Blätter – Käfer, direkter Effekt Gift – Blätter

sitiv, für das am stärksten vernetzte Szenario ($r = .44$) jedoch beträchtlich niedriger als bei mittlerer oder schwacher Vernetztheit ($r = .68$ bzw. $r = .70$). Nach Kluge könnte der schwächere statistische Zusammenhang bei größerer Vernetztheit auf einen erhöhten Einfluss von unsystematischen Situationsfaktoren, z.B. auf Unterschiede in der Bearbeitungszeit zurückgehen.

Für die geplante Untersuchung soll der Effekt von Vernetztheit auf Systemidentifikation und -steuerung geprüft und nach Möglichkeit repliziert werden. Zum anderen interessiert, ob der Zusammenhang zwischen Wissen und Steuergüte wie bei Kluge (2003) von der Systemvernetztheit abhängt. Auch für die noch folgenden Hypothesen wird Vernetztheit als vermittelnder Faktor in Betracht kommen.

3.3 Situationale Faktoren

Unter situationalen Faktoren werden in dieser Arbeit verschiedene Einflussfaktoren zusammengefasst, die zur Problemstellung im weiteren Sinn gehören. Darbietungsmerkmale wie Layout oder Semantik sind an die Computersimulation gebunden, jedoch in oberflächlicherer Weise als die formalen Systemmerkmale. Erweiterte Aufgabenanforderungen beziehen sich auf Zusatzinformationen, Erläuterungen oder verbale Instruktionen zum Umgang mit dem komplexen System.

3.3.1 Art der Informationsdarbietung

Drei Facetten der Informationsdarbietung erscheinen für die Einordnung des Szenarios ÖKOSYSTEM relevant: Grafische gegenüber numerischer Informationsdarbietung, simultane gegenüber sequenzieller Darstellung und semantische Einkleidung gegenüber Abstraktheit.

Grafische gegenüber numerischer Informationsdarbietung. Zustände eines komplexen Systems können einem Problemlöser sowohl numerisch als Zahlen in Tabellen als auch grafisch mitgeteilt werden. Verstellbare Balken, deren Größe einen Inputwert symbolisiert, oder Liniendiagramme, aus deren Höhe der Verlauf von Outputwerten ersichtlich wird, sind Beispiele für grafische Informationsdarbietung. Auch Kombinationen

beider Präsentationsformate sind möglich, wenn auch die rein numerische Darstellung meist die Regel ist (vgl. Beckmann, 1994, S. 65).

Auf Hübner (1988) gehen experimentelle Befunde zurück, nach denen sich die grafische, analoge Darbietung förderlich auf die Steuerleistung auswirkt. Eine DYNAMIS-Studie von Funke (1992a) hingegen brachte ein abweichendes Bild: Funke, der Problemlöseleistungen (Systemsteuerung und -identifikation) für formal und semantisch äquivalente Systeme bei unterschiedlichem Präsentationsformat verglich, konnte keine generelle Überlegenheit eines Darbietungsmodus feststellen. Vielmehr fand sich eine Interaktion zwischen dem Darbietungsmodus und der Schwierigkeit der Problemstellung. Bei einer einfachen Steueraufgabe ließ sich das System im numerischen, bei einer komplexeren Aufgabe im grafischen Format vollständiger identifizieren. Nach Funke ist das Darbietungsformat daher für sich betrachtet kein ausschlaggebender Faktor im Problemlöseprozess, sondern immer auf die Art der Problemstellung zu beziehen. Das Format ist für Probanden jedoch keineswegs gleichgültig, wie ein Nebenbefund in Funkes Untersuchung (s. Funke, 1992b) zeigt: Schon vor dem Beginn der Szenariobearbeitung berichteten Versuchspersonen, denen man ein Numerik-Szenario vorgab, höhere Angst zu versagen (Misserfolgsmotiviertheit) und geringere Anstrengungsbereitschaft als jene, deren Szenario eine grafische Oberfläche hatte.

Simultane gegenüber sequenzieller Darstellung. Jede Simulation erfordert eine Festlegung, welche Systemzustände zu welchem Zeitpunkt des dynamischen Verlaufs sichtbar sind. Im sequenziellen Fall ist nur der jeweils aktuelle Zustand einsehbar. Er wird im Folgetakt durch den nächsten ersetzt. Simultane Darbietung stellt neben dem aktuellen Zustand auch vergangene Systemzustände dar. Die Oberfläche von ÖKOSYSTEM beispielsweise zeigt die Werte exogener und endogener Variablen für die maximal letzten fünf Takte (von insgesamt sieben Takten) zugleich an.

In der DYNAMIS-Forschung findet sich überwiegend der simultane Modus – nach Müller (1993) aus gutem Grund. Nur bei simultaner Darstellung, argumentiert Müller, können Probanden die zeitliche Abfolge der Systemzustände ohne übermäßigen Gedächtnisaufwand analysieren. Die Voraussetzungen für den Wissenserwerb und darauf aufbauende Steuerprozesse sind günstig. Im sequenziellen Modus hingegen fiel der Wissenserwerb

in Müllers Untersuchung deutlich geringer aus, und substantielle Zusammenhänge zwischen Wissen und Steuern waren nicht nachweisbar.

Semantische Einkleidung gegenüber Abstraktheit. Computerbasierte Simulationen lassen sich auch danach charakterisieren, ob ihre Variablen aus einem inhaltlich bedeutsamen, an Vorwissen gebundenem Zusammenhang stammen (wie bei ÖKOSYSTEM) oder ob sie rein abstrakt bzw. vorwissensneutral sind. Ein abstraktes System kann als Variablenbezeichnungen reine Fantasienamen tragen (s. LINAS, z. B. bei Putz-Osterloh, 1993a, 1993b, Preußler, 2001) oder einzelne Buchstaben und Ziffern, wie es bei den simulierten fiktiven Maschinen MASCHINE (Beckmann, 1994) oder MULTIFLUX (Kröner, 2001) der Fall ist. Lösen Probanden eine semantisch eingekleidete Simulation anders als eine abstrakte Simulation und stimmen die verwendeten Systeme in allen übrigen Merkmalen überein, führt man die Lösungsunterschiede auf Vorwissenseffekte zurück.⁸

Um nach diesem Prinzip Vorwissens- und Semantikeffekte zu untersuchen, hat Beckmann (1994) das semantisch eingekleidete DYNAMIS-Szenario KIRSCHBAUM im Vergleich mit seinem semantikfreien Äquivalent MASCHINE eingesetzt. Probanden, die mit der Semantikversion arbeiteten, zeigten geringeren Wissenserwerb als die Bearbeiter der abstrakten MASCHINE. Fiel der Wissenserwerb hier deshalb niedriger aus, weil die im System umgesetzten semantischen Relationen den Vorerwartungen der Problemlöser widersprachen? Während eine Untersuchung von Funke (1992a, 1992b) eingeschränkten Wissenserwerb bei vorwissensinkompatiblen Systemen belegt, geht Beckmann davon aus, dass die Variablenbeziehungen im KIRSCHBAUM-Szenario mit dem Weltwissen der Probanden durchaus vereinbar waren. Unabhängig von kompatibelem oder inkompatiblen Vorwissen kann ein semantischer Kontext die Effizienz des Wissenserwerbs auf subtilere Weise beeinträchtigen: Jedes System mit inhaltlich bedeutsamen Variablennamen

⁸Dieser experimentelle Ansatz entspricht dem erwähnten allgemeinpsychologischen Vorgehen. Nicht interindividuelle Unterschiede zwischen Experten und vorwissensfreien Novizen werden geprüft; vielmehr ist die Rolle des Vorwissens in der Aufgabenstellung begründet. Semantisch eingekleidete Szenarios sprechen seltener spezielles Expertenwissen an als allgemeines Weltwissen, das ein Forscher bei allen Probanden voraussetzt.

ruft nach Beckmann die Erwartung hervor, „alles sei mit allem vernetzt“. Die ursprüngliche mentale Repräsentation entspricht einem Modell allgemeiner Intervernetzttheit. Dieses Modell wird ein Proband mit zunehmender Systemerfahrung schrittweise reduzieren, indem er nicht vorhandene Verknüpfungen zwischen Systemvariablen entfernt. Im Fall eines vorwissensneutralen Systems hingegen enthält die ursprüngliche Repräsentation nach Beckmann nur die Variablen ohne jegliche Verknüpfungen. Relationen werden mit zunehmender Erfahrung nicht entfernt, sondern zum mentalen Modell hinzugefügt. Da der vermutete reduktive Prozess bei semantisch eingekleideten Systemen offensichtlich kognitiv aufwändiger ist als der konstruktive Prozess des Wissenserwerbs bei abstrakten Systemen, sollte das erworbene Wissen in Semantik-Szenarios bei begrenzter Explorationsdauer unpräziser sein.

Insgesamt, obwohl ein vergleichbar einschränkender Semantikeffekt auf die Steuerleistung nur tendenziell zu finden war, geht Beckmann von höherer Problemlösequalität bei abstrakten Szenarios aus. Es sei „offenbar leichter, ein System zu verstehen und zu steuern, wenn ‚man weiß, dass man nichts weiß‘ (MASCHINE) als wenn ‚man glaubt, dass man etwas weiß‘ (KIRSCHBAUM)“ (Beckmann, 1994, S. 199).

Für den Einsatz von ÖKOSYSTEM als semantisch eingekleidetem Versuchsmaterial ergibt sich in dieser Hinsicht also möglicherweise keine optimale Voraussetzung, um den Wissenserwerb und sein Verhältnis zur Wissensanwendung zu untersuchen. Ein unbestrittener Vorteil dürfte in der simultanen Darstellung aktueller und vorheriger Systemzustände liegen. Auch das numerische Format sollte zumindest keinen negativen Effekt auf Wissen und Steuern haben.

3.3.2 Erweiterte Aufgabenanforderungen

Die Frage, welche zusätzlichen, situationalen Einflussgrößen den Problemlöseprozess fördern oder hemmen, ist in etlichen Studien thematisiert worden. Das typische Untersuchungsparadigma sieht einen Vergleich von Experimentalprobanden in einer angereicherten Problemlöseumgebung mit nicht unterstützten Kontrollprobanden vor. Die erweiterten Aufgabenanforderungen einer angereicherten Problemsituation können sowohl *inhaltsorientiert* systemspezifisches Wissen vermitteln als auch *strategieorientiert*

wirken, indem sie relativ unabhängig vom verwendeten System zu strategisch-systematischem Problemlösen anregen. V. a. die strategieorientierten Maßnahmen sollen hier von Interesse sein, da sie die Frage aufwerfen, ob möglicherweise auch die Methode der prozessbegleitenden Kausaldiagramm-Analyse ähnliche Auswirkungen auf den Problemlöseprozess hat.

Inhaltsorientierte Maßnahmen. In einer Standard-DYNAMIS-Situation erwerben Versuchspersonen Wissen, indem sie ein System frei explorieren. Diesen Prozess kann der Versuchsleiter unterstützen, indem er ergänzende, objektive Informationen über das System verfügbar macht. Eine einfache Operationalisierung sind Transparenzbedingungen, in denen der Proband das Strukturdiagramm des verwendeten Systems einsehen kann (Putz-Osterloh, 1993a, 1993b). Transparenz lässt sich erweitern durch verbale Erläuterungen oder konkrete Beispiele (Preußler, 1996) oder durch Übungsaufgaben im Analysieren von Output-Effekten (Preußler, 1998, Schoppek, 2002).

Problemlöser, die durch derartige Maßnahmen unterstützt wurden, zeigten in den zitierten Studien einheitlich einen gesteigerten Wissenserwerb und oftmals erhöhte Leistungen im Steuern der komplexen Systeme. Insbesondere in Transferaufgaben wurde die Überlegenheit instruierter Versuchspersonen gegenüber Kontrollpersonen offensichtlich, was man als Beleg für umfassendes Strukturwissen interpretiert.

Strategieorientierte Einflüsse. Man muss nicht notwendigerweise die Problemstellung durch inhaltliche Informationen anreichern, um den Wissenserwerb zu fördern. Ein anderer Ansatz besteht darin, den aktiven Prozess des Wissenserwerbs zu intensivieren, indem man möglichst günstige Bedingungen für strategisch effektive Operationen schafft. Strategien beim komplexen Problemlösen wiederum stehen in enger Verbindung zum Problemlöseerfolg unabhängig davon, ob Versuchspersonen sie spontan oder nach entsprechender Instruktion anwenden (s. auch Abschnitt 3.4.2). Die experimentellen Arbeiten von Vollmeyer und Kollegen (Vollmeyer, Burns & Holyoak, 1996; Vollmeyer & Burns, 1996) mit dem DYNAMIS-Szenario BIOLOGY LAB stellen drei Facetten strategieorientierter Einflüsse heraus: explizite Strategieinstruktion, Zielspezifität und Hypotheseninstruktion.

Eine *explizite Anleitung zum Strategiegebrauch* gaben Vollmeyer, Burns und Holyoak (1996), indem sie Problemlöser aufforderten, beim Explorieren des vorgegebenen Systems jeweils nur eine Input-Variable pro Takt zu verändern. Diese Strategie, abgekürzt VOTAT (= Vary One Thing At a Time), ist von vielen Autoren übereinstimmend als besonders erfolgversprechend beschrieben worden (s. z. B. Putz-Osterloh, 1993b, Vollmeyer & Rheinberg, 1998, Rollett, 2002). Auch bei Vollmeyer et al. führte sie zum erwarteten Effekt: Wissenserwerb und Steuergüte, insbesondere im Transfer, waren bei VOTAT-instruierten Versuchspersonen erhöht.

In derselben Studie wie auch in der Folgestudie (Vollmeyer & Burns, 1996) konnten die Autoren weiterhin zeigen, dass der Strategiegebrauch beim komplexen Problemlösen von der *Zielspezifität* der Problemstellung abhängt. Ein spezifisches Ziel benennt konkret zu erreichende Sollwerte, während die unspezifische Kontrastbedingung lediglich zielfreie Exploration verlangt. Im letzteren Fall sollen Problemlöser so viel wie möglich über Systemzusammenhänge herausfinden, ohne ihr Wissen zur Steuerung heranzuziehen.

Im Vergleich dieser beiden Bedingungen fanden Vollmeyer et al. (1996) das Strukturwissen ebenso wie die Transfer-Steuerleistung bei Zielspezifität erniedrigt. Auf Prozessebene diagnostizierten sie einen verminderten Gebrauch von VOTAT und anderen, für den Wissenserwerb förderlichen Strategien. Probanden der zielspezifischen Problemlösebedingung verfolgten eine mit VOTAT nicht kompatible Strategie der Unterschiedsreduktion bzw. eine Mittel-Zweck-Analyse: Sie versuchten ausschließlich, den Abstand zwischen aktuellem Ist-Zustand und angestrebtem Zielzustand von Takt zu Takt zu verringern. In Anlehnung an Sweller (1988) folgern die Autoren: Spezifische Zielvorgaben können den Erwerb von speziellem, lösungsbezogenem Wissen fördern; sie stellen jedoch eine ungünstige Voraussetzung dar, wenn Problemlöser breites, auch im Transfer anwendbares Strukturwissen erwerben sollen.

Eine dritte Möglichkeit, Intensität und Systematik des Explorationsverhaltens zu beeinflussen, bieten *Hypotheseninstruktionen*. Im Gegensatz zu expliziten Strategieinstruktionen beinhalten Hypotheseninstruktionen keine Vorgaben, durch welche konkreten Operationen eine Person ein System explorieren soll. Die Instruktion fungiert vielmehr als allgemeiner, orientierender Rahmen der Problemlösesituation: Den Versuchspersonen

wird eine Hypothese über eine spezifische Systembeziehung genannt, die im Fall von ÖKOSYSTEM z.B. lauten könnte: Ein Gärtner vermutet, dass Dünger die Blätterzahl beeinflusst, indem jede Gabe von Dünger mit zehn multipliziert zur Anzahl der Blätter addiert wird (für die Originalinstruktion s. Vollmeyer & Burns, 1996, S. 665). Inwieweit die Hypothese mit den tatsächlichen Systembeziehungen übereinstimmt, hat die Versuchsperson zu prüfen, während sie das System exploriert.

Kann eine solche Zusatzinstruktion den Problemlöseprozess fördern? Vollmeyer und Burns (1996) stützten ihre Annahmen diesbezüglich auf die *Dual Space Theory* nach Klahr und Dunbar (1988), eine Lerntheorie, die Lernprozesse über einen zusammengesetzten Problemraum aus zwei Komponenten definiert: Hypothesen- und Experimenterraum. Im Experimenterraum bewegt sich ein Problemlöser nach Klahr und Dunbar (1988) dann, wenn er ohne artikulierte Vorerwartungen Eingriffe im System vornimmt, Maßnahmen ausprobiert, aus denen sich später explizites Wissen ableiten lässt. Formuliert eine Person hingegen zunächst Vermutungen über Systemzusammenhänge, um diese anschließend in ausgewählten Experimenten zu prüfen, bearbeitet sie das Problem weitgehend im Hypothesenraum. Die vorwiegende Suche nach Zusammenhängen im Hypothesenraum entspricht dem offensichtlich systematischeren Zugang und sollte mit erfolgreicherem Problemlösen einhergehen.

Vollmeyer und Burns (1996) vermuteten, dass Hypotheseninstruktionen die systematische Suche im Hypothesenraum verstärken. Damit übereinstimmend stellten sie erhöhtes Strukturwissen und gesteigerte Steuerleistungen bei hypothesen-instruierten Versuchspersonen gegenüber Kontrollprobanden fest. Interessanterweise trat der Effekt selbst dann auf, wenn eine vorgegebene Hypothese falsche Informationen enthielt, aber nicht, wenn richtige Informationen in anderer Formulierung als im Hypothesenformat dargeboten wurden. Der gesteigerte Wissenserwerb beschränkte sich auch nicht auf die Systembeziehungen, die in der Hypotheseninstruktion ausdrücklich angesprochen wurden. Unabhängig vom informativen Gehalt scheinen die Hypotheseninstruktionen bei Vollmeyer und Burns (1996) hypothesentestendes Problemlösen begünstigt zu haben.

Kausaldiagramm-Analyse als Hypotheseninstruktion? In Abschnitt 2.3.1 wurden das Prinzip und der Anwendungsbereich von Kausaldiagramm-Analysen einschließ-

lich ihrer Kritik vorgestellt. Auch Vertreter dieser wissensdiagnostischen Methode haben eingeräumt, dass Kausaldiagramme Aufforderungscharakter haben, die erfragten Kausalzusammenhänge beim Bearbeiten eines komplexen Problemlösezenarios besonders zu beachten (s. Funke, 1992b, S. 78). Empirische Hinweise darauf, dass Wissensdiagnose reaktiv auf den komplexen Problemlöseprozess wirken kann, finden sich jedoch bislang nur für einen verbalen Wissenstest beim realitätsnahen Szenario SCHNEIDERWERKSTATT. Süß et al. (1993) stellten in einem Nebenbefund fest, dass Probanden ein System zunächst besser steuerten, wenn sie zuvor einen Wissenstest absolviert hatten, als wenn der Wissenstest entfiel. Die Autoren vermuten, der Test könne die Teilnehmer „zu Reflexion und Hypothesenbildung angeregt“ haben (Süß et al., 1993, S. 196).

Es liegt nahe, die Interpretation auf Kausaldiagramm-Analysen zu übertragen. Personen, die ein Kausaldiagramm bearbeiten, verbalisieren abstraktes Wissen, sofern sie es erworben haben. Nicht immer aber muss die Abstraktion schon während der aktiven Exploration spontan stattgefunden haben; sie kann auch erst in dem Moment erfolgen, in dem eine Person durch den wissensdiagnostischen Test zu formalen Angaben über Systemzusammenhänge aufgefordert wird. In diesem Fall bleiben Angaben im Kausaldiagramm zunächst hypothetisch, stellen jedoch zugleich explizite Festlegungen dar. Als solche sollten sie der Person im weiteren Problemlöseprozess ohne weiteres zugänglich sein. Die Vorläufigkeit der angenommenen Systembeziehungen wiederum kann Eingriffe verlangen, die im Sinne von Experimenten nach Klahr und Dunbar (1988) die Hypothesen prüfen sollen. Nach der Dual Space Theory wäre demnach für Bearbeiter von Kausaldiagrammen eine verstärkte Suche im Hypothesenraum zu erwarten.

Eine Parallele zur Hypotheseninstruktion bei Vollmeyer und Burns (1996) ist leicht zu ziehen. Der Unterschied, dass explizite Hypothesen bei Vollmeyer und Burns (1996) von außen vorgegeben bzw. fremdgeneriert waren, während Kausaldiagramme selbstgenerierte Hypothesen implizieren, fällt wenig ins Gewicht, wenn man beachtet, dass die Hypotheseninstruktion nach Vollmeyer und Burns nur als Auslöser für interne, subjektive Hypothesenbildung und Hypothesensuche interpretiert wird.

Für die eigene Untersuchung gehe ich deshalb von einer intensivierten Hypothesenbildung durch Kausaldiagramm-Analysen aus. Die Hypothesenbildung sollte den Wissen-

serwerb und darauf aufbauend die Steuerung des komplexen Systems unterstützen.

3.4 Personenfaktoren

Die dritte große Klasse von Einflüssen richtet ihren Blick weniger auf Eigenschaften von Problem oder Situation als auf Merkmale, die primär mit der problemlösenden Person in Verbindung gebracht werden: Intelligenz, spontaner Strategieeinsatz und motivationale bzw. emotionale Faktoren sollen als die wesentlichen Personenvariablen angesprochen werden.

3.4.1 Intelligenz

Inwieweit Intelligenz und komplexes Problemlösen verwandte Konstrukte darstellen, ist Gegenstand kontroverser Diskussionen geworden. Insbesondere Dörners (1986) Konzept der *operativen Intelligenz*, das wegen seiner Kritik an traditionellen Intelligenztests bekannt wurde, postuliert fundamentale Unterschiede zwischen einem herkömmlichen statisch-analytischen Intelligenzbegriff und einer dynamisch ausgerichteten, operativen (Problemlöse-)Intelligenz. Gestützt werden Dörners Annahmen u. a. durch frühe Studien von Putz-Osterloh (Putz-Osterloh, 1981, Putz-Osterloh & Lüer, 1981), die allenfalls schwache prädiktive Zusammenhänge zwischen akademischer Intelligenz und Problemlöseerfolg aufzeigen (für Kritik an den Untersuchungen: s. Funke, 1983). Auch Funke (1985), der in seiner ÖKOSYSTEM-Studie Korrelationsanalysen für Maße der Problemlösequalität und Skalen des Berliner-Intelligenzstrukturmodells (Jäger, 1982) durchführte, konnte positive Korrelationen nur in der Höhe von $r = .12$ bis $r = .37$ bei einer insgesamt geringen Varianzaufklärung nachweisen.

Neuere empirische Arbeiten belegen demgegenüber durchaus substantielle, positive Zusammenhänge. Beckmann und Guthke (1995) sprechen von Intelligenz als einem konsistenten Prädiktor für erfolgreiches Problemlösen, wenn sie auch einräumen, dass der Einfluss von Intelligenz durch situationale Variablen vermittelt sein kann und nicht immer gleichermaßen zum Tragen kommt: Systemtransparenz (s. dazu Putz-Osterloh, 1981), klar definierte Zielanforderungen und abstrakte, semantikfreie Systeme machen ein komplexes Problem einer analytischen Aufgabe im Intelligenztest eher vergleichbar

als intransparente, semantisch eingekleideten Systemen mit unklaren Zielvorgaben. Im ersteren Fall sollte die akademische Intelligenz daher ein besserer Prädiktor der Problemlöseleistung sein.

Die höchsten Zusammenhänge zwischen Wissenserwerb, Systemkontrolle und Intelligenz fanden sich dementsprechend in Untersuchungen mit MULTIFLUX, einem abstrakten DYNAMIS-System, das speziell als computerbasiertes, dynamisches Intelligenzdiagnostikum entwickelt und validiert wurde (Kröner, 2001). Ein globaler MULTIFLUX-Leistungswert korreliert bei Kröner zu $r = .65$ mit Maßen der Verarbeitungskapazität gemäß dem Berliner Intelligenz-Strukturmodell (Jäger, 1982). Intelligenz ist in Krönens Modell sowohl ein Prädiktor für geschickte Systemexploration, die den Wissenserwerb und Steuerprozess fördert, als auch ein zusätzlicher Prädiktor für gute Steuerleistung unabhängig von den Explorationsfertigkeiten.

3.4.2 Spontaner Strategieeinsatz

Wie oben erwähnt kann auch der Einsatz systematischer Problemlösestrategien erfolgreiches Identifizieren und Steuern dynamischer Systeme vorhersagen. Sowohl bei experimenteller Strategieinstruktion als auch bei spontanem Strategieeinsatz eignen sich VOTAT-Strategien bzw. Einfacheingriffe, um Effekte einzelner Variablen isoliert zu beobachten. Ähnlich geeignet sind sogenannte Nulleingriffe bzw. das Konstanthalten sämtlicher Input-Variablen, um Eigendynamiken festzustellen. Für VOTAT und das Variieren von nicht mehr als zwei Variablen pro Takt fand Rollett (2002) in einer Studie mit dem DYNAMIS-Szenario BIOLOGY LAB die höchsten korrelativen Zusammenhänge zum erworbenen Strukturwissen ($r = .51$), für Nulleingriffe zeigten sich ebenfalls hohe Zusammenhänge ($r = .35$). Allenfalls schwach mit dem Strukturwissen korrelierten Strategien, die mit der – oft zufällig anmutenden, unsystematischen – Veränderung von drei bzw. sämtlichen exogenen Variablen einhergingen.

Empirische Arbeiten, die den Strategieeinsatz nach der Anzahl simultan manipulierter Variablen klassifizierten, kamen fast immer übereinstimmend zu dem Ergebnis, dass der Gebrauch von Strategien Wissenserwerb und Steuerung positiv bedingt (s. z. B. Vollmeyer, Burns & Holyoak, 1996). Möglicherweise ist das strategische Wissen, d. h. Wissen,

wie man in systematischer Weise mit einem komplexen System umgeht, sogar in der Lage, Defizite im Strukturwissen auszugleichen, wenn es um eine geschickte Steuerung des Systems geht (s. Schoppek, 2002).

3.4.3 Motivation und Emotion

Dass kognitive Prozesse nicht nur voneinander und von kognitiven Personenmerkmalen des Problemlösers, sondern überdies von Emotionen abhängen, gilt als empirisch gut belegt (vgl. Funke, 2003, S. 203). Den anerkannten Rahmenmodellen nach (s. z. B. Fiedler, 1988, 2000) bedingen emotionale Zustände mit, welchen Zugang eine Person zu einer aktuellen kognitiven Anforderung wählt: In positiver Stimmung überwiegt ein intuitiv-heuristischer Zugang, in negativer Stimmung analytisches, informationsgeleitetes Vorgehen. Diese Grundtendenzen hinsichtlich Problemlöseverhalten finden sich speziell für den Umgang mit komplexen, dynamischen Systemen bei Spering (2001) bestätigt, wenn auch ohne Unterschiede in der resultierenden Problemlösegröße.

So wie Emotionen die Herangehensweise an ein Problem beeinflussen, wirkt sich auch Motivation nur vermittelt über das Problemlöseverhalten auf die Leistung aus: Die Dauer und die Häufigkeit, mit der sich jemand mit einem Problem auseinandersetzt, die Qualität der Problemlöseaktivität (z. B. strategisches Vorgehen), der Konzentrationszustand, in dem jemand ein Problem löst, sind Beispiele für vermittelnde Variablen (s. Vollmeyer & Rheinberg, 1998). Es bleibt zu beachten, dass Motivation niemals ein eindimensionales Konstrukt ist, sondern viele Komponenten umfasst. Ebenfalls zu berücksichtigen ist, inwieweit man motivationale Einflüsse als aktuelle Zustände oder als stabilere Persönlichkeitsmerkmale (Zielorientierungen) untersucht (vgl. Niemivirta, 2002).

Relativ gut untersucht ist der Motivationsfaktor „Erfolgszuversicht gegenüber Misserfolgsmotiviertheit“. Vollmeyer und Rheinberg (1998) fanden in einem pfadanalytischen Modell positive Auswirkungen von Erfolgszuversicht einerseits auf die Systematik verwendeter Strategien in einem Szenarioproblem, andererseits auf die anstrengungsfreie Konzentration. Beide Faktoren, Systematik und Konzentration, standen wiederum in enger Verbindung mit der Güte des Wissenserwerbs und der Steuerleistung.

Trotz ihrer unbetrübten Bedeutung wird die Frage nach Personenfaktoren in der eige-

nen (allgemeinpsychologisch ausgerichteten) Untersuchung kein Schwerpunktthema sein. Die Analyse von Personenfaktoren wird sich darauf beschränken, ausgewählte emotionale und motivationale Zustandsvariablen als Kontrollfaktoren einzubeziehen. So kann geprüft werden, ob Probanden unterschiedlicher experimenteller Bedingungen in Ausgangsmotivation und emotionaler Befindlichkeit vergleichbar sind. Hinsichtlich Intelligenz wird ein ähnliches Vorgehen gewählt. Da ein psychometrischer Test aus ökonomischen Gründen nicht in Frage kommt, werden – wenn auch als unstandardisiertes Maß – schulische Leistungen erfasst werden, die nach Amelang (2001) mit dem psychologischen Intelligenzkonstrukt in enger Verbindung stehen und zumindest indirekte Rückschlüsse auf die intellektuelle Befähigung einer Person zulassen. Dabei werden sowohl Gesamtleistungen erfragt als auch Schulleistungen speziell in Mathematik, die zumal beim numerischen Format des verwendeten Szenarios Prädiktoren der Problemlösequalität sein könnten (vgl. Beckmann, 1994, S. 66; Schulz, 2003). Es werden also positive Zusammenhänge zwischen Systemidentifikation, Systemsteuerung und schulischen Leistungen erwartet.

3.5 Hypothesen

Auf der Grundlage der vielfältigen Determinanten im Problemlöseprozess lassen sich die inhaltlichen Hypothesen dieser Arbeit wie folgt zusammenfassen:

Hypothese 1: Die Steuerung eines komplexen, dynamischen Systems kann man als Anwendung erworbenen Strukturwissens auffassen. Strukturwissen ist ein Prädiktor für erfolgreiche Systemsteuerung. („*Steuern impliziert Wissen*“)

Hypothese 2: Die Vernetztheit eines Systems stellt einen Schwierigkeitsindikator dar: Mit zunehmendem Vernetztheitsgrad werden Systemidentifikation und -steuerung erschwert. (*Vernetztheitshypothese*)

Hypothese 3: Strukturwissen ist mit zunehmender Vernetztheit des Systems ein geringerer Prädiktor für erfolgreiche Systemsteuerung. (*Interaktionshypothese*)

Hypothese 4: Der Einsatz von prozessbegleitenden Kausaldiagramm-Analysen unterstützt Wissenserwerb und Steuerprozesse durch verstärkte Hypothesenbildung. (*Reaktivitätshypothese*)

Hypothese 5: Schulische Leistungen, sowohl insgesamt als auch spezifisch auf mathematische Fähigkeiten bezogen, sagen den Problemlöseerfolg hinsichtlich Systemidentifikation und -steuerung vorher. („*Intelligenz entspricht Problemlösekompetenz*“)

4 Methoden

Das folgende Kapitel beschreibt, mit welchen Methoden die theoretisch begründeten Hypothesen empirisch zu prüfen sind. Dargelegt werden das experimentelle Design, die verwendeten Materialien einschließlich der Computersimulation, eine Beschreibung der Stichprobe und des Versuchsablaufs sowie generelle Anmerkungen zur statistischen Datenanalyse.

4.1 Versuchsdesign

Überprüfung der Reaktivitätshypothese. Die zentrale Hypothese, dass Kausaldiagramm-Analysen im Problemlöseprozess reaktiv wirken, erfordert den experimentellen Vergleich von Probanden, die Kausaldiagramme bearbeiten (Kausal-EG), mit Kontrollprobanden, die dasselbe Szenario ohne intermittierende Wissensdiagnose bearbeiten (Baseline-KG). Mit diesem Vergleich sollten sich generelle Reaktivitätseffekte von Kausaldiagramm-Analysen gegenüber dem im Rahmen eines experimentellen Settings „natürlichsten“ Bearbeitungsmodus prüfen lassen.

In Erweiterung des einfachen Experimental-Kontrollgruppen-Designs werden zwei zusätzliche Bedingungen aufgenommen. Sie sollen helfen, den Wirkmechanismus der vermuteten Reaktivität genauer zu untersuchen, indem sie zwei, wenn auch weniger plausible Alternativinterpretationen berücksichtigen. Für den Fall, dass sich Kausal-EG-Teilnehmer den Baseline-KG-Probanden in Systemidentifikation und -steuerung überlegen zeigen, wäre die naheliegendste Interpretation, dass die besonderen Anforderun-

gen der Kausaldiagramm-Analyse (systematisches Analysieren von Systembeziehungen, Formulieren expliziter, abstrakter Hypothesen) zu gesteigertem Strukturwissen durch hypothesengeleitetes Vorgehen führen (s. o.). Doch vielleicht unterstützt jede Art prozessbegleitender Wissensdiagnostik den Wissenserwerb, auch ohne die speziellen Anforderungen der Kausaldiagramm-Analyse? Diese Überlegung führt ausgehend von den in Abschnitt 2.3.1 vorgestellten Messverfahren dazu, als zweite Experimentalgruppe eine wissensdiagnostische Bedingung zu realisieren, die anstelle von abstraktem Regelwissen oberflächliches Eingriffswissen erfragt (Oberflächenwissen-Experimentalgruppe, OW-EG).

Eine Minimalinterpretation möglicher Reaktivitätseffekte könnte weiterhin darin bestehen, Leistungsunterschiede im Problemlösen auf die bloße Unterbrechung des Prozesses zurückzuführen, unabhängig davon, ob die intermittierende Aufgabe systembezogenes Wissen abfragt oder für die Problemstellung inhaltlich irrelevant ist. Es könnte sein, dass sich Problemlöser nach unterbrochener Systemexploration und -steuerung schon allein deshalb verstärkt an abstrakten Hypothesen orientieren, weil konkrete Einzelepisoden, die Grundlage einer alternativen Suche im Experimenterraum, im Arbeitsgedächtnis nicht mehr hinreichend verfügbar sind. Emotionale Einflüsse wirken möglicherweise in dieselbe Richtung, wenn man mit Mandlers (1964, 1975) Emotionstheorie physiologische Erregung durch Unterbrechung mentaler Strukturen vorhersagt und die Erregung als Ärger über die unterbrochene Problemlöseaktivität attribuiert. Negative Emotionen wie Ärger könnten in diesem Fall eine analytisch-abstrakte Herangehensweise an das Problem bedingen (s. Fiedler, 1988, 2000). Die teilweise emotionsbezogene Alternativinterpretation legt nahe, als szenario-irrelevante Distraktoraufgabe einer Unterbrechungskontrollgruppe (UB-KG) die Abfrage von aktuellen emotionalen Zuständen zu wählen. Diese erlaubt für einen Teil der Stichprobe zugleich einen Einblick in eventuelle emotionale Veränderungen im Verlauf der Szenariobearbeitung.

Überprüfung der Vernetztheitshypothese. Effekte der Systemvernetztheit sollen wie bei Funke (1985) durch variierte Anzahlen von Systemrelationen überprüft werden (s. S. 27). Um das Design zu vereinfachen, wird der vormals dreistufige Faktor Vernetztheit auf zwei Abstufungen reduziert. In einer informellen Voruntersuchung erwies sich

schon die sogenannte mittelschwere ÖKOSYSTEM-Variante als durchaus anspruchsvoll, so dass in dieser Untersuchung neben der mittelschweren Variante (sechs Teilprozesse) nur die einfache Version (vier Teilprozesse) zum Einsatz kommen soll. Insbesondere zwischen diesen beiden Vernetztheitsgraden hatten sich bei Funke (1985) klare Bedingungsunterschiede hinsichtlich der Problemlösegröße gezeigt.

Da jede Bedingung des Faktors Wissensdiagnose auf jeder der beiden Vernetztheitsstufen realisiert werden soll, ergibt sich ein vollständig gekreuztes 2×4 -Design. Um für alle Versuchspersonen klar definierte Problemanforderungen zu schaffen, wird abweichend von Funke (1985) auf reine Explorationsphasen verzichtet. Die Anforderung „Exploriere das System!“ ohne Nennung von Erfolgskriterien könnte andernfalls für Kontrollprobanden ohne Wissenstest allzu unspezifisch sein. Stattdessen werden allen Teilnehmern bereits vor Beginn der Simulation spezifische Zielwerte genannt, die in kombinierten Explorations- und Steuerphasen zu erreichen sind.

Es ist davon auszugehen, dass sich in einer solchen zielspezifischen Problemsituation Effekte der Wissensdiagnosebedingung gerade darin äußern, dass Problemlöser die parallelen Anforderungen Wissenserwerb und Steuerung des Systems unterschiedlich gewichten und unterschiedlich gut miteinander verknüpfen. Während in den Kontrollbedingungen ohne Wissenstests steuerungsbezogene Mittel-Zweck-Analysen überwiegen sollten, sollten Teilnehmer der Kausal-EG über die Ziel-Erreichungs-Strategien hinaus Gebrauch von Strategien wie VOTAT machen, die den Erwerb von breitem Strukturwissen unterstützen. Die Bearbeitung von Kausaldiagrammen sollte also den Beeinträchtigungen entgegenwirken, die Zielspezifität auf den Wissenserwerb ausübt.

4.2 Materialien

4.2.1 Beschreibung der verwendeten Simulation

Bei dem verwendeten System ÖKOSYSTEM handelt es sich um ein semantisch eingekleidetes, komplexes, dynamisches System, das in JAVA nach dem DYNAMIS-Ansatz mit den in Abschnitt 2.3 beschriebenen Systemeigenschaften programmiert wurde. Die drei exogenen Variablen sind per Tastatureingabe oder per Mausclick am Scrollbalken ma-

ÖKOSYSTEM 1985 (A) (Durchgang 1/5)

	Takt 1/7	Takt 2/7	Takt 3/7	Takt 4/7	Takt 5/7	Takt 6/7
Käfer	200	180	162	145	181	
Wasserverschmutzung.	10	9	9	9	8	
Blätterzahl	5000	5000	5000	7000	7000	
Gif	10	0	0	10	<input type="text"/>	<input type="text"/>
Schädlingsefresser...	0	0	0	50	<input type="text"/>	<input type="text"/>
Dünger	0	0	200	0	<input type="text"/>	<input type="text"/>

OK!

Abbildung 2: Benutzeroberfläche der Simulation ÖKOSYSTEM.

nipulierbar. Eingegeben werden können alle ganzzahligen Werte zwischen -999999 und $+999999$. Keine Eingabe wird vom System als Nulleingriff interpretiert. Über Mausklicke auf einen „OK“-Button teilen Probanden dem System mit, wann eine Intervention abgeschlossen ist, wobei das System neben den Eingriffen auch automatisch die Bearbeitungsdauer für jeden Takt aufzeichnet. Die Outputwerte der drei endogenen Variablen werden ebenfalls als Zahlen im Tabellenformat dargestellt (s. Abb. 2).

4.2.2 Abfrage von Strukturwissen

Strukturwissen wird mit Hilfe von Kausaldiagramm-Analysen im Paper-and-Pencil-Format erfasst. Ein standardisierter Vorlagebogen, um Kausalbeziehungen zwischen den Systemvariablen einzutragen, findet sich einschließlich der Instruktion im Anhang (S. 87-90). Die Instruktion gibt Aufschluss über die drei Wissenskomponenten Relation, Richtung und Stärke von Systembeziehungen sowie über das Konzept der Eigendynamik.

4.2.3 Abfrage von Oberflächenwissen

Um oberflächliches Systemwissen zu erfragen, wird ein computerbasierter Test verwendet, der sich an den Rekognitionsmaßen für spezifisches Eingriffswissen (s. S. 22; s. Preußler, 2001) orientiert. Am Bildschirm werden den Probanden nach entsprechender Vorbereitung und Instruktion sogenannte Ausgangs- und Folgezustände in tabellarischer Anordnung dargeboten (s. Abb. 3, Instruktionstext im Anhang S. 92). Ein

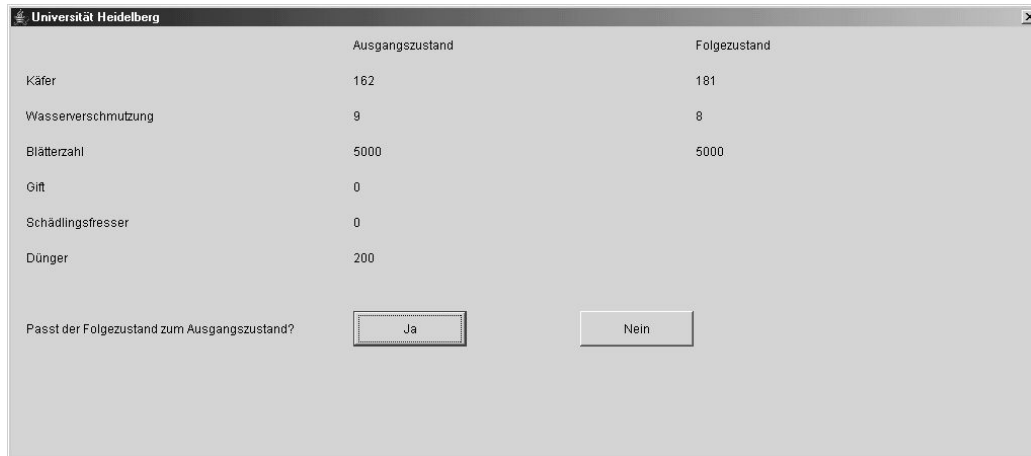


Abbildung 3: Benutzeroberfläche des computerbasierten Rekognitionstests.

Ausgangszustand zeigt einen zufällig ausgewählten Takt,⁹ mit dem eine Versuchsperson im vorherigen Bearbeitungsdurchgang konfrontiert war: oben die Zustandswerte der drei endogenen Variablen, darunter die Werte der exogenen Variablen, d. h. die Eingriffe, die eine Versuchsperson im selben Takt vorgenommen hat. Ein Folgezustand enthält demgegenüber nur Werte für die drei endogenen Variablen. Diese Werte können entweder vollständig denen des tatsächlichen Folgetakts entsprechen (*Target*), oder sie können dem Folgezustand lediglich ähneln (*Distraktoren*). Distraktoren werden erzeugt, indem das System für jede endogene Variable per Zufall Zustandswerte aus einem „unpassenden“ Takt auswählt, der weder der Ausgangstakt noch der tatsächliche Folgetakt ist. Mit einer Wahrscheinlichkeit von jeweils .5 werden Targets und Distraktoren in zufälliger Abfolge vorgelegt, vier Abfragen nach jedem Durchgang. Es ist die Aufgabe der Versuchspersonen zu unterscheiden, ob ein Folgezustand Target oder Distraktor ist. Neben dieser Einschätzung im Ja-Nein-Format geben Probanden auch ihre Antwortsicherheit auf einer zehnstufigen Skala von 1 (sehr unsicher) bis 10 (sehr sicher) an.

4.2.4 Abfrage emotionaler Zustände

In einem eigenkonstruierten, verbalen Kurzfragebogen sollen Probanden über ihre aktuelle emotionale Befindlichkeit vor, nach und – im Fall der UB-KG – auch während

⁹Von insgesamt sieben Takten zog das Programm jeweils einen der Takte zwei bis sechs. Der erste und der letzte Takt konnten aus technischen Gründen nicht gezogen werden.

der Szenariobearbeitung Auskunft geben. Der Fragebogen besteht aus einer Liste von 20 Adjektiven, die der Eigenschaftswörterliste (EWL) von Janke und Debus (1978) entnommen sind und sich fünf relevanten Dimensionen zuordnen lassen: allgemein positive Stimmung (z. B. unbesorgt, glücklich; negativ gepolt: pessimistisch, ratlos), Ärger (gereizt, verärgert), Aufregung (aufgeregt, nervös), Ängstlichkeit (ängstlich, zaghaft) und Aktiviertheit oder Konzentration (z. B. interessiert, aktiv; negativ gepolt: gelangweilt, müde). Für jedes Eigenschaftswort in der Liste gibt ein Teilnehmer auf einer fünffach abgestuften Skala mit den Polen „nicht vorhanden“ und „sehr stark vorhanden“ an, ob und wie intensiv er ein Gefühl zum Befragungszeitpunkt empfindet. Für die mehrfache Emotionsabfrage gibt es sechs verschiedene Varianten, in denen die Reihenfolge der 20 Adjektive variiert; so sollen Antworttendenzen durch Wiederholung vermieden werden. Eine Beispielfassung ist im Anhang (S. 91) gegeben.

4.2.5 Schulische Leistungen und soziodemografische Variablen

Auch die hypothesenrelevanten schulischen Leistungen werden als verbale Angaben zu Abiturnote¹⁰ und zur letzten Schulnote in Mathematik mit einem Paper-and-Pencil-Fragebogen erfasst. Weitere Kontrollvariablen sind Geschlecht, Alter und Beruf der Teilnehmer.

4.3 Teilnehmer

4.3.1 Vorüberlegungen zur Stichprobengröße

Ein geeigneter Stichprobenumfang, für den die geplanten statistischen Verfahren weder inhaltlich bedeutsame Effekte unterschätzen noch praktisch unbedeutsame Effekte überschätzen, wurde per a priori Poweranalyse mit der Statistik-Software GPower (Erdfelder et al., 1996) ermittelt. Richtwerte für erwartete Effektgrößen lieferten post hoc Poweranalysen von Funkes (1985) Daten: Effektgrößen schwanken hier zwischen $f = .47$ bei Kontrasten für den Faktor Vernetztheit, was nach den Konventionen von Cohen (1977) großen Effekten entspricht, und Effektgrößen nahe null. Überwiegend finden sich kleine bis mittlere Effekte um $f = .17$. Wenn auch mittlere Effekte für die geplante

¹⁰Es war mit ausschließlich studentischen Probanden zu rechnen.

Untersuchung ohne Zweifel am realistischsten scheinen, orientierte sich die Poweranalyse dennoch an großen Effekten, um den durchführungstechnischen Aufwand für das 2×4 -ANOVA-Zwischengruppen-Design in einem ökonomischen Rahmen zu halten.¹¹ Angenommen einen Alpha-Fehler von .05 und einen gemäß Cohens (1977) Konventionen vervierfachen Beta-Fehler von .20, gegeben mittlere Effekte, wären im geplanten ANOVA-Design 128 Probanden erforderlich gewesen, um eventuelle Signifikanzen für den zweistufigen Faktor Vernetztheit valide zu interpretieren, für den vierstufigen Faktor Wissensdiagnose sogar 179 Probanden. Für den Fall fester Stichprobengrößen bzw. erhebungsbedingter Beschränkungen schlugen Erdfelder et al. (1996) vor, in sogenannten Kompromissanalysen die üblichen Setzungen für Alpha- und Beta-Niveau zugunsten auch unkonventioneller Irrtumswahrscheinlichkeiten aufzugeben, indem man Poweranalysen für einen vorgegebenen Stichprobenumfang berechnet, üblicherweise bei Gleichgewichtung des Alpha- und Beta-Fehlers. Eine solche Kompromissanalyse im Acht-Gruppen-ANOVA-Design für große Effekte ($f = .4$) bei einem Zählerfreiheitsgrad liefert bei einer Stichprobengröße von $n = 64$ einen Alpha- und Beta-Fehler von .08, d. h. eine statistische Power von .92. Diese Werte stimmen weitgehend mit der Power und dem Signifikanzniveau bei Funke (1985) überein ($\alpha = \beta = .10$), so dass der Stichprobenumfang von $n = 64$ als gerade noch vertretbare Minimallösung angestrebt wurde unter dem Vorbehalt, dass die statistische Power geringer ausfällt, wenn signifikante oder insignifikante insbesondere kleine oder mittlere Effekte für ein konventionelles Alpha-Niveau von .05 interpretiert werden.

4.3.2 Stichprobenauswahl und -zusammensetzung

Insgesamt nahmen 69 Personen (52 weiblich, 17 männlich) an der Untersuchung teil. Nach dem Ausschluss von fünf Personen¹² aus der Auswertung verblieben 64 vollständige

¹¹Die Poweranalyse stützte sich auf varianzanalytische Auswertung, da dieses Verfahren die zentralen Hypothesen überprüft. Nebenhypothesen erfordern allerdings eine Auswertung über ANOVA mit Messwiederholung oder Korrelationsanalysen.

¹²In zwei Fällen führten Computerabstürze zum völligen Verlust der Daten, in einem Fall ließen Nachfragen bei der Bearbeitung vermuten, dass die Instruktion aufgrund von Sprachschwierigkeiten nicht richtig verstanden worden war, in einem Fall wurde ein Kausaldiagramm ausgelassen, ein weiteres Mal wurde beim Steuern des Systems versehentlich ein falscher Zielwert angestrebt; die Versuchsperson

Datensätze von 48 Teilnehmerinnen (75 %) und 16 Teilnehmern (25 %), acht Probanden pro Bedingung bei gleichem Geschlechterverhältnis in jeder Versuchsbedingung.

Knapp vier Fünftel (78.1 %) der Probanden waren Studentinnen und Studenten im Diplomstudiengang Psychologie ($n = 50$). Die übrigen Probanden waren in anderen Fächern eingeschrieben ($n = 11$), u. a. als Nebenfachstudierende in Psychologie, oder hatten ein Studium bereits abgeschlossen ($n = 3$). Der hohe Anteil an Psychologiestudierenden einschließlich der charakteristischen Geschlechterverteilung (75 % weibliche, 25 % männliche Versuchspersonen) ergibt sich aus einer vorwiegenden Rekrutierung am Psychologischen Institut. Mit Ankündigungen in Vorlesungen des Grundstudiums und Aushängen wurde auf das Experiment aufmerksam gemacht. Psychologiestudenten im Grundstudium erhielten eine Bescheinigung über Versuchspersonenstunden. Unabhängig davon wurde unter allen Teilnehmern ein USB-Memory-Stick verlost.

Das mittlere Alter in der Stichprobe lag bei $M = 25.7$ Jahren ($SD = 6.52$). Überwiegend (84.4 %) nahmen Probanden bis 30 Jahre einschließlich teil. Da die Teilnahme nicht an Altersvoraussetzungen gebunden war, fand sich jedoch eine große Variationsbreite bei einem Minimum von $min = 19$ und einem Maximum von $max = 48$ Jahren. Der Durchschnitt der Abiturnoten in der Stichprobe lag bei $M = 1.7$ ($SD = .59$). Fast drei Viertel der Teilnehmer (74.6 %; $n = 47$)¹³ hatten einen Abiturdurchschnitt von 2.0 oder besser bei einem Range von 1.0 bis 3.5. Als letzte Mathematiknote in der Schule hatte über die Hälfte der Probanden (54.1 %; $n = 33$)¹⁴ eine Note im Einserbereich. Im Mittel fanden sich mit $M = 11.2$ Punkten ($SD = 3.54$) Mathematikleistungen im Zweierbereich.

Sowohl Abiturdurchschnitt als auch Mathematiknoten lassen vermuten, dass es sich um überdurchschnittlich leistungsfähige Probanden handelt, die diesbezüglich sehr wahrscheinlich zwar keine repräsentative Auswahl aus einer studentischen Population darstellen, aber insgesamt über das erforderliche Interesse und die nötige Motivation zum Bearbeiten eines komplexen Problemlöseszenarios verfügen.

bekanntes auf Nachfrage, sich um eine Kommastelle „verlesen“ zu haben.

¹³ von 63 gültigen Angaben

¹⁴ von 61 gültigen Angaben

4.4 Versuchsdurchführung

Die Erhebung fand vom 1. bis zum 23. Juni 2005 im Psychologischen Institut der Universität Heidelberg statt. Nach Möglichkeit wurden Einzelsitzungen vereinbart, für die ein kleiner, separater Testraum zur Verfügung stand. Ließen sich Terminüberschneidungen nicht vermeiden, konnten in einem Nebenraum bis zu zwei weitere Personen parallel teilnehmen, ohne einander bei der Szenariobearbeitung zu beobachten. Eine Sitzung dauerte für jeden Probanden ungefähr 45 Minuten. Dieser Zeitrahmen wurde den Teilnehmern bereits vorab als Richtwert mitgeteilt. Sie wurden jedoch darauf hingewiesen, dass der Versuch unter Umständen deutlich weniger (ca. 30 Minuten) oder mehr (ca. 60 Minuten) Zeit in Anspruch nehmen könnte. Um nicht in Termindruck zu geraten, wurden alle Teilnehmer gebeten, auf einen einstündigen Versuch vorbereitet zu sein. Eine Zeitbegrenzung für die Bearbeitung von ÖKOSYSTEM gab es nicht.

Probanden wurden den acht experimentellen Bedingungen zufällig zugeteilt mit der Einschränkung, dass das Anzahlverhältnis männlicher und weiblicher Versuchspersonen pro Bedingung gleich sein sollte (Parallelisierung nach Geschlecht). Zu Beginn lasen alle Probanden eine zweiseitige Einführung in die Rahmengeschichte des Szenarios. Diese verlangt, sich in die Lage eines Kleingärtners hineinzusetzen, in dessen Garten ein einziger Baum steht. Der Baum soll über einen simulierten mehrwöchigen Zeitraum gepflegt werden, so dass folgender Zielzustand möglichst früh und exakt eintritt und gehalten wird: Im Idealfall hat der Baum *10000 Blätter*, *100 Käfer* sitzen darauf, und die *Grundwasserverschmutzung* hat den Wert *null*. Gleichzeitig verweist die Instruktion auf die Notwendigkeit, das System zu explorieren:

„Nehmen Sie sich Zeit zu beobachten, wie sich Gift, Schädlingsfresser und Dünger auf Blätterzahl, Käfer und Wasserverschmutzung auswirken [...] Wichtig ist, dass Sie sich *nicht auf eine Vorstellung verlassen, wie etwas wirken sollte*, sondern beobachten, was geschieht.“

Die Gewichtung und Koordination der beiden Ziele Wissenserwerb und Wissensanwendung blieb – im Gegensatz zur früheren Untersuchung (Funke, 1985) – den Probanden überlassen.

Nach einem Fragebogen zur Erfassung von Eingangsmotivation und emotionaler Stimmung bearbeiteten alle Versuchspersonen das Szenario ÖKOSYSTEM selbstständig am Computer für fünf kombinierte Explorations- und Steuerphasen mit jeweils sieben Takten bzw. sieben simulierten Wochen.¹⁵ Jeder Durchgang begann mit denselben Startwerten (Käfer: 200, Wasserverschmutzung: 10, Blätter: 5000), die von vorherigen Durchgängen unabhängig waren. Für Notizen lagen Kugelschreiber und Papier bereit.

Die vier Intervalle zwischen den fünf Szenariodurchgängen gestalteten sich je nach Bedingung des Faktors Wissensdiagnose verschieden: Teilnehmer der Kausal-EG bearbeiteten jedesmal ein Kausaldiagramm, Teilnehmer der OW-EG die Entscheidungsaufgabe am Computer, Teilnehmer der UB-KG einen Emotionsfragebogen; Teilnehmer der Baseline-KG übersprangen das Intervall, um sofort mit dem nächsten Durchgang fortzufahren. In der Kausal-EG konnten die Probanden in späteren Durchgängen frühere Kausaldiagramme einsehen und bekannte Variablenzusammenhänge gegebenenfalls in das neue Diagramm übertragen.

Im Anschluss an den fünften Durchgang folgte zunächst der Oberflächenwissen-Test, dann ein Kausaldiagramm für sämtliche Teilnehmer,¹⁶ zuletzt ein abschließender Emotionsfragebogen und ein Fragebogen bezüglich soziodemografischer Kontrollvariablen und schulischer Leistungen. Probanden erhielten Süßigkeiten zur Belohnung, die Möglichkeit zur Teilnahme am Gewinnspiel, auf Wunsch eine mündliche Aufklärung über Thema und Ziel der Untersuchung sowie die Gelegenheit, eine individuelle Leistungsrückmeldung anzufordern. Individualergebnisse hinsichtlich Steuerleistung und Kausalwissen wurden etwa drei Wochen nach der Erhebung per E-Mail verschickt. Der Versuchsablauf ist zusammenfassend in Tabelle 1 wiedergegeben.

¹⁵Zustände der exogenen Variablen im achten Takt wurden noch angezeigt, aber ohne weitere Eingriffsmöglichkeiten.

¹⁶Diese Reihenfolge wurde bevorzugt, da das Anfertigen eines Kausaldiagramms von einer vorherigen Oberflächenwissen-Abfrage weniger beeinflusst werden sollte als im umgekehrten Fall.

Tabelle 1: Überblick über den Versuchsablauf.

Experimentalgruppen		Kontrollgruppen	
Kausal-EG	OW-EG	UB-KG	Baseline-KG
Instruktion: Einführung in das Szenario			
Emotionsbogen A			
Szenario Durchgang 1			
Kausaldiagramm	Rekognitionstest	Emotionsbogen B	
Szenario Durchgang 2			
Kausaldiagramm	Rekognitionstest	Emotionsbogen C	
Szenario Durchgang 3			
Kausaldiagramm	Rekognitionstest	Emotionsbogen D	
Szenario Durchgang 4			
Kausaldiagramm	Rekognitionstest	Emotionsbogen E	
Szenario Durchgang 5			
Rekognitionstest			
Kausaldiagramm			
Emotionsbogen F			
Fragebogen: schulische Leistung, soziodemografische Variablen			

4.5 Datenanalyseverfahren und -instrumente

Die Auswertung der Teilnehmerdaten erfolgte mit dem Statistikprogramm SPSS (V 11.0, V 12.0), für spezielle Prozeduren ergänzend mit SAS (V 8, SAS Enterprise Guide V 2.0) und AMOS 5.0. Varianzanalytische Methoden (lineare ANOVA-Kontraste, univariate mehrfaktorielle ANOVA mit und ohne Messwiederholung, MANOVA), Korrelations- bzw. Regressionsanalysen wurden zur inferenzstatistischen Absicherung von Effekten herangezogen. Hierfür wurden standardmäßig die folgenden Voraussetzungen geprüft: für Korrelations- und Regressionsverfahren die Annahme univariater *Normalverteilung der Fehlerkomponenten* mit dem Kolmogorov-Smirnov-Anpassungstest (K-S-Test),¹⁷ für varianzanalytische Verfahren darüber hinaus die *Homogenität der Fehlervarianzen* mit dem Levene-Test, der als relativ robust gegenüber Verletzungen der Normalverteilungsannahme gilt (s. Bortz, 1999, S. 275). Auswertungsgrundlage dieser beiden Tests waren die Stichproben-Rohwerte, die sich von den Fehlertermen nur um den konstanten Mittelwert unterscheiden.

Eine dritte varianzanalytische Voraussetzung, die *Unabhängigkeit der Fehlerkomponenten*, wurde ohne empirische Prüfung angenommen, da sie laut Bortz (1999, S. 275) bei zufälliger Probandenzuteilung zu den experimentellen Bedingungen zu erwarten ist. Nur im Messwiederholungsfall testete Mauchlys Sphärizitätskriterium auf die geforderte Zirkularität. Für den multivariaten Fall konnte die Homogenität der Varianz-Kovarianz-Matrizen mit dem Box-M-Test geprüft werden, der jedoch anfällig gegenüber Verletzungen der Normalverteilung ist (Werner, 1997, S. 494). Von Tests auf multivariate Normalverteilung musste dabei wegen eingeschränkter empirischer Prüfmöglichkeiten Abstand genommen werden.

Abweichungen von den Voraussetzungen, Anwendung eines Analyseverfahrens unter Vorbehalt oder Verwendung nicht-parametrischer Alternativverfahren werden an entsprechender Stelle berichtet.

¹⁷Zur Überprüfung der Normalverteilung einzelner Treatmentstufen ließen sich multiple Kolmogorov-Smirnov-Tests und eine mögliche Alpha-Fehler-Kumulierung leider nicht vermeiden.

5 Ergebnisse

Die Darstellung der Ergebnisse beginnt mit einer Übersicht über die abhängigen Variablen. Daran anschließend soll zunächst auf Befunde zur Beziehung zwischen Wissen und Steuern (Hypothese „Steuern impliziert Wissen“ und Interaktionshypothese) eingegangen werden. Es folgen Befunde zu Effekten von Systemvernetztheit (Vernetztheitshypothese), Kausaldiagramm-Analysen (Reaktivitätshypothese) und den indirekten Intelligenzmaßen (Hypothese „Intelligenz entspricht Problemlösekompetenz“). Der letzte Abschnitt des Kapitels ist Ergebnissen zum Einfluss von Emotionen auf die Problemlösefähigkeit gewidmet.

5.1 Abhängige Variablen

5.1.1 Güte der Kausaldiagramme

Als abhängiges Maß für den Erwerb von Strukturwissen wurde die oben beschriebene *Güte der Kausaldiagramme* (GdK; Funke, 1992b, Müller, 1993) herangezogen (s. S. 18). Analysen für einzelne Komponenten des Kausalwissens werden im Folgenden nur bei Auffälligkeiten dargestellt oder wenn spezielle Effekte zu erwarten sind. In der Regel werden Befunde für das Summenmaß GdK_{sum} berichtet, das nicht nur am umfassendsten ist, sondern auch die Bedingung normalverteilter Stichprobenwerte am besten erfüllt. Abweichend von der vorgeschlagenen Gleichgewichtung der drei Komponenten wurden GdK_{rel} und GdK_{vor} hier mit dem Faktor 1 gewichtet, GdK_{num} mit dem Faktor .5, was sich faktisch als Gleichgewichtung interpretieren lässt, wenn man die Ratewahrscheinlichkeiten (jeweils $p = .5$) der ersten beiden Komponenten berücksichtigt. Der pragmatische Grund für die geringere Gewichtung des Numerikwissens lag darin, dass GdK_{num} bei signifikanten K-S-Tests eine deutlich rechtsschiefe Verteilung aufwies.

5.1.2 Güte des Oberflächenwissens

Das in Rekognitionstests erfragte Oberflächenwissen eignet sich prinzipiell für eine Auswertung nach der Signalentdeckungstheorie (s. z. B. Velden, 1982). In einer gesonderten Betrachtung der OW-EG wird auf diese Möglichkeit eingegangen. Für die varianzanaly-

tischen und korrelativen Hauptanalysen, d. h. für den Vergleich mehrerer experimenteller Bedingungen, ist mit lediglich vier Entscheidungstrials aber keine Auswertungsgrundlage für signalentdeckungstheoretische Methoden gegeben. Es wurde der Index *Güte des Oberflächenwissens* (GdOW) gebildet, der Punkte für richtige und falsche Antworten mit ihren zugeordneten Antwortsicherheiten (S_i) nach der folgenden Formel verrechnet:

$$\text{GdOW} = \sum_{i=0}^{n(\text{richtig})} (S_i | \text{richtige Antwort}) - \sum_{j=0}^{n(\text{falsch})} (S_j | \text{falsche Antwort}) \quad (9)$$

wobei $n(\text{richtig})$ = Anzahl richtiger Antworten,

$n(\text{falsch})$ = Anzahl falscher Antworten

Bei maximal vier richtigen oder vier falschen Antworten mit einer maximalen Antwortsicherheit von 10 ergibt sich für GdOW ein ganzzahliger Wertebereich von +40 bis -40. Je höher der Wert, desto höher das Oberflächenwissen. GdOW-Werte von null oder weniger sind als Nullsensitivität aufzufassen. Hinsichtlich der Verteilung in der Stichprobe zeigte der K-S-Test für GdOW keine signifikanten Abweichungen von der Annahme einer Normalverteilung.

5.1.3 Güte der Systemsteuerung

Die *Güte der Systemsteuerung* (GdS) berechnete sich nach der von Müller (1993) entwickelten Formel (s. S. 23). Um nicht-definierte Logarithmen bei Istwert-Sollwert-Differenzen von null zu umgehen, wurde von Vollmeyer, Burns und Holyoak (1996, S. 85) das Vorgehen übernommen, zu jedem Differenzenbetrag pauschal den Wert 1 zu addieren. Trotz logarithmischer Transformation von GdS fanden sich in der Stichprobe einige zum Teil hohe Ausreißerwerte, die jedoch offensichtlich nicht auf planloses Herumexperimentieren der Probanden zurückgingen. Da ein sinnvolles Eliminationskriterium fehlte, mussten für GdS erhöhte Varianzen in Kauf genommen werden.

5.1.4 Güte der Strategieranwendung

Die *Güte der Strategieranwendung* (GdStA) wurde in Anlehnung an Vollmeyer et al. (1996) durch einen Strategiepunktwert operationalisiert. GdStA ist der Mittelwert aller

Strategiepunkte, die für jeden Interventionstakt vergeben werden: Zwei Punkte für einen strategisch besonders wirkungsvollen Null- oder Einfach-Eingriff, ein Punkt für einen Zweifach-Eingriff, null Punkte für einen oft wenig systematischen Dreifach-Eingriff. Hohe GdStA-Werte (Maximum: 2) gehen nach dieser Klassifikation mit systematischem Strategieeinsatz einher. Es bleibt anzumerken, dass der Strategiewert streng genommen nicht als normalverteilt vorausgesetzt werden kann, da der K-S-Test unter der Nullhypothese einer Normalverteilung für mehrere Treatmentstufen Signifikanzen an der Grenze zum 10 %-Niveau erbrachte.

5.1.5 Bearbeitungsdauer

Auch die Bearbeitungsdauer wurde als abhängige Variable herangezogen, wobei zwischen der Brutto-Bearbeitungszeit einer gesamten Experimentalsitzung und einer Netto-Bearbeitungszeit für das Szenario ohne wissensdiagnostische Tests und Instruktionen zu unterscheiden ist. Für einen sinnvoll interpretierbaren Gruppenvergleich kommt nur die Netto-Bearbeitungszeit in Frage, will man nicht die triviale Annahme bestätigen, dass es mehr Zeit in Anspruch nimmt, das ÖKOSYSTEM zu bearbeiten und Kausaldiagramme zu zeichnen als nur das ÖKOSYSTEM zu bearbeiten. Die Netto-Zeit berechnet sich als Summe des Zeitbedarfs pro Trial.¹⁸

5.2 Systemwissen und Systemsteuerung

5.2.1 Lerneffekte und Leistungen

GdS. Für GdS, separat berechnet pro Übungsdurchgang, zeigte eine Messwiederholungs-ANOVA in der Gesamtstichprobe ($n = 64$) hochsignifikante, globale Niveauunterschiede zwischen den fünf Messzeitpunkten ($F(4, 252) = 15.99, p < .0001$). Das Signifikanzniveau wurde trotz Verletzung der Zirkularitätsbedingung (Mauchly: $W = .51; p < .0001$) und Korrektur der Freiheitsgrade auf die Untergrenze erreicht. Eine alternative MANOVA, die Werner (1997, S. 494) für diesen Fall empfiehlt, erbrachte denselben

¹⁸Nur die im ersten Takt des ersten Durchgangs benötigte Zeit ging nicht in die Berechnung ein, weil einige Probanden zum besseren Aufgabenverständnis bereits beim Lesen der Instruktion die Computersimulation öffneten.

Schluss (Pillai-Spur: $V = .34$, $F(4, 60) = 7.63$, $p < .0001$).

Auf deskriptiver Ebene war in der Gesamtstichprobe eine kontinuierliche Abnahme von GdS zu finden bei gleichzeitig zunehmender Varianz wie für Übungseffekte typisch (s. Tabelle 2). Bei gesonderter Betrachtung der acht Treatment-Kombinationen wies der Verlauf dieselbe Tendenz auf, wenn auch z. T. mit diskontinuierlichen „Sprüngen“. Möglicherweise aufgrund von hohen Ausreißerwerten kam es vor, dass GdS in einem früheren Durchgang geringer war als im Folgedurchgang.

Interkorrelationen der fünf Steuerdurchgänge fielen durchweg positiv aus. Fanden sich für GdS im ersten Durchgang lediglich schwache Signifikanzen für den Zusammenhang zu den beiden Folgedurchgängen, waren ab dem zweiten Durchgang alle Interkorrelationskoeffizienten statistisch hochsignifikant ($p < .0001$) in der Größenordnung von $r = .56$ (zweiter und fünfter Durchgang) bis $r = .79$ (vierter und fünfter Durchgang). Die Höhe der GdS-Korrelationskoeffizienten schien dem Trend zu folgen: Je fortgeschrittener und je enger benachbart zwei Messzeitpunkte, desto höher der statistische Zusammenhang.

GdK. Mögliche Lerneffekte hinsichtlich Kausalwissen konnten nur für die Teilnehmer der Kausal-EG ($n = 16$), die je fünf Kausaldiagramme bearbeiteten, untersucht werden. In dieser Teilstichprobe zeigte sich ein signifikanter Einfluss des Messzeitpunkts auf GdK_{sum} in einer MANOVA, die wegen verletzter Zirkularität (Mauchly: $W = .14$; $p < .0017$) der Messwiederholungs-ANOVA vorgezogen werden musste (Pillai-Spur: $V = .67$, $F(4, 12) = 6.19$, $p = .0061$). Bei geringer Vernetztheit ($n = 8$) nahmen die GdK_{sum} -Mittelwerte monoton zu, bei hoher Vernetztheit ($n = 8$) fand sich nur in den ersten drei Durchgängen eine monotone Zunahme, danach eine annähernde Stagnation. Im Mittel steigerten Probanden der Kausal-EG ihre Identifikationsleistung in GdK_{sum} von $M = .45$ ($SD = .27$) im ersten Durchgang auf $M = .82$ ($SD = .38$) im fünften Durchgang (s. Tabelle 2).

GdOW. Für GdOW in der OW-EG ($n = 14$)¹⁹ stellte sich ein ANOVA-Messwiederholungsfaktor als insignifikant heraus ($F(4, 52) = 1.98$, $p < .112$); die voraussetzungärmere MANOVA, die üblicherweise weniger effizient testet, verfehlte knapp die

¹⁹Zwei Probanden hatten einzelne Abfragen im Computertest ausgelassen.

Signifikanz auf dem 5 %-Niveau ($F(4, 10) = 3.12, p = .066$). Die deskriptive Betrachtung der GdOW-Mittelwerte (mit relativ großen Streuungen) lässt jedoch in keinem Fall ein systematisches Verlaufsmuster erkennen (s. Tabelle 2).

Auch die Teilnehmer, die bis zum fünften Durchgang keine Erfahrungen mit dem Rekognitionstest gemacht hatten, beantworteten die Abfrage überzufällig gut: Die mittlere Rate richtiger Antworten nach dem fünften Durchgang lag in der Gesamtstichprobe ($n = 64$) bei $M = .78$ ($SD = .25$), signifikant höher als eine zu erwartende Rate-wahrscheinlichkeit von $p = .5$ ($t(63) = 9.15, p < .0001$). Noch höher fiel die mittlere Trefferrate aus ($M = .82, SD = .29$), während die mittlere Falschalarmrate bei $M = .28$ ($SD = .37$) lag. Die Antwortsicherheit für richtige Antworten ($M = 6.97, SD = 2.72$) war im Mittel signifikant höher als die Sicherheitseinschätzung bei Falschantworten ($M = 5.66, SD = 2.50$; t-Test für abhängige Stichproben:²⁰ $t(250) = 3.18, p = .002$).

Für Steuerleistung und Kausalwissen lassen sich also deutliche Leistungszuwächse, insbesondere in den Anfangsdurchgängen, feststellen. Obwohl vergleichbare Effekte für Oberflächenwissen unter der vorliegenden Operationalisierung unterblieben, wurde auch diese Aufgabe zuverlässig bearbeitet. Es ist davon auszugehen, dass sich die Teilnehmer ernsthaft mit dem ÖKOSYSTEM-Szenario und den verbundenen Fragestellungen auseinander gesetzt haben.

5.2.2 Zusammenhänge zwischen Strukturwissen und Systemsteuerung

In Anlehnung an Müller (1993) wurde der Zusammenhang zwischen latenten Variablen für Systemidentifikation (I) und Systemsteuerung (S) mit Hilfe von linearen Strukturgleichungsmodellen untersucht. Als manifeste Variablen wurden GdK_{rel} , GdK_{vor} und GdK_{num} nach dem fünften Durchgang sowie GdS pro Durchgang herangezogen ($GdS_1, GdS_2, GdS_3, GdS_4, GdS_5$). Da sich der Einfluss von GdS_1 als nicht signifikant herausstellte, wurde diese Variable aus dem Modell entfernt, so dass nur noch substantielle Faktoren mit einer Indikatorreliabilität von mindestens .4 verblieben, wie als Gütekri-

²⁰Der t-Test wurde trotz nicht gegebener Normalverteilung eingesetzt, da der vergleichbare nicht parametrische Wilcoxon-Test die ungleiche Anzahl richtiger gegenüber falscher Antworten nicht berücksichtigt hätte. Das tatsächliche Signifikanzniveau dürfte also höher sein als oben angegeben.

Tabelle 2: GdS, GdK_{sum} und GdOW im Verlauf der fünf Durchgänge.

Durchgang		1	2	3	4	5	MANOVA ^a
GdS (<i>n</i> = 64)	<i>M</i>	5.60	4.68	4.42	4.08	3.84	$F(4, 60) = 7.63, p < .0001$
	<i>SD</i>	1.35	1.57	2.15	2.18	2.24	
GdK _{sum} (<i>n</i> = 16)	<i>M</i>	.45	.61	.76	.74	.82	$F(4, 12) = 6.19, p = .0061$
	<i>SD</i>	.27	.50	.40	.42	.39	
GdOW (<i>n</i> = 14)	<i>M</i>	15.21	12.64	10.79	2.86	14.50	$F(4, 10) = 3.12, p = .066$
	<i>SD</i>	14.53	13.80	15.04	15.29	12.63	

^a*F*-Werte und Signifikanzniveaus stimmen für die vier gebräuchlichen multivariaten Statistiken überein.

terium lokaler Anpassung gefordert (für Details s. Tabelle 6 im Anhang S. 92).

Die globale Anpassungsgüte ist demgegenüber mit einem beinahe signifikanten χ^2 -Kriterium ($\chi^2 = 21.53, df = 13, p = .063$) als kritisch anzusehen: Trotz geringer Stichprobengröße ($n = 64$) ist die Wahrscheinlichkeit, dass die Datenstruktur von den Modellannahmen abweicht, relativ hoch. Nach Hair et al. (1998) ist eine Nichtsignifikanz im χ^2 -Anpassungstest überdies für Stichprobenumfänge unter $n = 100$ nicht sinnvoll interpretierbar. Ein *RMSEA* von .102 deutet ebenfalls auf eine suboptimale Modellpassung hin, während der Goodness-of-Fit-Index und die inkrementellen Maße *NFI* und *TLI* für ein angemessenes Modell sprechen ($GFI = .920, NFI = .927, TLI = .950$). Mit einem normierten $\chi^2 < 2$ ist von einem sparsamen Modell auszugehen ($\frac{\chi^2}{df} = 1.656$).

Obwohl die Anpassung des Modells teilweise diskutabel ist, sollte eine inhaltliche Interpretation möglich sein. Um die Lesbarkeit zu erleichtern, werden Korrelationen zwischen GdK und GdS im Folgenden für ein umgepoltes GdS berichtet, so dass inhaltlich positive Zusammenhänge auch rechnerisch positiv erscheinen. Im unstandardisierten Pfadmodell (Abb. 4) ist mit einem Pfadkoeffizienten von 12.72 ein substantiell positiver Prädiktionwert der latenten Variable Identifikation (I) auf die Steuerung (S) abzulesen. Die Höhe des Zusammenhangs entspricht einem Korrelationskoeffizienten von $r = .66$, beinahe übereinstimmend mit der zusätzlich ermittelten Korrelation zwischen der manifesten GdK_{sum} und der mittleren GdS über alle Durchgänge ($r = .64, p < .0001$).

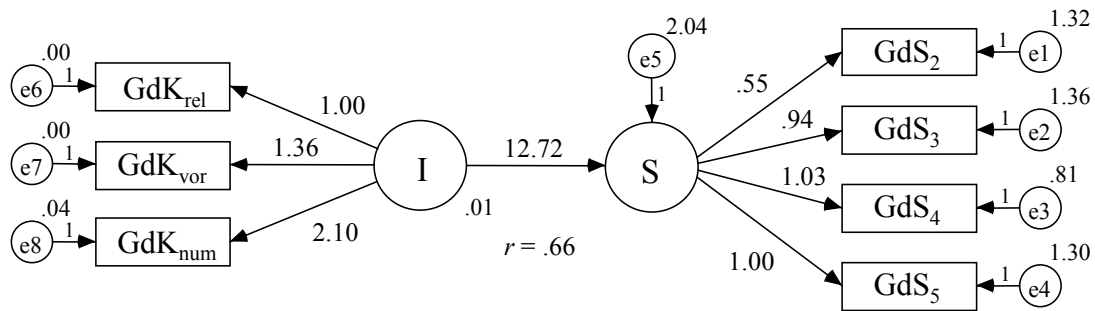


Abbildung 4: Pfadmodell zur Vorhersage von Steuerleistung aus Identifikationsleistung.

In der standardisierten Lösung zeigt GdK_{vor} unter den drei Wissenskomponenten das größte Einflussgewicht auf die Identifikationsvariable. Von den Steuerdurchgängen werden die Durchgänge drei bis fünf zur maximalen Datenpassung am stärksten gewichtet. Da sich die Identifikationsleistung aufgrund des Versuchsdesigns auf den fünften Durchgang bezieht, ist der Befund kaum überraschend. Untersuchungen über die Höhe des Zusammenhangs im Verlauf der Szenariobearbeitung konnten nur für die Kausal-EG erfolgen: Die Korrelationskoeffizienten zwischen GdK_{sum} am Ende eines Durchgangs und GdS im nachfolgenden Durchgang ließen jedoch keinen Trend wie etwa wachsende Zusammenhänge mit zunehmender Übung erkennen.

Ein Gruppenvergleich im Rahmen des Pfadmodells, um den vermuteten moderierenden Effekt von Vernetztheit auf den Zusammenhang zwischen Identifikation und Steuerung inferenzstatistisch zu prüfen, erlaubte keine zuverlässigen Schätzungen, da die maximale Iterationszahl des Verfahrens überschritten wurde. Die manifeste Korrelation zwischen GdK_{sum} und dem mittleren GdS beträgt bei geringer Vernetztheit $r = .57$ ($p = .001$), bei hoher Vernetztheit $r = .64$ ($p < .0001$). Der insgesamt positive Zusammenhang zwischen Wissen und Steuern ist also in der vorliegenden Stichprobe entgegen der Erwartung bei hoher Vernetztheit keineswegs niedriger, sondern sogar höher als unter der schwach vernetzten Systembedingung.

Hypothese 3 hat sich demnach nicht bestätigt, während man im Hinblick auf Hypothese 1 davon ausgehen kann, dass Strukturwissen ein Prädiktor für Steuerleistung ist.

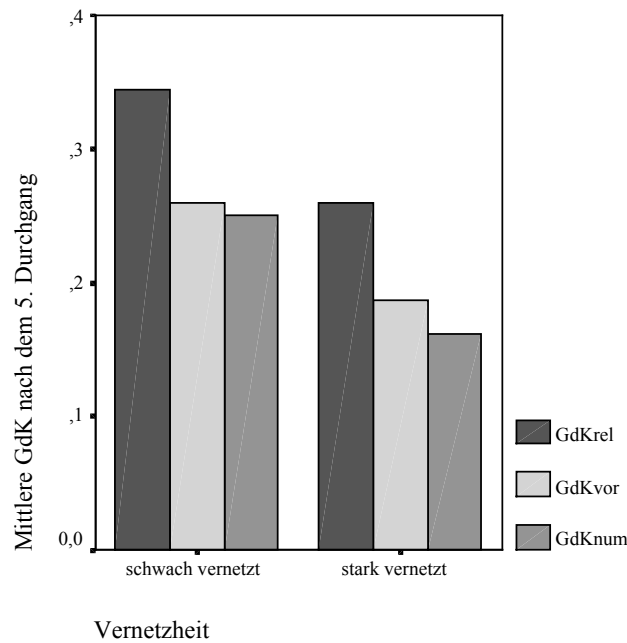


Abbildung 5: GdK_{rel}, GdK_{vor} und GdK_{num} unter den zwei Vernetztheitsbedingungen.

5.3 Vernetztheit

Hypothese 2 ging von einem Effekt der Systemvernetztheit auf die Problemlösequalität aus. ANOVA-Kontraste bzw. *t*-Tests prüften diese Annahme für die drei abhängigen Variablen GdS, GdK und GdOW. Wie bereits bei Funke (1985) unterschieden sich Teilnehmer der zwei Vernetztheitsbedingungen signifikant: Bei geringer Vernetztheit war GdS (Mittelwert über alle fünf Durchgänge) signifikant niedriger als bei großer Vernetztheit ($t(62) = -5.49, p < .0001, d = 1.39$) und GdK_{sum} nach dem fünften Durchgang signifikant höher ($t(62) = 2.28, p = .026, d = .57$). Mittelwerte und Standardabweichungen erscheinen in Tabelle 3 auf S. 61 in Verbindung mit den Effekten für den Faktor Wissensdiagnose. Dieselben Tendenzen findet man, wenn man die drei Wissenskomponenten GdK_{rel}, GdK_{vor} und GdK_{num} als separate abhängige Variablen untersucht. Die Ausprägung nimmt dabei auf beiden Vernetztheitsstufen von GdK_{rel} über GdK_{vor} zu GdK_{num} hin ab (s. Abb. 5).

Für GdOW in der Gesamtstichprobe war kein vergleichbarer Effekt festzustellen ($t(59) = .72, p < .474$). Der Einfluss von Vernetztheit auf Oberflächenwissen wurde zusätzlich ge-

sondert für die OW-EG nach dem signalentdeckungstheoretischen Ansatz auf der Basis von 20 Rekognitionsitems pro Versuchsperson untersucht. Mittelwertsvergleiche deuten darauf hin, dass d' , die Diskriminationsleistung bezogen auf bekannte gegenüber unbekanntem Items, mit größerer Vernetztheit abnimmt (schwach vernetzt: $M = 1.21$, $SD = .83$; stark vernetzt: $M = .74$, $SD = .46$). Eine mittlere Reaktionsneigung β von weniger als eins ($M = .87$, $SD = .39$) bei geringer Vernetztheit kennzeichnet ein eher laxes Antwortkriterium bzw. eine Ja-Sage-Tendenz, die bei hoher Vernetztheit nicht zu finden war ($M = 1.05$, $SD = .29$). Keine dieser Tendenzen war jedoch inferenzstatistisch abzusichern (für d' : $t(14) = 1.42$, $p = .178$; für β : $t(14) = -1.04$, $p = .317$). Das Ergebnis ist aufgrund von der geringen Teilstichprobengröße ($n = 16$) und aufgrund der wenigen Rekognitionsitems²¹ mit großer Vorsicht zu interpretieren. Weitere Kennwerte und ein Überblick zum Nachvollziehen der Auswertungsprozedur befinden sich daher nur im Anhang (S. 93).

5.4 Kausaldiagramm-Analysen, Problemlösegüte und Problemlöseverhalten

5.4.1 Systemwissen

Als Hauptindikator für den vermuteten Reaktivitätseffekt von Kausaldiagramm-Analysen wurde GdK_{sum} nach dem fünften Szenariodurchgang herangezogen. Die Auswertung erfolgte sowohl über eine globale 2×4 -ANOVA (Typ III-Quadratsummen), um eventuelle Interaktionen zwischen Wissensdiagnosebedingung und Vernetztheit zu prüfen, als auch über ANOVA-Kontraste. Um die Hypothese zu testen, dass sich Teilnehmer der Kausal-EG in GdK_{sum} vom Mittel der Teilnehmer anderer Gruppen unterscheiden, wurde ein a priori Kontrast formuliert. In der ANOVA zeigten sich auf dem 10%-Niveau signifikante Mittelwertsunterschiede zwischen den vier Diagnosebedingungen ($F(3, 56) = 2.41$, $p = .077$, $f = .33$). Die Interaktion mit Vernetztheit war nicht signifikant ($F(3, 56) = .79$, $p = .507$). Der a priori Kontrast wurde annähernd auf dem 1%-Niveau signifikant ($p = .012$). Von post hoc eingesetzten Vergleichen für die sechs Mittelwertspaare erreichte nur

²¹Nach Velden (1982, S. 46) sind für zuverlässige Analysen 500 Beobachtungen empfohlen.

Tabelle 3: GdK_{sum}, GdS und Netto-Bearbeitungszeit im 2 × 4 ANOVA-Design mit den Faktoren Vernetztheit und Wissensdiagnose.

Bedingung		<i>n</i>	GdK _{sum}		GdS		Zeit (<i>ms</i>)	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
schwach	Kausal-EG	8	.92	.42	3.35	1.18	1070.72	356.71
	vernetzt	8	.78	.38	4.21	1.66	761.45	294.63
	UB-KG	8	.60	.49	3.66	1.53	791.05	323.12
	Baseline-KG	8	.62	.27	3.47	1.38	870.05	196.87
	Gesamt	32	.73	.39	3.67	1.42	873.32	309.35
stark	Kausal-EG	8	.72	.36	5.02	.98	1084.78	556.50
	vernetzt	8	.39	.17	5.84	.83	740.24	176.34
	UB-KG	8	.58	.40	5.29	1.55	1387.64	609.12
	Baseline-KG	8	.41	.19	5.36	.55	922.68	324.52
	Gesamt	32	.53	.31	5.38	1.04	1033.83	492.87
Gesamt	Kausal-EG	16	.82	.39	4.18	1.36	1077.75	451.62
	OW-EG	16	.59	.32	5.03	1.52	750.84	234.82
	UB-KG	16	.59	.44	4.47	1.71	1089.35	562.83
	Baseline-KG	16	.51	.25	4.42	1.41	896.36	260.71
	Gesamt	64	.63	.37	4.53	1.50	953.58	416.13

der Paarvergleich zwischen Kausal-EG und Baseline-KG 10%-Signifikanz (Tukey HSD: $p = .068$). Probanden der Kausal-EG erzielten mit einer mittleren GdK_{sum} von $M = .82$ ($SD = .39$) den höchsten Wert, Probanden der Baseline-KG mit $M = .51$ ($SD = .25$) den niedrigsten Wert (s. Tabelle 3).

Die Analyse der abhängigen Variable GdOW erbrachte demgegenüber weder signifikante Gruppenunterschiede für den Faktor Wissensdiagnose (ANOVA: $F(3, 53) = .44$, $p = .727$) noch für die Interaktion mit dem ebenfalls nicht signifikanten Komplexitätsfaktor ($F(3, 53) = .77$, $p = .516$).

5.4.2 Steuerleistung

Analog zum Vorgehen für GdK_{sum} und GdOW wurden Effekte von Wissensdiagnose auf die mittlere Steuerleistung GdS geprüft. Die globale 2 × 4-ANOVA wies einen nicht signifikanten Faktor Wissensdiagnose auf ($F(3, 56) = 1.28$, $p = .291$) und eine definitiv zu vernachlässigende Interaktion ($F(3, 56) = .037$, $p = .990$). Die Irrtumswahrscheinlich-

keit für einen vorhandenen Mittelwertsunterschied zwischen der Kausal-EG und den drei Vergleichsgruppen lag im a priori Kontrast bei $p = .218$ mit einem negativen Kontrastgewicht, d. h. kleinerem mittlerem GdS für Versuchspersonen der Kausal-EG. Deskriptiv für die vorliegende Stichprobe betrachtet, war die Kausal-EG mit dem geringsten GdS allen anderen Versuchsbedingungen überlegen (s. Tabelle 3), jedoch jenseits jeder interpretierbaren Signifikanzgrenze.

5.4.3 Strategieeinsatz

Hypothese 4 postulierte nicht nur, dass prozessbegleitende Kausaldiagramm-Analysen mit höherem abschließend erworbenen Strukturwissen einhergehen, sondern auch dass dem Wissenserwerb ein verstärkt hypothesengeleitetes Vorgehen zu Grunde liegt. Als empirischer Indikator auf Hypothesenbildung kann der Strategieindex GdStA eingesetzt werden. Zu erwarten wäre ein systematischerer Strategieeinsatz von Probanden der Kausal-EG gegenüber denen der Reststichprobe.

Die mittlere GdStA über alle fünf Durchgänge wurde mit dem nicht parametrischen Kruskal-Wallis-Test auf Mittelwertsunterschiede zwischen den vier Diagnosebedingungen geprüft. Signifikante Unterschiede waren nicht zu finden ($\chi^2(3) = 3.015, p_{asy} = .389$). Genauso wenig konnte ein statistisch effizienterer a priori ANOVA-Kontrast (unter Inkaufnahme fehlender Normalverteilung) die Vermutung bestätigen, dass GdStA in der Kausal-EG gegenüber dem Mittel der drei anderen Gruppen systematisch erhöht ist ($p = .366$). Die deskriptive Betrachtung des Strategieeinsatzes lässt vermuten, dass ein Großteil der Probanden das Szenario unabhängig von der experimentellen Bedingung relativ systematisch bearbeitete, so dass die Varianz in der Gesamtstichprobe gering ist. Beinahe in der Hälfte aller Trials (47.8 %) wurden systematische Null- oder Einfach-Interventionen vorgenommen, weitere 30.7 % waren Zweifach-Eingriffe, und nur gut ein Fünftel (21.7 %) sind als strategisch ungünstige Dreifach-Eingriffe zu klassifizieren (s. Abb. 6).

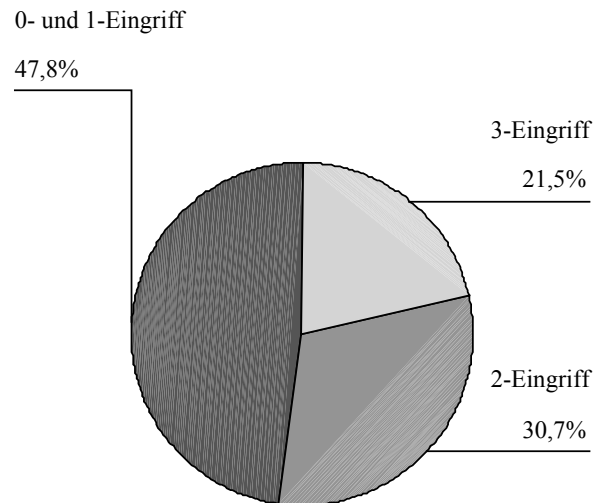


Abbildung 6: Deskriptive Strategieanalyse. Anteile von Null- und Einfach-, Zweifach- und Dreifach-Eingriffen beim Steuern von ÖKOSYSTEM ($n = 2240$ Takte).

5.4.4 Bearbeitungsdauer

Im Zusammenhang mit Hypothese 4 wurde weiterhin vermutet, dass Teilnehmer der Kausal-EG bei hypothesengeleitetem Vorgehen mehr Zeit zur Szenariobearbeitung benötigen als Vergleichsprobanden. Der Mittelwertsvergleich der Bearbeitungszeiten erbrachte allerdings ein uneindeutiges Muster. In einer 2×4 -ANOVA fand sich neben einem schwachen Haupteffekt für Vernetztheit ($F(1, 56) = 2.81, p = .099$) ein signifikanter Haupteffekt für den Faktor Wissensdiagnose ($F(3, 56) = 2.84, p = .046$) und ein auf dem 10 %-Niveau signifikanter Interaktionseffekt ($F(3, 56) = 2.33, p = .084$). Der Haupteffekt ist dabei keineswegs eindeutig auf eine höhere Bearbeitungsdauer der Kausal-EG gegenüber den anderen drei Treatment-Stufen zurückzuführen (a priori Kontrast: $p = .14$). Ohne Berücksichtigung des Vernetztheitsfaktors ist für die Kausal-EG gegenüber der UB-KG kein Unterschied in der Bearbeitungsdauer festzustellen (Tukey HSD: $p = 1.000$). Die genannten Gruppen unterscheiden sich im paarweisen post hoc Vergleich jeweils schwach signifikant von der OW-EG (Tukey HSD: Kausal-EG vs. OW-EG $p = .086$, UB-KG vs. OW-EG $p = .071$). Auffällig ist, dass die UB-KG als einzige Gruppe bei hoher Vernetztheit deutlich mehr Zeit benötigte verglichen mit der Bedingung niedriger Vernetztheit (s. Tabelle 3), was möglicherweise den Effekt für Vernetz-

heit und die signifikante Interaktion bedingt. Eine plausible Erklärung kann an dieser Stelle nicht gegeben werden. Bei einer Gruppengröße von $n = 8$ sind auch zufällige Abweichungen vom Mittel relativ wahrscheinlich.²²

Die vierte und zentrale Hypothese bezüglich der Reaktivität von Kausaldiagramm-Analysen konnte demnach nur im Kern bekräftigt werden, insofern als das „Produkt“ des Wissenserwerbsprozesses, das nachweisbare Strukturwissen, nach mehrfacher Bearbeitung von Kausaldiagrammen in der Kausal-EG erhöht ist. In Prozessmaßen, die einen erfolgreichen Wissenserwerb charakterisieren (Strategieeinsatz, Bearbeitungsdauer, im weiteren Sinn auch Oberflächenwissen), war die Kausal-EG höchstens tendenziell überlegen. Entgegen der Erwartung fiel auch die Steuerungsleistung in der Kausal-EG nicht signifikant höher als in den Vergleichsgruppen aus, auch wenn eine Tendenz in diese Richtung vorlag und der positive prädiktive Einfluss von Kausalwissen auf Steuerleistung als bestätigt angesehen werden kann.

5.5 Schulischer Erfolg, soziodemografische Variablen und Problemlöseerfolg

Laut Hypothese 5 sollte sich der Problemlöseerfolg u. a. aus allgemeinen und spezifisch mathematischen schulischen Leistungen vorhersagen lassen, d. h. in dieser Untersuchung aus den Indikatorvariablen Abiturnote und letzter Schulnote in Mathematik. Um auch die soziodemografischen Variablen in die Analyse einzubinden, wurden Regressionsmodelle für die fünf Prädiktoren Abiturnote, letzte Schulnote in Mathematik, Beruf, Geschlecht und Alter jeweils getrennt für die abhängigen Variablen GdK_{sum} und GdS aufgestellt. Die kategorialen Variablen Geschlecht und Beruf wurden dabei als dummy-kodierte Variablen berücksichtigt (Geschlecht: eine Dummy-Variable; Beruf: zwei Dummy-Variablen). Anzumerken bleibt, dass die Kategorie Beruf mit 50 Studen-

²²Eine einzelne Versuchsperson aus der UG-KG mit hoher Vernetztheit benötigte über 2500 ms, 3.9 Standardabweichungen über dem Gesamtmittelwert von 953.6 ms. Um den Effekt von Ausreißerwerten zu prüfen, wurde die Analyse zum Vergleich auch ohne diesen Datensatz durchgeführt. Dennoch blieb die mittlere Bearbeitungsdauer unter der genannten Treatment-Kombination unerwartet hoch. Auch die logarithmische Transformation der Bearbeitungszeiten veränderte das Verteilungsmuster nicht.

tinnen und Studenten im Diplomstudiengang Psychologie relativ einheitlich ausgeprägt war. Abweichungen von der Normalverteilungsannahme wurden für die Prädiktoren in Kauf genommen, da die zu erwartenden Verzerrungen laut Bortz (1999, S. 184) meist tolerabel sind.

Das Ausgangsmodell mit fünf Prädiktoren wurde über Rückwärts-Selektion (Ausschlusskriterium: $\alpha \geq .10$) auf ein sparsameres Modell mit ausschließlich signifikanten Prädiktoren reduziert. Es wurden in diesem Fall genau die Prädiktoren ins reduzierte Modell übernommen, die sich schon im Ausgangsmodell als signifikant erwiesen hatten, so dass die oft kritisierte Beliebigkeit schrittweiser Regressionsverfahren (vgl. Hays, S. 732) hier kaum ins Gewicht fallen dürfte. Die Tabellen 4 und 5 stellen nur die reduzierten Modelle dar. Weder die dummy-kodierten Berufsvariablen noch Abitur- oder Mathematiknote konnten GdS oder GdK_{sum} vorhersagen. Die Schulnoten korrelierten auch bei getrennter Betrachtung der Geschlechter nicht signifikant mit den Maßen der Problemlösegüte. Das Alter einer Versuchsperson steht in signifikant negativer prädiktiver Beziehung zu GdK_{sum} , jedoch in keiner systematischen Beziehung zur Steuerleistung GdS. Das Geschlecht kann – zu einem gewissen Anteil – sowohl GdK_{sum} als auch GdS vorhersagen. Die männlichen Teilnehmer waren den Teilnehmerinnen in beiden Gütemaßen überlegen bei einer signifikant höheren Varianz unter den männlichen Teilnehmern (s. auch Tabelle 6), ohne dass es signifikante Geschlechterunterschiede in Abiturnote, Mathematiknote oder im Strategieeinsatz gab. Allein auf der Emotionsskala zeigten Frauen einen höheren Wert für Ängstlichkeit vor der Szenariobearbeitung ($t(62) = -2.01, p = .048$).

Mit einer aufgeklärten Varianz (R^2) von rund 8 % im Ein-Prädiktoren-Modell für GdS und etwa 21 % im Zwei-Prädiktoren-Modell für GdK_{sum} ist die Vorhersagekraft durch demografische Einflüsse vergleichsweise gering. Die Varianzaufklärung lässt sich steigern, wenn man zusätzliche Prädiktoren aufnimmt, die aufgrund der berichteten Befunde in Betracht kommen: Das Modell für GdS wurde um eine Dummy-Variable für Vernetztheit und um die quantitative Variable GdK_{sum} erweitert und erreichte eine Varianzaufklärung von $R^2 = .577$ (korrigiert: $R_{adj} = .563$; s. Tabelle 7). Allerdings verfehlte die dummy-kodierte Geschlechtsvariable in diesem Modell die Signifikanz und wird nicht mehr aufgeführt. Die abhängige Variable GdK_{sum} ist zu gut 35 % prädizierbar, wenn man

Tabelle 4: Zwei Regressionsmodelle zur Vorhersage von GdK_{sum} . Modell GdK_{con} : nur Kontrollvariablen als Prädiktoren, Modell GdK_{con+ex} : Experimentalfaktoren und Kontrollvariablen als Prädiktoren.

Modell		b	SE_b	β	Signifikanz	
GdK_{con}					$F = 7.60$.001
	Geschlecht	.34	.100	.40	$t = 3.39$.001
	Alter	-.015	.007	-.28	$t = -2.36$.022
GdK_{con+ex}					$F = 8.21$	< .0001
	Geschlecht	.33	.088	.39	$t = 3.68$.001
	WiDi 1 ^a	.24	.088	.28	$t = 2.66$.010
	Vernetztheit	.21	.076	.29	$t = 2.77$.008
	Alter	-.013	.006	-.24	$t = -2,23$.029

^aFaktor Wissensdiagnose: Dummy-Prädiktor für Zugehörigkeit zur Kausal-EG

Tabelle 5: Zwei Regressionsmodelle zur Vorhersage von GdS. GdS wurde zur größeren Übersichtlichkeit umgepolt (hohe Werte entsprechen hoher Steuerleistung). Modell GdS_{con} : nur Kontrollvariablen als Prädiktoren, Modell GdS_{con+ex} : Experimentalfaktoren und Kontrollvariablen als Prädiktoren.

Modell		b	SE_b	β	Signifikanz	
GdS_{con}					$F = 1.62$.025
	Geschlecht	1.00	.435	.29	$t = 2.30$.025
GdS_{con+ex}					$F = 41.63$	< .0001
	GdK_{sum}	2.13	.355	.52	$t = 6.01$	< .0001
	Vernetztheit	1.27	.258	.43	$t = 4.92$	< .0001

Tabelle 6: Geschlechtereffekte für GdK_{sum} und GdS.

	Männer ($n = 16$)		Frauen ($n = 48$)		t -Test
	M	SD	M	SD	
	GdK_{sum}	.86	.41	.55	
GdS	3.86	1.79	4.75	1.34	$t(20.895) = -1.80, p = .086$

Tabelle 7: Varianzaufklärung durch die Regressionsmodelle für GdK_{sum} und GdS.

Modell	df	R^2	R^2_{adj}	SE_{R^2}
GdK _{con}	2	.208	.180	.332
GdK _{con+ex}	4	.357	.314	.304
GdS _{con}	1	.082	.067	1.464
GdS _{con+ex}	2	.577	.563	.993

im demografischen Modell zusätzlich die experimentellen Manipulationen Vernetztheit und Wissensdiagnose berücksichtigt. Von drei dummy-kodierten Variablen für Wissensdiagnose erscheint hier nur jene, die die Zugehörigkeit zur Kausal-EG kennzeichnet; übereinstimmend mit den Ergebnissen in Abschnitt 5.4.1 blieben Dummy-Prädiktoren für die Zugehörigkeit zur OW-EG oder UB-KG insignifikant.

Das Regressionsmodell bestätigt insofern die Befunde zum Einfluss von Vernetztheit und Wissensdiagnose auf die Problemlösegüte sowie den engen Zusammenhang zwischen GdK_{sum} und GdS. Soziodemografische Merkmale der Versuchspersonen haben demgegenüber einen schwächeren Effekt. Geschlechtereffekte sind zweifellos vorhanden, sollten aber die Interpretation anderer Ergebnisse nicht einschränken, da pro Treatment-Kombination das gleiche Geschlechterverhältnis vorlag. Entgegen der Erwartung fand sich kein positiver Zusammenhang zwischen Maßen der Problemlösegüte, Abitur- oder Mathematiknote, so dass Hypothese 5 nicht gestützt wurde.

5.6 Emotionen und Problemlösegüte

Um stimmung induzierte Einflüsse, die mit den beschriebenen Effekten interferieren, als Alternativerklärungen so weit wie möglich auszuschließen, sollten Probanden unterschiedlicher Treatment-Bedingungen in ihrer Eingangsmotivation und emotionalen Befindlichkeit vergleichbar sein. Zur Überprüfung herangezogen wurden mittlere Ausprägungen auf den fünf theoretisch bestimmten Dimensionen positive Stimmung, Ärger, Aufregung, Ängstlichkeit und Aktiviertheit (für die Zuordnung der Adjektiv-Items zu den Dimensionen s. S. 93 im Anhang). Diese bestehen allerdings aus ungleichen und

z. T. sehr geringen Itemzahlen. Von normalverteilten Emotionsindizes ist überwiegend nicht auszugehen.

Ein Acht-Gruppen-Vergleich per Kruskal-Wallis-Test erbrachte für keinen der fünf Emotionsindizes vor der Szenariobearbeitung signifikante oder ansatzweise signifikante Mittelwertsunterschiede. Ebenso fiel der Gruppenvergleich der Emotionsindizes nach dem Bearbeiten des Systems aus. Systematische Stimmungsunterschiede zwischen den Treatment-Kombinationen sind sowohl vor als auch nach der Problembearbeitung unwahrscheinlich, wenn eine Nichtsignifikanz in der non-parametrischen Analyse auch weniger aussagekräftig ist als in einer parametrischen ANOVA.

Eine Analyse der Emotionsentwicklung *während* des Problemlöseprozesses wurde für die Teilnehmer der UB-KG ($n = 16$) vorgenommen. Signifikante globale Unterschiede zwischen den sechs Befragungszeitpunkten zeigten sich im Friedman-Test für Ärger ($\chi^2(5) = 11.63$, $p_{asy} = .040$), Aufregung ($\chi^2(5) = 13.40$, $p_{asy} = .020$) und Aktiviertheit ($\chi^2(5) = 14.70$, $p_{asy} = .012$). Der Verlauf der fünf Emotionsdimensionen ist in Abb. 7 veranschaulicht: Positive Stimmung und Aktiviertheit (Interesse) sind bei einem theoretisch möglichen Wertebereich von 1 bis 5 mit dem Erwartungswert 3 über die ganze Versuchsdauer überdurchschnittlich vorhanden. Die Aktiviertheit nimmt mit dem Beginn der Computersimulation (Befragungszeitpunkt 2) deutlich zu, danach wieder ab. Zu Beginn mäßig ausgeprägt ist Aufregung, die im Verlauf des Versuchs fast kontinuierlich abnimmt. Ärger und Ängstlichkeit sind anfangs kaum vorhanden. Während Ängstlichkeit tendenziell weiter abnimmt, steigt Ärger in nachfolgenden Befragungen sichtbar an.

Über Korrelationsanalysen wurde – weitgehend explorativ – das Verhältnis zwischen emotionaler Befindlichkeit und Problemlöseerfolg untersucht. Von den eingangs erfassten Dimensionen korrelierten Aufregung ($r = .30$, $p = .015$) und tendenziell auch Aktiviertheit ($r = .24$, $r = .057$) signifikant positiv mit dem später erworbenen Strukturwissen GdK_{sum} . Ähnlich hoch korrelierten diese zwei Dimensionen mit der mittleren GdS (Aufregung: $r = .25$, $p = .047$; Aktiviertheit: $r = .35$, $p = .005$), während sich Korrelationskoeffizienten für die übrigen Dimensionen als nicht signifikant herausstellten. Auch die abschließend erfassten Emotionsindizes standen z. T. in systematischer Bezie-

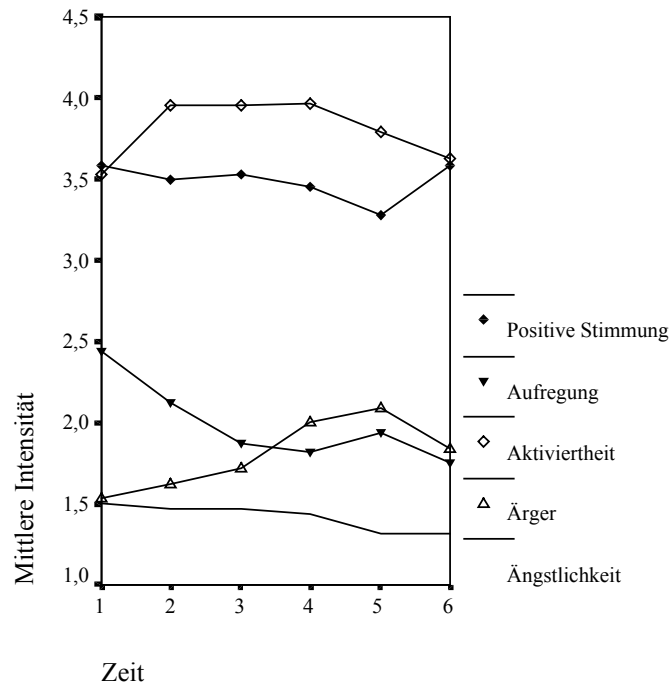


Abbildung 7: Emotionsindizes im Verlauf des Versuchs.

lung zum Problemlöseerfolg: Gute Stimmung korrelierte signifikant positiv, Ärger und Ängstlichkeit korrelierten signifikant negativ mit der durchschnittlichen GdS (positive Stimmung: $r = .31$, $p = .013$; Ärger: $r = -.27$, $p = .033$; Ängstlichkeit: $r = -.33$, $p = .008$).

6 Diskussion

Hauptziel der empirischen Arbeit war es zu untersuchen, ob und in welcher Weise sich die Bearbeitung von Kausaldiagramm-Analysen parallel zur Steuerung eines komplexen Problemlöseszenarios vom DYNAMIS-Typ auf Problemlösegüte und -verhalten auswirkt. Ausgehend von Kluwes (1988) Kritik zur Reaktivität dieses wissensdiagnostischen Verfahrens und empirischen Befunden bei Süß et al. (1993) wurde auf der Grundlage der Dual Space Theory von Klahr und Dunbar (1988) angenommen, dass prozessbegleitende Kausaldiagramm-Analysen zu einem überwiegend hypothesentestenden Vorgehen im Problemlöseprozess anregen. Indem sie zum Verbalisieren von kausalen Systemzusammenhängen aufgefordert werden, sollten Problemlöser systematisch eigene Hypothesen

generieren, die einen vergleichbar leistungssteigernden Effekt haben wie explizite, vom Versuchsleiter vorgegebene Hypothesen (vgl. Vollmeyer & Burns, 1996). Als primäres Ergebnis war daher für jene Probanden, die wiederholt Kausaldiagramme anfertigten, ein erhöhter Wissenserwerb gegenüber Kontrollprobanden zu erwarten.

Die Hypothese bezüglich Effekten auf den Wissenserwerb ließ sich bestätigen. Die Teilnehmer der Kausal-EG, die nach jedem von fünf kombinierten Steuer- und Explorationsdurchgängen ein Diagramm bearbeiteten, konnten nach dem abschließenden fünften Durchgang signifikant mehr und präziseres Systemwissen im Diagramm formalisieren als Teilnehmer, die – bei gleicher Erfahrung mit ÖKOSYSTEM – erstmalig nach dem fünften Durchgang ein Kausaldiagramm bearbeiteten. Die Beobachtung, dass sich speziell die Probanden der Kausal-EG von den drei relativ leistungshomogenen Vergleichsgruppen abhoben, spricht für einen Reaktivitätseffekt, der auf den besonderen Aufforderungscharakter der Kausaldiagramm-Abfrage zurückgeht. Die zunächst in Betracht gezogenen Alternativinterpretationen, dass ein oberflächlicher Rekognitionstest oder eine Unterbrechung des Problemlöseprozesses für eine inhaltlich nicht relevante Abfrage emotionaler Zustände vergleichbare förderliche Auswirkungen wie ein Kausaldiagramm hat, erscheinen kaum mehr plausibel.

Weniger eindeutig lässt sich die Frage beantworten, inwieweit die gesteigerte Systemidentifikation, d. h. das „Produkt“ des Problemlöseprozesses, bei Kausaldiagramm-Analysen tatsächlich durch eine verstärkt hypothesenorientierte Herangehensweise bedingt ist. Es wurde vermutet, dass selbstständiges Hypothesengenerieren und -prüfen verglichen mit einer Trial-and-Error-Strategie bzw. einer Suche im Experimenterraum mit einem höheren kognitiven und zeitlichen Aufwand verbunden ist, der sich in einer längeren Bearbeitungsdauer manifestiert. Weiterhin sollte der Strategiegebrauch bei hypothesengeleitetem Vorgehen systematischer sein, d. h. Null- und Einfacheingriffe (VOTAT) sollten gegenüber den weniger systematischen Doppel- und Dreifacheingriffen überwiegen, da gerade die Manipulation nur einer Variablen oder das Konstanthalten aller Variablen es erlaubt, gezielt Hypothesen über isolierte Variableneffekte zu testen. Keine dieser beiden Vermutungen konnte jedoch statistisch bekräftigt werden.

Der zusammenfassende Strategie-Indikator „Güte der Strategieanwendung“ deutet un-

abhängig von der experimentellen Bedingung auf ein relativ systematisches Explorieren hin. Möglicherweise waren die Probanden, zum großen Teil Studentinnen und Studenten der Psychologie, durch ihre methodische Ausbildung mit dem Konzept des systematischen Hypothesentestens so vertraut, dass sie es spontan auf die komplexe Problemlösesituation übertrugen. Eine nicht erfahrungswissenschaftlich geprägte Stichprobe (evtl. Schülerstichprobe) könnte eine größere Varianz im spontanen Strategieeinsatz aufweisen und damit eher geeignet sein, um experimentell bedingte Strategieeffekte aufzuzeigen. Das zweite Prozessmaß, die Netto-Bearbeitungsdauer, wies nur unter der Bedingung geringer Vernetztheit den erwarteten Unterschied zwischen den vier Wissensdiagnosegruppen auf. In der Gesamtstichprobe wie auch unter der Bedingung hoher Systemvernetztheit benötigten die Probanden der Emotions-Kontrollgruppe überraschend mindestens ebenso viel Zeit wie die Probanden der Kausal-EG. Es wäre nicht auszuschließen, dass die wiederholte Reflexion über emotionale Zustände die Aufmerksamkeit verstärkt auf die eigene Befindlichkeit hin- und von der Problemstellung ablenkt, was durch einen erhöhten Zeitbedarf kompensiert würde. Dies allerdings begründet nicht schlüssig, warum eine solche Verlangsamung ausschließlich bei hoher Systemkomplexität zu beobachten ist. Für die hier durchgeführte Untersuchung legen Einblicke in die Individualdaten eine Verzerrung durch Ausreißerwerte nahe. Unter dieser Annahme scheinen die Bearbeitungszeiten nicht durchweg valide interpretierbar. Die (nach wie vor plausible) Hypothese eines erhöhten Zeitbedarfs durch hypothesengeleitetes Vorgehen bei Kausaldiagrammen lässt sich so weder bestätigen noch eindeutig von der Hand weisen.

In einem weiterführenden Schritt wurde die Annahme aufgestellt und geprüft, dass Kausaldiagramm-Analysen nicht nur den Erwerb expliziten Strukturwissens fördern, sondern vermittelt über den Wissenserwerb auch die Steuerleistung. Konkret untersucht wurden das Verhältnis zwischen Systemidentifikation und -steuerung ebenso wie der vermutete Effekt von Kausaldiagramm-Analysen auf die Steuerleistung.

Für beide Leistungsindikatoren, Güte des Kausalwissens und Güte der Systemsteuerung, gilt dabei, dass die Leistung vom Grad der Systemerfahrung und vom Grad der Systemvernetztheit abhängig ist: Neben systematischen Leistungssteigerungen mit zunehmender Anzahl von Explorationsdurchgängen im anfangs völlig unbekanntem System war der

Problemlöseerfolg für die weniger komplexe, schwächer vernetzte ÖKOSYSTEM-Variante größer als für die stark vernetzte Variante. Die Ergebnisse von Funke (1985) zum Einfluss von Systemvernetztheit konnten damit (im Rahmen eines Zwei-Gruppen-Vergleichs) klar repliziert werden, so dass offensichtlich auch hier die formale Komplexität des Systems mit der subjektiven Komplexität der Problemanforderung übereinstimmt. Zugleich stellt die Replikation eine erfolgreiche Validitätsprüfung und günstige Voraussetzung für die anschließenden Analysen dar.

Ebenfalls interessant ist es, neben dem Einfluss von abstraktem Strukturwissen die Rolle von anwendungsnäherem, spezifischem Eingriffswissen im Problemlöseprozess zu bestimmen. Einen Versuch, Eingriffswissen zu diagnostizieren, stellt der Rekognitionstest dar. Um Vergleichbarkeit zwischen den experimentellen Bedingungen herzustellen, war es erforderlich, Rekognitionsabfragen gestaffelt nach jedem Durchgang vorzulegen. Damit sich die Steuersituation nicht vollständig wiederholte, beschränkte sich die Darstellung mit vier Rekognitionsitems bei sieben Eingriffsmöglichkeiten pro Durchgang auf eine stichprobenartige Auswahl. Diese ist trotz Einberechnung der Antwortsicherheiten in einen Wissensindex sehr sparsam gewählt und mag anfällig für Verzerrungen sein. Insbesondere für jene Probanden, die im Abschlussdurchgang alle Steuerziele exakt erreichten, stellte sich die Abfrage als überaus einfach heraus: Im Extremfall waren alle dargebotenen Targets und Distraktoren identisch; vermeintliche, vom Computer als Distraktoren deklarierte Items waren in Wirklichkeit Targets und mussten als solche ausgewertet werden. In anderen Fällen variierte die Schwierigkeit der Distraktoraufgaben stark: Es gab Items, deren formale Unmöglichkeit beinahe intuitiv erkennbar war, und „knifflige“ Distraktoren, die mit Target-Items fast vollkommen übereinstimmten. Eine Vereinheitlichung im Schwierigkeitsgrad wäre theoretisch durch einen strengeren Konstruktionsalgorithmus denkbar (z. B. ausschließlich Verwendung von Distraktorelementen aus benachbarten Takten, z. B. Verwendung von Distraktorzuständen, die sich nur in einem falschen Element vom Target unterscheiden).

Derartige methodische Beschränkungen in den Rekognitionstests könnten erklären, weshalb Oberflächenwissen hier (im Gegensatz zu den verwandten Maßen Strukturwissen und Steuerleistung) weder nachweislich von der Vernetztheit des Systems noch von der

Systemerfahrung abhing. Probanden der Oberflächenwissen-Experimentalgruppe zeigten keine systematischen Leistungsanstiege im Verlauf der fünf Testdurchgänge und waren den Vergleichspersonen in der abschließenden Abfrage keineswegs überlegen. Auch die Teilnehmer, deren Strukturwissen durch Kausaldiagramme unterstützt worden war, konnten vergangene Systemzustände nicht signifikant zuverlässiger wiedererkennen, obwohl das erhöhte Strukturwissen Plausibilitätseinschätzungen im Rekognitionstest erleichtern sollte.

Generell sinnvoll, wenn auch hier nicht durchführbar, sind größere Itemzahlen, die eine fehlende Standardisierung nivellieren. Preußler (2001) z.B. erfasste spezifisches Eingriffswissen nicht gestaffelt im Anschluss an einzelne Szenariodurchgänge, sondern in einem Block von 72 Items nach dem letzten von fünf Szenariodurchgängen. Fänden sich auch unter diesen Bedingungen keine statistischen Gruppenunterschiede, könnte man die Schlussfolgerung wagen, dass Probanden konkrete Steuersituationen insgesamt gut im Gedächtnis behalten und von Distraktorsituationen abgrenzen können, unabhängig von Übung und Szenarioschwierigkeit. Für die folgende Interpretation des Zusammenhangs von Wissen und Steuerleistung soll wegen der problematischen Operationalisierung allerdings Abstand vom spezifischen Eingriffswissen genommen werden. Als Wissensmaß wird ausschließlich das in Kausaldiagrammen erfasste Strukturwissen berücksichtigt.

Für deutliche positive Zusammenhänge zwischen Systemidentifikation und -steuerung sprechen sowohl die Korrelationsanalyse der manifesten Güteindikatoren als auch die pfadanalytisch ermittelte Korrelation zwischen angenommenen latenten Variablen. Anders als vermutet und bei Kluge (2003) aufgezeigt hatte die Vernetztheit keinen moderierenden Einfluss auf die Höhe des Zusammenhangs. Probanden konnten das erworbene Systemwissen offenbar bei einer relativ komplexen Problemstellung mindestens genauso gut anwenden wie bei einer einfacheren Problemstellung. Zum Einsatz der Pfadanalyse bleibt anzumerken, dass der Stichprobenumfang mit nur 64 Versuchspersonen die empfohlenen Richtwerte weit unterschreitet. Inhaltlich darf von einer Vorhersage der Steuerleistung durch erworbenes Strukturwissen nicht im zeitlich strikten Sinn gesprochen werden, da die Steuerleistung in vier aufeinander folgenden Steuer- und Explorationsdurchgängen gemessen wurde, der Wissenserwerb hingegen erst nach dem letzten

Steuerdurchgang zu einem einzigen Zeitpunkt. Dies mag ein Grund für die problematische globale Modellpassung sein.

Ausgehend vom gesteigerten Wissenserwerb bei Kausaldiagrammen und dem positiven Einfluss von Strukturwissen auf die Steuerleistung wäre zu erwarten gewesen, dass Probanden, die wiederholt Kausaldiagramme bearbeiten, auch Steuerziele besser erreichen als Vergleichspersonen. Diese Hypothese ließ sich aber statistisch nicht bestätigen. Selbst wenn die kausal-instruierte Gruppe eine tendenzielle, bei größerer Stichprobe eventuell signifikante Überlegenheit im Steuern zeigte, dürfte der Effekt gering ausfallen gegenüber dem Effekt auf den Wissenserwerb. Wie kann man diese Dissoziation erklären?

Eine Interpretation, die Beckmann (1994) für einen vergleichbaren Befund in Betracht zieht, geht davon aus, dass ein wissenserwerbsbasierter Steuermechanismus wie im Pfadmodell beschrieben nicht der einzig denkbare Mechanismus ist. Insbesondere Personen, die nur unzureichendes Strukturwissen erwerben, können nach Beckmann ein System alternativ *ad hoc* steuern. Interessanterweise bringt Beckmann dieses zweite Steuerprinzip mit semantisch eingekleideten Systemen in Verbindung. Es ist also möglicherweise kein Zufall, dass die in Abschnitt 3.1 vorgestellten hohen, positiven Korrelationen zwischen Wissens- und Steuergüte überwiegend für vorwissensneutrale Szenarios ermittelt wurden. Eine *ad hoc* Steuerung könnte für die vorliegende Untersuchung ansatzweise erklären, weshalb sich die pfadanalytische Modellpassung als kritisch herausstellte und weshalb die Umsetzung von Wissen in Steuerleistung hier zwar substantiell, aber nicht maximal ist.

Während Beckmann (1994) von „systemunspezifischen Heuristiken“ spricht, wäre für ÖKOSYSTEM auch heuristisches Operieren aufgrund von semantischen Vorerwartungen (Weltwissen) möglich. Genauere Voruntersuchungen und Analysen einzelner Teilrelationen hinsichtlich Erwartungskonformität wären sicherlich aufschlussreich gewesen. Möglicherweise gab es im verwendeten ÖKOSYSTEM nur eine einzige, offensichtlich erwartungsdiskonforme Relation (Einsatz von Gift senkt die Wasserverschmutzung), so dass heuristisches Problemlösen, basierend auf ungeprüften Vorannahmen, einen Teil des Steuererfolgs mitzuverantworten hat. Die überwiegend positive Ausgangsstimmung der Probanden sollte nach Fiedler (1988, 2000) mit dem vermuteten intuitiven Problem-

lösestil vereinbar sein, wenn auch keine nachweisbaren Effekte positiver Stimmung auf die Problemlösegröße vorlagen.

Die explorativ ermittelten Zusammenhänge zwischen Stimmungs- und Problemlösemaßen reflektieren in erster Linie bekannte und alltagsnahe Annahmen: Hoher Problemlöseerfolg ging mit hohem anfänglichem Interesse und hoher Konzentration einher ebenso wie mit einem erhöhten Aufregungsniveau bei insgesamt geringer Aufregung. Ähnliche positive Effekte von anstrengungsfreier Konzentration auf den Umgang mit einem DYNAMIS-Szenario haben z.B. Vollmeyer und Rheinberg (1998) aufgezeigt. Angaben über die emotionale Befindlichkeit unmittelbar nach dem Experiment lassen vermuten, dass die Teilnehmer ihre Problemlöseleistung überwiegend realistisch einschätzten, wobei sich die Einschätzung auf die aktuelle Stimmung auswirkte: Erfolgreiche Probanden waren besser gestimmt, weniger ängstlich und weniger ärgerlich. Die sinkende Aktiviertheit und Aufregung im Verlauf der Szenariobearbeitung deutet gleichermaßen auf Ermüdung und zunehmende Vertrautheit im Umgang mit dem anfangs völlig unbekanntem ÖKOSYSTEM hin. Dass die Ärgerlichkeit der Teilnehmer in der Unterbrechungs-Kontrollgruppe von Messzeitpunkt zu Messzeitpunkt anstieg, kann mit der empfundenen Langwierigkeit der Problemstellung zusammenhängen oder mit wahrgenommenem Misserfolg. Alternativ kann der Ärger eine Reaktion auf die wiederkehrende Unterbrechung der Problemlöseaktivität sein (vgl. S. 41). Welche der Erklärungen zutreffender ist, lässt sich ohne Vergleichsdaten nicht entscheiden.

Entgegen der Erwartung konnten weder schulische Gesamtleistungen der Probanden noch Mathematikleistungen den Problemlöseerfolg vorhersagen. Die nicht signifikanten Zusammenhänge zwischen Abiturdurchschnitt und Problemlösegröße sollten nahelegen, dass die Bewältigung eines DYNAMIS-Problems andere, stärker praxisorientierte Fähigkeiten und Fertigkeiten erfordert als jene, die mit guten akademischen Leistungen der gymnasialen Oberstufe einhergehen (vgl. Dörner, 1986). Interpretiert man die Abiturnote als Intelligenzmaß, fügt sich die vorliegende Untersuchung in eine Reihe von Studien ein, die den immer wieder postulierten und empirisch aufgezeigten Einfluss von Intelligenz beim komplexen Problemlösen (vgl. Kröner, 2001) nicht klar stützen konnten (s. z. B. Funke, 1985). Die Analyse von Abiturnoten ist allerdings mit methodischen Schwie-

rigkeiten verbunden, da gleiche Abiturdurchschnittsnoten durch unterschiedlichste Spezialbegabungen zustande kommen können, die im unstandardisierten Gesamtwert nicht berücksichtigt werden. In der vorliegenden Stichprobe, die sich weitgehend aus Studentinnen und Studenten eines Numerus-Clausus-Fachs zusammensetzte, wies nur etwa ein Viertel der Probanden eine Abiturnote, die schlechter als 2.0, auf. Es ist von einer relativ geringen Varianz bei insgesamt hohem schulischen Leistungsniveau auszugehen; d. h. die statistischen Voraussetzungen für Korrelationsanalysen waren keineswegs optimal.

Da mathematische Schulleistungen in keinem signifikanten positiven Zusammenhang zur Problemlösegüte standen, sind auch diese Konstrukte offensichtlich weniger eng verwandt als z. B. von Beckmann (1994, S. 66) angenommen. Interessant für ergänzende Studien wäre es zweifellos zu untersuchen, ob elementare Rechenfertigkeiten (Kopfrechnen), die im Oberstufen-Mathematikunterricht von untergeordneter Bedeutung sind, eventuell höher mit dem Problemlöseerfolg korrelieren oder ob komplexes Problemlösen mit DYNAMIS generell ein strategisch-exploratives Moment erfordert (vgl. Schoppek, 2002), das über mathematisch-analytisches Denken hinausgeht.

Unter den soziodemografischen Kontrollvariablen stellte sich v. a. das Geschlecht der Probanden als einflussreich heraus. Männliche Teilnehmer waren den Frauen sowohl im Wissenserwerb als auch im Steuern überlegen, unabhängig von ihren Mathematik- oder Abiturnoten. Geschlechterunterschiede in der allgemeinen kognitiven Leistungsfähigkeit und in schulischen mathematischen Fertigkeiten sind also unwahrscheinlich. Es ist jedoch nicht auszuschließen, dass speziell die numerische Verarbeitungskapazität bei den Männern in der Stichprobe erhöht war (vgl. Wagener, 2001) und den Problemlöseerfolg begünstigte. Ebenfalls möglich wäre eine leistungsmindernde geringere Erfolgserwartung bei den Frauen (vgl. Rustemeyer & Jubel, 1996). Ein Indiz dafür könnte die stärker ausgeprägte Ängstlichkeit der Frauen zu Beginn der Problemstellung sein.

Fazit. Der unterstützende Einfluss von Kausaldiagramm-Analysen auf den Erwerb von Strukturwissen deutet darauf hin, dass spontan gebildete, subjektive Kausalmodelle als mentale Repräsentationen beim komplexen Problemlösen weniger zwingend sind als im ursprünglichen DYNAMIS-Konzept angenommen (vgl. Funke, 1985). Bei einer Mehrheit

der Problemlöser scheinen präzisierte subjektive Kausalmodelle erst dann verfügbar zu sein, wenn die Problemstellung vorsieht, den Aufbau des Systems formal zu analysieren und zu benennen. Konkrete Steuerziele werden andernfalls möglicherweise mit einem Minimum abstrakter Repräsentation angestrebt, ergänzt um alternative Problemrepräsentationen oder experimentegeteilte ad hoc Maßnahmen.

Wie eingangs angedeutet steht die kognitive Prozessforschung vor der Schwierigkeit, komplexe, mentale Repräsentationen in ihrer Entstehung beschreiben und erklären zu wollen. Sogenannte non-reaktive Verfahren setzen in diesem Bemühen auf der Ebene basaler Wahrnehmung (z. B. Blickbewegungsanalysen) oder bei einfachen Verarbeitungsprozessen (z. B. Reaktionszeitanalysen) an. Beobachtungen auf dieser Ebene können jedoch bestenfalls Aufschluss über Vorstufen artikulierbarer, mentaler Repräsentationen geben. Die interessierenden, vollständigen Repräsentationen bleiben der Beobachtung durch sprachlich oder symbolisch vermittelte Methoden wie Kausaldiagramm-Analysen vorbehalten.

Oben bemerkt wurde, dass die meisten Problemlöser abstraktes Wissen über kausale Systemzusammenhänge nur nach entsprechender Aufforderung benennen. Nicht bemerkt wurde, dass die meisten Problemlöser auch konkrete Steuerziele nur nach entsprechender Aufforderung zu erreichen versuchen. Es wäre sicherlich nicht nur überspitzt, sondern auch unlogisch, die Steueranforderung als reaktiven Einfluss auf einen spontan ablaufenden Problemlöseprozess zu interpretieren, liegt doch gerade in der Interaktion zwischen Anforderung und Person ein wesentliches Merkmal komplexen Problemlösens (s. Frensch & Funke, 1995). Übertragen auf die Bedeutung von Kausaldiagramm-Analysen und ihre Reaktivität auf den Wissenserwerb lässt das Argument eine durchaus konstruktive Schlussfolgerung für die Forschungspraxis zu: Kausaldiagramm-Analysen sind *mehr* als nur wissensdiagnostische Instrumente; sie sind ein integrativer Bestandteil der Problemstellung mit hoher diagnostischer Aussagekraft.

Danksagung

Mein abschließender Dank gilt allen, die zum Entstehen dieser Arbeit beigetragen haben: Herrn Prof. Dr. Joachim Funke für seine umfassende Betreuung, Carsten Schulz, der die JAVA-programmierte ÖKOSYSTEM-Version zur Verfügung stellte, Frau Lammarsch, die das Szenario auf den Rechnern des Psychologischen Instituts lauffähig machte und mir wertvolle Hinweise zur Grafikformatierung in L^AT_EXgab. Meinem Bruder Jan Olaf danke ich für spezielle Anpassungen von ÖKOSYSTEM an die aktuelle Fragestellung. Nicht zuletzt danke ich allen Versuchsteilnehmerinnen und -teilnehmern für ihre Kooperation, ihre Geduld und das Interesse an der Untersuchung.

Literaturverzeichnis

- Amelang, M. & Bartussek, D. (2001). *Differentielle Psychologie und Persönlichkeitsforschung*. Stuttgart: Kohlhammer.
- Beckmann, J. F. (1994). *Lernen und komplexes Problemlösen. Ein Beitrag zur Konstruktvalidierung von Lerntests*. Bonn: Holos.
- Beckmann, J. F. & Guthke, J. (1995). Complex problem solving, intelligence and learning ability. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 177-200). Hillsdale, NJ: Erlbaum.
- Berry, D. C. & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology*, *36A*, 209-231.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Heidelberg: Springer.
- Brehmer, B. (1995). Feedback delays in complex dynamic decision tasks. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 103-130). Hillsdale, NJ: Erlbaum.
- Brehmer, B. & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, *9*, 171-184.
- Buchner, A. (1995). Basic topics and approaches to the study of complex problem solving. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 27-63). Hillsdale, NJ: Erlbaum.
- Buchner, A. & Funke, J. (1993). Finite-state automata: Dynamic task environments in problem-solving research. *Quarterly Journal of Experimental Psychology*, *46A*, 83-118.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences (rev. ed.)*. New York: Academic Press.
- Dienes, Z. & Fahey, R. (1995). Role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 848-862.

- Dörner, D. (1976). *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.
- Dörner, D. (1981). Über die Schwierigkeiten menschlichen Umgangs mit Komplexität. *Psychologische Rundschau*, 32, 163-179.
- Dörner, D. (1986). Diagnostik der operativen Intelligenz. *Diagnostica*, 32, 290-308.
- Dörner, D. (1992). Über die Philosophie der Verwendung von Mikrowelten oder „Computerszenarios“ in der psychologischen Forschung. In H. Gundlach (Ed.), *Psychologische Forschung und Methode: Das Versprechen des Experiments. Festschrift für Werner Trausel* (pp. 53-87). Passau: Passavia-Universitäts-Verlag.
- Dörner, D., Kreuzig, H. W., Reither, F. & Stäudel, T. (1983). LOHHAUSEN. *Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Huber.
- Dörner, D., Schaub, H. Stäudel, T. & Strohschneider, S. (1988). Ein System zur Handlungsregulation oder – Die Interaktion von Emotion, Kognition und Motivation. *Sprache & Kognition*, 7, 217-232.
- Dörner, D. & Wearing, A. J. (1995). Complex problem solving: Toward a (computer-simulated) theory. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 65-99). Hillsdale, NJ: Erlbaum.
- Erdfelder, E., Faul, F. & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments & Computers*, 28, 1-11.
- Fiedler, K. (1988). Emotional mood, cognitive style, and behavior regulation. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition, and social behavior* (pp. 100-119). Göttingen: Hogrefe.
- Fiedler, K. (2000). Toward an integrative account of affect and cognitive phenomena using the BIAS computer algorithm. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 223-252). Cambridge: Cambridge University Press.
- Frensch, P. A. & Funke, J. (Eds.). (1995). *Complex problem solving: The European perspective*. Hillsdale, NJ: Erlbaum.
- Funke, J. (1983). Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Prädiktor? *Diagnostica*, 29, 283-302.

- Funke, J. (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle. *Zeitschrift für Psychologie*, 193, 435-457.
- Funke, J. (1992a). Dealing with dynamic systems: Research strategy, diagnostic approach and experimental results. *German Journal of Psychology*, 16(1), 24-43.
- Funke, J. (1992b). *Wissen über dynamische Systeme: Erwerb, Repräsentation und Anwendung*. Berlin: Springer.
- Funke, J. (1993). Microworlds based on linear equation systems: A new approach to complex problem solving and experimental results. In G. Strube & K.-F. Wender (Eds.), *The cognitive psychology of knowledge* (pp. 313-330). Amsterdam: Elsevier Science Publishers.
- Funke, J. (2003). *Problemlösendes Denken*. Stuttgart: Kohlhammer.
- Funke, J. & Buchner, A. (1992). Finite Automaten als Instrumente für die Analyse von wissensgeleiteten Problemlöseprozessen: Vorstellung eines neuen Untersuchungsparadigmas. *Sprache und Kognition*, 11, 27-37.
- Funke, J. & Müller, H. (1988). Eingreifen und Prognostizieren als Determinanten von Systemidentifikation und Systemsteuerung. *Sprache & Kognition*, 7, 176-186.
- Goshen-Gottstein, Y. & Moscovitch, M. (1995). Repetition priming for newly formed and preexisting associations: Perceptual and conceptual influences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1229-1248.
- Hair, J. F., Anderson, R. E., Tatham, R. L. & Black, W. (1998). *Multivariate Data Analysis*. Upper Saddle River, NJ: Prentice Hall.
- Hays, W. L. (1991). *Statistics*. Fort Worth, TX: Harcourt Brace.
- Hübner, R. (1988). Die kognitive Regelung dynamischer Systeme und der Einfluss analoger versus digitaler Informationsdarbietung. *Zeitschrift für Psychologie*, 196, 161-170.
- Jäger, A. O. (1982). Mehrmodale Klassifikation von Intelligenzleistungen: Experimentell kontrollierte Weiterentwicklung eines deskriptiven Intelligenzstrukturmodells. *Diagnostica*, 28, 195-225.
- Janke, W. & Debus, G. (1978). *Die Eigenschaftswörterliste EWL. Eine mehrdimensionale*

nale Methode zur Beschreibung von Aspekten des Befindens. Göttingen: Hogrefe.

Klahr, D. & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12, 1-55.

Kluge, A. (2003). *Problemlöseleistungen und Szenarioschwierigkeit: Über Wissenserwerb für das Steuern technischer Systeme.* Unveröffentlichte Habilitationsschrift, RWTH, Aachen.

Kluwe, R. H. (1988). Methoden der Psychologie zur Gewinnung von Daten über menschliches Wissen. In H. Mandl & H. Spada (Eds.), *Wissenspsychologie* (pp. 359-385). München: Psychologie Verlags Union.

Kröner, S. (2001). *Intelligenzdiagnostik per Computersimulation.* Münster: Waxmann.

Logan, G. (1988). Towards an instance theory of automatization. *Psychological Review*, 95, 492-527.

Mandler, G. (1964). The interruption of behavior. In D. Levine (Ed.), *Nebraska symposium on motivation, 1964* (pp. 163-219). Lincoln, NE: University of Nebraska Press.

Mandler, G. (1975). *Mind and emotion.* New York: Wiley.

Müller, H. (1993). *Komplexes Problemlösen: Reliabilität und Wissen.* Bonn: Holos.

Nickerson, R. S. (1998). The Teaching of Thinking and Problem Solving. In R. J. Sternberg (Ed.), *Thinking and Problem Solving.* London: Academic Press.

Niemivirta, M. (2002). Motivation and performance in context: The influence of goal orientation and instructional setting on situational appraisals and task performance. *Psychologia*, 45, 250-270.

Preußler, W. (1996). Zur Rolle expliziten und impliziten Wissens bei der Steuerung dynamischer Systeme. *Zeitschrift für Experimentelle Psychologie*, 43, 399-434.

Preußler, W. (1997). Effekte des Kontexts auf den Wissenserwerb bei der Steuerung eines dynamischen Systems. *Sprache & Kognition*, 16, 48-59.

Preußler, W. (1998). Strukturwissen als Voraussetzung für die Steuerung komplexer dynamischer Systeme. *Zeitschrift für Experimentelle Psychologie*, 45, 218-240.

- Preußler, W. (2001). Problemlöseerfolg als Bedingung für den Erwerb von Eingriffswissen bei der Systemsteuerung. *Zeitschrift für Experimentelle Psychologie*, 48, 3, 214-225
- Putz-Osterloh, W. (1981). Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg. *Zeitschrift für Psychologie*, 189, 79-100.
- Putz-Osterloh, W. (1993a). Strategies for knowledge acquisition and transfer of knowledge in dynamic tasks. In G. Strube & K.-F. Wender (Eds.), *The cognitive psychology of knowledge* (pp. 331-350). Amsterdam: Elsevier Science Publishers.
- Putz-Osterloh, W. (1993b). Unterschiede im Erwerb und in der Reichweite des Wissens bei der Steuerung eines dynamischen Systems. *Zeitschrift für Experimentelle und angewandte Psychologie*, 40, 386-410.
- 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.
- Rinne, H. & Specht, K. (2002). *Zeitreihen. Statistische Modellierung, Schätzung und Prognose*. München: Vahlen.
- Rollett, W. (2003). *Strategieinsatz, Informationsnutzung und Motivation beim Explorieren und Steuern komplexer Systeme*. Unveröffentlichte Dissertation, Humanwissenschaftliche Fakultät der Universität Potsdam, Potsdam.
- Rustemeyer, R. & Jubel, A. (1996). Geschlechtsspezifische Unterschiede im Unterrichtsfach Mathematik hinsichtlich der Fähigkeitseinschätzung, Leistungserwartung, Attributionen sowie im Lernaufwand und im Interesse. *Zeitschrift für Pädagogische Psychologie*, 10, 13-25.
- Schaub, H. & Strohschneider, S. (1992). Die Auswirkungen unterschiedlicher Problemlöseerfahrungen auf den Umgang mit einem unbekanntem komplexen Problem. *Zeitschrift für Arbeits- und Organisationspsychologie*, 36, 117-126.
- Schoppek, W. (2002). Examples, rules, and strategies in the control of dynamic systems. *Cognitive Science Quarterly*, 2, 63-92.
- Schulz, C. A. (2003). *Strategien zum komplexen Problemlösen. Eine experimentelle Ana-*

lyse der Rolle von Analogien beim Wissenserwerb in komplexen, dynamischen Systemen. Unveröffentlichte Diplomarbeit, Psychologisches Institut der Universität, Heidelberg.

Spering, M. (2001). *Emotionen und Kontrollüberzeugungen beim komplexen Problemlösen: eine experimentelle Untersuchung anhand des computersimulierten Problemlösenszenarios FSYS 2.0*. Unveröffentlichte Diplomarbeit, Psychologisches Institut der Universität, Heidelberg.

Sternberg, R. J. (1995). Expertise in complex problem solving: A comparison of alternative conceptions. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 295-321). Hillsdale, NJ: Erlbaum.

Süß, H., Kersting, M. & Oberauer, K. (1993). Zur Vorhersage von Steuerungsleistungen an computersimulierten Systemen durch Wissen und Intelligenz. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 14, 189-203.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, 12, 257-285.

Velden, M. (1982). *Die Signalentdeckungstheorie in der Psychologie*. Stuttgart: Kohlhammer.

Vollmeyer, R. & Burns, B. D. (1996). Hypotheseninstruktion und Zielspezifität: Bedingungen, die das Erlernen und Kontrollieren eines komplexen Systems beeinflussen. *Zeitschrift für Experimentelle Psychologie*, 43, 657-683.

Vollmeyer, R., Burns, B. D. & Holyoak, K. J. (1996). The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognitive Science*, 20, 75-100.

Vollmeyer, R. & Rheinberg, F. (1998). Motivationale Einflüsse auf Erwerb und Anwendung von Wissen in einem computersimulierten System. *Zeitschrift für Pädagogische Psychologie*, 12, 11-23.

Wagener, D. (2001). *Psychologische Diagnostik mit komplexen Szenarios*. Lengerich: Pabst.

Werner, J. (1997). *Lineare Statistik – Allgemeines Lineares Modell*. Weinheim: Psychologie Verlags Union.

Anhang

Anhang A: Materialien zur Versuchsdurchführung

Instruktionstext zum Szenario ÖKOSYSTEM

Liebe Teilnehmerin, lieber Teilnehmer,

vielen herzlichen Dank schon vorab für Ihre Teilnahme an der Untersuchung zum ÖKOSYSTEM!

Für die nächsten 45-60 Minuten werden Sie in die Rolle eines Kleingärtners/einer Kleingärtnerin schlüpfen – selbstverständlich nur virtuell in einer Computersimulation.

Stellen Sie sich bitte Ihren Kleingarten vor, in dem sich ein einziger Baum befindet.

- Der Baum hat eine bestimmte Anzahl von Blättern;
- auf dem Baum sitzen eine bestimmte Anzahl von Käfern;
- die Grundwasserverschmutzung in diesem Ökosystem hat einen bestimmten Betrag.

Als Kleingärtner/in ist es Ihre Aufgabe, den Baum zu „pflegen“ bzw. möglichst gut zu versorgen. Von Woche zu Woche haben Sie die Möglichkeit, Pflegemaßnahmen durchzuführen und so Einfluss auf den Zustand Ihres Baums zu nehmen. Um die Blätterzahl Ihres Baums, die Käferzahl und die Grundwasserverschmutzung zu beeinflussen, können Sie auf drei Arten in das Ökosystem eingreifen: Sie können

- Gift geben,
- Schädlingsfresser auf den Baum setzen und
- Dünger geben.

Es muss jedoch nicht immer nötig sein, dass Sie alle drei Maßnahmen auf einmal durchführen. Genauso können Sie in einer Woche nur eine oder zwei Maßnahmen aufgreifen oder ganz auf einen Eingriff verzichten.

Sie haben immerhin fünf mal sieben Wochen Zeit, um den Kleingarten kennen zu lernen und zu versorgen.

Ziel ist es, den folgenden Zustand in Ihrem Kleingarten zu erreichen:

- Der Baum soll **10.000 Blätter** haben.
- Auf dem Baum sollen sich **100 Käfer** befinden.
- Die **Grundwasserverschmutzung** soll **0** sein.

Blätterzahl:	10.000
Käfer:	100
Wasserverschmutzung:	0

Ideal ist es, wenn Sie diese Werte exakt und so früh wie möglich erreichen. Aber auch Annäherungen an die Zielwerte sind schon ein gutes Ergebnis – Sie wissen schließlich zunächst nicht, wie sich Ihre Pflegemaßnahmen auf den Kleingarten auswirken.

Nehmen Sie sich Zeit zu beobachten, wie sich Gift, Schädlingsfresser und Dünger auf Blätterzahl, Käfer und Wasserverschmutzung auswirken. Achten Sie auch darauf, ob es Abhängigkeiten zwischen der Blätterzahl, der Käferzahl und der Wasserverschmutzung gibt. Wichtig ist, dass Sie sich **nicht auf eine Vorstellung verlassen, wie etwas wirken sollte**, sondern beobachten, was geschieht. Es können durchaus Wirkungen eintreten, die in einem realen Ökosystem nicht zu erwarten wären.

Ein Tipp: Alle Maßnahmen haben nur Auswirkungen auf den Zustand der unmittelbar folgenden Woche.

Zuletzt noch einige technische Hinweise, wie Sie am Computer Einfluss auf Ihren Kleingarten nehmen:

Für jede Woche zeigt der Computer drei Eingabefelder an: Ein Feld für „Grif“, eines für „Schädlingsefresser“, eines für „Dünger“. In diese Felder können Sie Ihre gewünschten Maßnahmen als Zahlenwerte entweder direkt über die Tastatur eingeben, oder Sie können den Scrollbalken unter dem Feld dazu benutzen.
Ein leer gelassenes Feld entspricht einem Wert von null.

Die Veränderungen treten ein, sobald Sie auf den „OK“-Button drücken. Der Computer zeigt Ihnen dann, was aus den Maßnahmen geworden ist. Es beginnt eine neue Woche, für die Sie die nächsten Maßnahmen auswählen.

Anders als in einem realen Ökosystem kann es passieren, dass auch negative Zahlen erscheinen, z. B. eine Käferzahl von -500 , eine Wasserverschmutzung von -10 , eine Blätterzahl von -700 . Lassen Sie sich dadurch nicht verwirren; negative Werte sind eine Eigenart Ihres virtuellen Kleingartens.

Genauso können auch Sie „negative Maßnahmen“ treffen, z. B. könnten Sie 10 Einheiten Dünger hinzugeben (10) oder 10 Einheiten wegnehmen (-10).

Nach der achten Woche beginnt jeweils ein neuer Durchgang. Die Maßnahmen, die Sie in den früheren Durchgängen getroffen haben, wirken sich hier nicht mehr aus. Jeder Durchgang ist unabhängig vom vorherigen und hat dieselben Startwerte.

Alles klar?

Wenn Sie noch Fragen haben, wenden Sie sich bitte an die Versuchsführung.
Sie können diesen Text bei Bedarf jederzeit nachlesen und sich z. B. die Zielvorgaben in Erinnerung rufen. Wenn Sie möchten, können Sie sich auch Notizen auf einem zusätzlichen Blatt machen.

Viel Vergnügen!

Vorlagebogen zur Kausaldiagramm-Analyse

Vpn-no.: _____
Durchgang: _____

Käfer		
Wasserverschmutzung		
Blätterzahl		
Gift		
Schädlingfresser		
Dünger		

Instruktion zur Kausaldiagramm-Analyse

Bitte erinnern Sie sich an den gerade gelösten Durchgang.

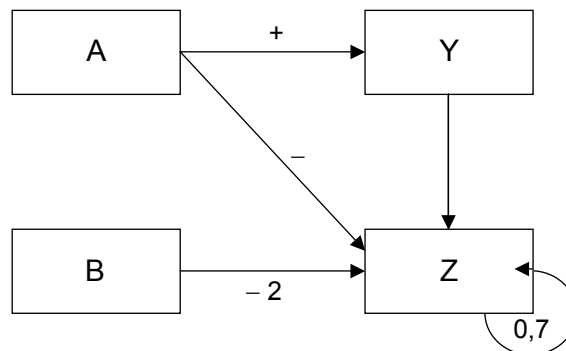
Vielleicht haben Sie Vermutungen darüber aufgestellt, in welcher Weise die Käferzahl, die Wasserverschmutzung und die Blätterzahl (die Variablen Ihres Ökosystems) mit anderen Variablen Ihres Ökosystems (Gift, Schädlingsfresser, Dünger) zusammenhängen bzw. wie Käferzahl, Wasserverschmutzung und Blätterzahl auch untereinander zusammenhängen könnten.

Es geht um **kausale Zusammenhänge**, d. h. Fragen wie:

- Welche Auswirkungen hat eine (Ursachen-)Variable eine andere Variable?
- Welche Ursache(n) hat eine Wirkung, die Sie bei einer bestimmten Variable beobachtet haben?

Ich möchte Sie bitten, Ihre Vermutungen über Kausalzusammenhänge als Pfeildiagramm in den separaten Vordruck einzutragen.

Hier ist ein Beispiel, wie ein solches Pfeildiagramm – ganz formal mit abstrakten Buchstabenymbolen – aussehen kann:



Sie können Vermutungen über Kausalzusammenhänge auf unterschiedliche Weise in Ihr Pfeildiagramm eintragen, je nachdem wie genau Ihre Vermutungen sind. Es gibt drei Arten, Zusammenhänge einzuzeichnen:

- a) Angabe der Relation
 - b) Angabe der Richtung der Relation
 - c) Angabe der Stärke und Richtung der Relation
- a) Relation

Sie können zwei Kästchen mit Variablennamen durch einen einfachen Pfeil miteinander verbinden. Zeichnen Sie einen einfachen Pfeil dann ein, wenn Sie vermuten, dass zwei Variablen voneinander abhängen ohne dass Sie über die Art des Zusammenhangs etwas Genaues wissen.

Beispiel:



„Y bedingt Z.“

„Z hängt von Y ab.“

„Wenn sich Y verändert, verändert sich Z.“

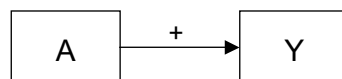
b) Richtung der Relation

Sie können zwei Kästchen mit Variablennamen durch einen Pfeil miteinander verbinden und die Pfeile mit einem Plus- oder Minuszeichen beschriften. Tragen Sie ein Plus- oder Minuszeichen neben einem Pfeil ein, wenn Sie vermuten, dass zwei Variablen positiv bzw. negativ voneinander abhängen.

Ein positiver Zusammenhang (+) zwischen zwei Variablen bedeutet: Beide Variablen verändern sich in der gleichen Richtung. Entweder erhöht sich der Betrag bei beiden Variablen oder er verringert sich bei beiden Variablen.

Ein negativer Zusammenhang (–) zwischen zwei Variablen bedeutet: Die beiden Variablen verändern sich in entgegengesetzten Richtungen. Erhöht sich der Betrag bei einer Variable, verringert er sich bei der anderen Variable; verringert sich der Betrag bei der einen Variable, erhöht er sich bei der anderen Variable.

Beispiele:

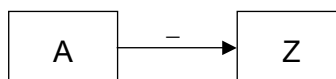


„A bedingt Y positiv.“

„Y hängt positiv von A ab.“

„Wenn sich A erhöht, erhöht sich Y.“

„Wenn sich A verringert, verringert sich Y.“



„A bedingt Z negativ.“

„Z hängt negativ von A ab.“

„Wenn sich A erhöht, verringert sich Z.“

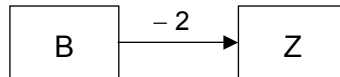
„Wenn sich A verringert, erhöht sich Z.“

c) Richtung und Stärke der Relation

Sie können zwei Kästchen mit Variablennamen durch einen Pfeil miteinander verbinden und die Pfeile mit einem Wirkfaktor, d. h. mit einer positiven oder negativen Zahl beschriften. Tragen Sie eine positive oder negative Zahl neben dem Pfeil ein, wenn Sie vermuten, dass diese Zahl einen exakten mathematischen Zusammenhang zwischen zwei Variablen beschreibt.

Ein exakter mathematischer Zusammenhang zwischen zwei Variablen bedeutet: Die eine Variable verändert sich um einen bestimmten Betrag, die andere Variable ändert sich um diesen Betrag multipliziert mit dem Wirkfaktor.

Beispiel:



„B bedingt Z mit einem Wirkfaktor von $- 2$.“

„Z hängt von B mit einem Wirkfaktor von $- 2$ ab.“

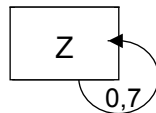
„Wenn sich B um den Betrag x erhöht, verringert sich Z um das Doppelte von x ($- 2x$).“

„Wenn sich B um den Betrag x verringert, erhöht sich Z um das Doppelte von x ($+ 2x$).“

- d) Und noch eine Besonderheit: Sie können auch ein Kästchen mit Variablenamen durch einen kreisförmigen Pfeil „mit sich selbst verbinden“, wenn Sie vermuten, dass eine Variable sich „von selbst“ ändert ohne dass andere Variablen Einfluss haben. Auch hier können Sie den Pfeil mit einem Wirkfaktor für einen exakten mathematischen Zusammenhang beschriften.

Ein exakter mathematischer Zusammenhang bedeutet hier: Die Variable hat in einer Woche einen bestimmten Betrag, in der Folgewoche hat die Variable diesen Betrag multipliziert mit dem Wirkfaktor.

Beispiel:



„Z hängt mit einem Wirkfaktor von $0,7$ von sich selbst ab.“

„Wenn Z in einer Woche den Betrag x hat, hat Z in der Folgewoche den Betrag $0,7 \cdot x$.“

„Z verringert sich in der Folgewoche auf 70% vom Betrag der Vorwoche.“

Wenn Sie Ihre Vermutungen über Zusammenhänge in Ihrem Ökosystem nicht so exakt angeben können wie in den letzten Beispielen, ist das nicht weiter tragisch!

Versuchen Sie aber auf jeden Fall, alle Vermutungen in irgendeiner Weise im Pfeildiagramm wiederzugeben, wenn Sie möchten auch ergänzend mit schriftlichen Kommentaren am Rand oder auf der Rückseite.

Fragebogen zur Erfassung von Emotion und Motivation

Vpn-no.: _____

Wie fühlen Sie sich *jetzt*?

Bitte beschreiben Sie Ihren Gefühlszustand mit Hilfe der vorgegebenen Wörter. Ist ein Gefühl bei Ihnen gerade nicht vorhanden, kreuzen Sie den Kreis ganz links an. Ist ein Gefühl vorhanden, kreuzen Sie einen Kreis weiter rechts an, umso weiter rechts, je stärker das Gefühl ist, das Sie gerade erleben. Ist ein Gefühl bei Ihnen gerade sehr stark vorhanden, kreuzen Sie den Kreis ganz rechts an.

traurig	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
unbesorgt	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
aufgeregt	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
interessiert	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
verärgert	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
nervös	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
geistesabwesend	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
ratlos	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
konzentriert	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
zufrieden	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
müde	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
ängstlich	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
zaghaf	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
pessimistisch	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
gereizt	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
fröhlich	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
gelangweilt	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
aktiv	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
heiter	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden
glücklich	nicht vorhanden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr stark vorhanden

Bildschirm-Instruktionstext zum Rekognitionstest

„Im Folgenden sehen Sie einige Zustände aus dem gerade gelösten Durchgang noch einmal.

Der Ausgangszustand wird Ihnen vermutlich bekannt vorkommen.

Er entspricht den Werten für Käfer, Wasserverschmutzung und Blätterzahl, wie sie Ihnen in einem der sieben Takte angezeigt wurden.

Die Werte für Gift, Schädlingsfresser und Dünger entsprechen den Werten, die Sie selbst in diesem Takt eingegeben haben, um Einfluss auf den Ausgangszustand zu nehmen.

Neben dem Ausgangszustand zeigt der Bildschirm einen Folgezustand, ebenfalls mit Werten für Käfer, Wasserverschmutzung und Blätterzahl.

Diese Werte können als Folge des Ausgangszustands eingetreten sein, müssen es aber nicht.

Sie selbst sollen sich erinnern und entscheiden, ob der Folgezustand zum Ausgangszustand passt oder nicht.

Achten Sie dabei auf alle drei Angaben: Käfer, Wasserverschmutzung und Blätterzahl.“

Anhang B: Materialien zur Versuchsauswertung

Lokale Anpassungsgütekriterien des Pfadmodells auf S. 58

Indikatorreliabilität		<i>t</i> -Wert		
S	.439	S	← I	-5.000 * **
GdK _{rel}	.706	GdK _{rel}	← I	
GdK _{vor}	.831	GdK _{vor}	← I	8.014 * **
GdK _{vor}	.501	GdK _{vor}	← I	6.168 * **
GdS2	.458	GdS2	← S	6.045 * **
GdS3	.701	GdS3	← S	8.325 * **
GdS4	.826	GdS4	← S	9.399 * **
GdS5	.736	GdS5	← S	

Klassifikation der Emotions-Adjektive nach Dimensionen

Dimension	Adjektiv-Item ^a
Positive Stimmung	unbesorgt glücklich heiter fröhlich zufrieden traurig (-) pessimistisch (-) ratlos (-)
Ärger	gereizt verärgert
Aufregung	aufgeregt nervös
Ängstlichkeit	ängstlich zaghaf
Aktiviertheit	interessiert aktiv konzentriert gelangweilt (-) müde (-) geistesabwesend (-)

^aNegativ gepolte Items sind durch „-“ gekennzeichnet.

Signalentdeckungstheoretische Auswertung der Daten der OW-EG

Die im Haupttext auf S. 60 vorgestellte Analyse von Oberflächenwissen nach der Signalentdeckungstheorie wurde folgendermaßen durchgeführt:

1. Datengrundlage: Je 20 Antworten auf Rekognitionsitems für Probanden der OW-EG ($n = 16$)
2. Berechnung von individuellen Trefferraten(TR) und Falschalarmraten(FR)

$$TR = \frac{n(\text{Treffer})}{n(\text{Treffer}) + n(\text{Verpasser})}$$

$$FR = \frac{n(\text{Falsche Alarme})}{n(\text{Falsche Alarme}) + n(\text{Korrekte Zurückweisungen})}$$

Bei einzelnen Versuchspersonen fanden sich Extremwerte von null oder eins für Treffer- oder Falschalarmsraten. Nach Stanislaw und Todorov (1999, S. 144) sind diese zu transformieren.

3. Transformierte individuelle Trefferraten (TR') und Falschalarmsraten (FR')

Es wurde folgende, von Stanislaw und Todorov (1999) generell empfohlene Transformation durchgeführt:

$$TR' = \frac{n(\text{Treffer}) + .5}{n(\text{Treffer}) + n(\text{Verpasser}) + 1}$$

$$FR' = \frac{n(\text{Falsche Alarmer}) + .5}{n(\text{Falsche Alarmer}) + n(\text{Korrekte Zurückweisungen}) + 1}$$

4. Schätzung der signalentdeckungstheoretischen Parameter d' und β für jede Versuchsperson (nach Stanislaw & Todorov, 1999):

$$d' = \text{PROBIT}(TR') - \text{PROBIT}(FR')$$

$$\beta = \text{EXP} \left(\frac{\text{PROBIT}(FR')^2 - \text{PROBIT}(TR')^2}{2} \right)$$

Verteilung von d' und β in der OW-EG:

	<i>min</i>	<i>max</i>	<i>M</i>	<i>SD</i>
d'	.18	2.77	.97	.69
β	.38	1.57	.96	.35

Die empirischen Verteilungen für d' und β in der OW-EG weichen nicht signifikant von der Annahme einer Normalverteilung ab (K-S: $p < .20$).

5. Statistischer Mittelwertsvergleich für d' und β auf zwei Vernetztheitsstufen:

	schwach vernetzt ($n = 8$)		stark vernetzt ($n = 8$)		<i>t</i> -Test
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
d'	1.21	.83	.74	.46	$t(14) = 1.42, p = .178$
β	.87	.39	1.05	.29	$t(14) = -1.04, p = .317$

Literatur:

Stanislaw, H. & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments & Computers*, 31(1), 137-149.

Erklärung

Hiermit erkläre ich, dass ich diese Arbeit selbstständig und ohne Benutzung von anderen als den angegebenen Quellen und Hilfsmitteln angefertigt habe. Alle den Quellen wörtlich und inhaltlich entnommenen Stellen wurden als solche kenntlich gemacht.

Leimen, 15. Februar 2006

Christine Blech

Freiwillige Erklärung

Ich stimme ausdrücklich zu, dass meine durch Herrn Prof. Dr. Joachim Funke und Herrn Dr. Jörg Zumbach betreute Diplomarbeit mit dem Titel:

Die Reaktivität von Kausaldiagramm-Analysen beim komplexen Problemlösen.

Eine experimentelle Untersuchung anhand des komplexen, dynamischen Systems ÖKOSYSTEM.

nach Beendigung der Diplomprüfung wissenschaftlichen Zwecken zugänglich gemacht und in der Institutsbibliothek aufgestellt wird (Veröffentlichungen nach § 6 Abs. 1 UrhG), sowie hieraus im Rahmen des § 51 UrhG zitiert werden kann.

Sämtliche Verwertungsrechte nach § 15 UrhG verbleiben bei der Verfasserin der Diplomarbeit.

Leimen, 15. Februar 2006

Christine Blech

Christine Blech

Sudetenweg 8
69181 Leimen