

# Zur Reaktivität von Kausaldiagramm-Analysen beim komplexen Problemlösen

Christine Blech und Joachim Funke

Universität Heidelberg

**Zusammenfassung.** Die Produktion von Kausaldiagrammen gilt als eine Methode zur Wissensdiagnostik beim Bearbeiten komplexer Probleme. Ein Experiment sollte untersuchen, inwieweit Kausaldiagramm-Analysen den Umgang mit dynamischen Systemen beeinflussen. Es wurde vermutet, dass prozessbegleitende Kausaldiagramm-Analysen einen hypothesentestenden Problemlösestil sensu Klahr und Dunbar (1988) unterstützen und intensivieren. Als beobachtbare Konsequenz wurden ein erhöhter Erwerb von Strukturwissen und gesteigerte Leistungen im Steuern des Systems erwartet. 64 studentische Versuchspersonen bearbeiteten das Szenario „Ökosystem“ über fünf Durchgänge. Personen, die nach jedem Durchgang ein Kausaldiagramm anfertigten, zeigten sich im abschließend erworbenen Strukturwissen Vergleichspersonen überlegen, die einen oberflächlichen Rekognitionstest, eine nicht szenariobezogene Aufgabe oder gar keine zusätzliche Aufgabe absolviert hatten. Dies spricht für einen Reaktivitätseffekt. Allerdings benötigten kausal-instruierte Personen weder mehr Zeit noch explorierten sie das System geschickter. Auch im Steuern des Systems profitierten sie nicht von ihrem Wissensvorteil, sondern erbrachten Leistungen auf dem Niveau der Kontrollprobanden. Die abschließende Diskussion sieht in der Reaktivität des Instruments einen diagnostischen Gewinn.

Schlüsselwörter: Problemlösen, Wissen, Diagnose, System, reaktive Messungen

On the reactivity of causal diagrams in complex problem solving

**Abstract.** Producing causal diagrams during a dynamic problem solving task is a diagnostic measure of structural knowledge. For our experiment, we hypothesised that causal diagrams would affect complex problem solving, supporting and intensifying cognitive processes guided by self-generated hypotheses sensu Klahr and Dunbar (1988). In consequence, enhanced structural knowledge and increased control performance should occur. 64 students explored the computer simulation „ecosystem“ during five cycles, trying to reach goal values. Subjects confronted with causal diagrams after each cycle finally revealed deeper structural knowledge than 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, a likely effect of reactivity. The so-instructed subjects, however, did not spend more time on the task nor did they explore the system more skillfully or benefit from their enhanced knowledge when controlling the system. The final discussion stresses the value of causal diagrams.

Key words: problem solving, knowledge, diagnosis, system, reactive measures

Die Frage nach einem empirischen Zugang zu Denkinhalten und Denkvorgängen ist keine genuine Frage der modernen Problemlöseforschung. Dennoch hat die Frage gerade hier – im Rahmen der informationstheoretischen Forschungsperspektive – einen besonderen Stellenwert erhalten: Computersimulierte Szenarios, Nachbildungen realer, komplexer Probleme in der virtuellen Laborwelt, eignen sich hervorragend, um Bearbeitungsprozesse, etwa eine schrittweise Annäherung an einen vorgegebenen Zielzustand, sichtbar zu machen. Keinen unmittelbaren Einblick erlauben sie in das, was sich im Kopf einer problemlösenden Person abspielt: Welche innere, mentale Repräsentation hat eine Person von einer Problemsituation und wie verändert sich die Repräsentation im Verlauf des Lösungsprozesses?

Ein offensichtlich praktikabler Ansatz besteht darin, Personen nach definierten Lösungsabschnitten zu unterbrechen und über ihre kognitiven Repräsentationen, d. h. über ihr Wissen bezüglich der Problemsituation zu befragen. Üblicherweise erfolgt eine solche prozessbegleitende Wissensdiagnostik nach standardisierten Vorgaben, die sich sowohl am Untersuchungsmaterial als auch an Vorannahmen der Wissenschaftler und Wissenschaftlerinnen orientieren. Unabhängig von konkreten Operationalisierungen lässt sich dabei jedoch nicht mit Sicherheit entscheiden: Ist das erfasste Wissen tatsächlich das Abbild einer inneren, spontan generierten und verfügbaren, mentalen Repräsentation? Oder ist es erst der Einsatz einer wissensdiagnostischen Abfrage, der als externer Stimulus zum Nachdenken und zum Ausbilden einer bestimmten, gedanklichen Problemrepräsentation anregt? Im

letzteren Fall würde es sich um ein reaktives Messverfahren handeln.

Der vorliegende Beitrag untersucht die Problematik reaktiver, wissensdiagnostischer Verfahren exemplarisch anhand von Kausaldiagramm-Analysen, die in Verbindung mit komplexen Problemlösenszenarios zum Einsatz kommen. Zunächst geben wir einen kurzen Überblick über den Forschungsansatz und die Methode der Kausaldiagramm-Analyse, bevor theoretische Vermutungen zur Reaktivität dargelegt und experimentell geprüft werden.

## DYNAMIS und die Methode der Kausaldiagramm-Analyse

Erstmals von Funke (1985; vgl. auch Funke, 1993) beschrieben und entwickelt, entstand DYNAMIS als Rahmenprogramm, um computerbasierte, realitätsnahe, komplexe, dynamische Systeme zu erstellen, ohne auf die Vorteile mathematisch exakter Beschreibbarkeit zu verzichten: Ein formaler Algorithmus, der jedem DYNAMIS-Szenario zu Grunde liegt, erlaubt es, Befunde aus Studien an unterschiedlichen Systemen bzw. Teilsystemen zu vergleichen. Ebenso lassen sich spezielle Systemeigenschaften, etwa Eigendynamiken oder Nebenwirkungen gezielt manipulieren, um ihre Auswirkungen auf den Problemlöseprozess aufzuzeigen (z. B. Funke, 1993). Dass sich der Einsatz von DYNAMIS-Szenarios nicht in der Untersuchung formaler Systemeigenschaften erschöpft, zeigen vielfältige empirische Studien mit diesem Ansatz etwa zum Einfluss des Darbietungsformats (z. B. Beckmann, 1994), des Strategieeinsatzes (z. B. Schoppek, 2002; Vollmeyer, Burns & Holyoak, 1996), der Motivation (z. B. Niemivirta, 2002; Vollmeyer & Rheinberg, 1998) und der Intelligenz (Kröner, 2001) auf Variablen des komplexen Problemlösens.

Das mathematische Kernstück des DYNAMIS-Prinzips bilden multivariate, in der Praxis meist deterministische, autoregressive (AR) Prozesse erster oder seltener zweiter Ordnung. Mit Hilfe der AR-Prozesse werden lineare Abhängigkeiten zwischen den konstitutiven Variablen eines Systems numerisch exakt definiert. Jeder aktuelle Zustand einer endogenen Variable (Outputvariable) wird als Linearkombination gewichteter Vorgängerkonstrukte modelliert, wobei sowohl Zustände der betreffenden endogenen Variable selbst als auch Zustände exogener Variablen (Inputvariablen) oder anderer endogener Variablen in die Systemgleichungen eingehen. Anders als bei den weniger strukturierten, so genannten realitätsnahen Szenarios ist eine inhaltlich-semantische Einkleidung des Systems gegenüber der formalen Struktur sekundär (zur Klassifikation computersimulierter Szenarios s. z. B. Buchner, 1995; Frensch & Funke, 1995). Erst nach der Festlegung abstrakter Regelbeziehungen können Systemvariablen mit inhaltlich relevanten, vorwissensassoziierten Begriffen belegt werden.

Die Anforderungen an den Problemlöser sind bei DYNAMIS in zweierlei Hinsicht klar definiert: Ziel der Be-

arbeitung ist es, das komplexe, dynamische System zu steuern, d. h. durch Manipulation der exogenen Variablen in einen zuvor benannten Zielzustand zu überführen. Auf Grund der mathematisch exakten Systemdefinition existiert für einen beliebigen (Zwischen-)Zustand im Problemlöseprozess eine eindeutig bestimmbare, optimale, zielführende Intervention. Das Kriterium erfolgreicher Systemsteuerung ist ein möglichst geringer Abstand zwischen tatsächlich erreichten Ist-Zustand und angestrebtem Sollzustand (vgl. Müller, 1993).

Die zweite Anforderung – eigentlich eine Voraussetzung für die erfolgreiche Steuerung – besteht im Identifizieren der Systemstruktur. In einem typischen DYNAMIS-Experiment werden Probanden aufgefordert, nach jedem von mehreren Explorations- oder Steuerzyklen ihr bislang erworbenes Wissen zu formalisieren: Was konnte eine Person über die Funktionsweise des Systems herausfinden, welche Zusammenhänge zwischen welchen Variablen vermutet die Person? Neben anderen Abfrageformaten wie der „Paar-Aufgabe“ (Preußler, 1996, 1997, 1998) oder der lexikalischen Entscheidungsaufgabe zur Erfassung impliziten Systemwissens (Preußler, 1996) hat sich für explizites Systemwissen die Methode der *Kausaldiagramm-Analyse* als besonders differenziertes, diagnostisches Verfahren etabliert.

Den Versuchspersonen wird hierbei ein lückenhaftes Strukturdiagramm zur Vervollständigung vorgelegt. Ein Vorlagebogen enthält die Namen aller Systemvariablen einbeschrieben in einer Anordnung von Rechtecken, jedoch ohne Angaben bezüglich Relationen zwischen den Variablen. Zu ergänzen sind daher – je nach individuellem Kenntnisstand einer Versuchsperson – Pfeile, Vorzeichen und Zahlen, die symbolisieren, welche Systemrelationen eine Versuchsperson vermutet. Ein einfacher Pfeil kennzeichnet das Vorliegen einer Systemrelation ohne weitere Spezifikationen über die Art dieser Relation; zusätzliche positive bzw. negative Vorzeichen geben an, dass ein gleichsinniger bzw. gegensinniger Zusammenhang erkannt wurde (z. B. „je mehr – desto mehr“; „je mehr – desto weniger“); eine Zahl neben einem Pfeil steht für eine numerisch exakte Aussage über die Stärke des Zusammenhangs zwischen zwei Variablen. Selbstverständlich dürfen Versuchspersonen die drei Symboltypen miteinander kombinieren, um ihr zusammenhängendes Wissen über das System so präzise wie möglich zu artikulieren. Eine ideale Lösung würde ein vollständig bestimmtes Strukturdiagramm wie in Abb. 1 – das in dieser Arbeit verwendete Szenario – darstellen, das später als Bewertungsmaß herangezogen wird.

Strukturwissen anhand von Kausaldiagrammen lässt sich somit nicht nur als zusammenfassender Globalwert, sondern differenziert auf drei Genauigkeitsstufen ermitteln: als Relationswissen (Angabe einer vorliegenden Relation), als Richtungs- oder Vorzeichenwissen (Angabe eines korrekten Vorzeichens) und als Stärkewissen oder numerisches Wissen (Angabe eines korrekten numerischen Gewichts). Numerisches Wissen setzt Vorzeichen- und Relationswissen, Vorzeichenwissen setzt Relationswissen logisch voraus.

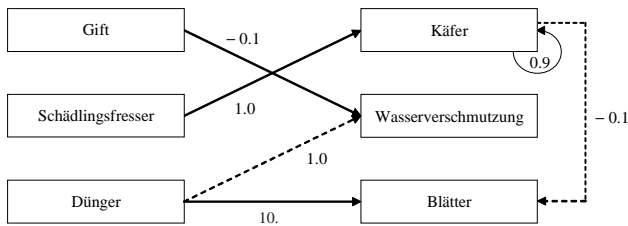


Abbildung 1. Struktur des DYNAMIS-Szenarios „Ökosystem“. Version mit vier Teilprozessen (geringe Vernetztheit): nur durchgezogene Linien. Version mit sechs Teilprozessen (erhöhte Vernetztheit): durchgezogene und gestrichelte Linien. Die Zahlen an den Kanten geben die Stärke der Verbindung an, die Vorzeichen deren Richtung.

Trotz ihrer Gebräuchlichkeit und offensichtlicher Vorzüge hat die Methode der Kausaldiagramm-Analyse theoretisch fundierte Kritik herausgefordert: Nicht allein der – unbestreitbar vorhandene – intellektuelle Anspruch der Diagramm ist nach Kluwe (1988) ein Schwachpunkt des Verfahrens; insbesondere könnten „Fragen nach bestimmten Informationen den Versuchspersonen nahe legen [...], bestimmte Komponenten der Aufgabenstellung oder Sachverhalte besonders zu beachten“ (S. 370/ 371). Die Vorgabe von Kausaldiagrammen parallel zum Explorieren eines DYNAMIS-Szenarios würde die Aufmerksamkeit der Teilnehmer demnach in erhöhtem Maß auf kausale Abhängigkeiten und auf die Vernetztheit des Systems lenken. Während Funke (1985) annimmt, dass die externe Wissensdiagnose einer internen, mentalen Repräsentation im selben Format (subjektives Kausalmodell) entspricht, führt Kluwe erfolgreich bearbeitete Kausaldiagramme auf Reaktivitäts- oder Demand-Effekte zurück.

Empirische Anhaltspunkte für derartige Effekte von Wissensdiagnostik beim komplexen Problemlösen sind selten, aber vorhanden. Eine Untersuchung von Süß, Kersting und Oberauer (1993) mit dem komplexen Szenario „Schneiderwerkstatt“ belegt einen schwachen Reaktivitätseffekt für einen prozessbegleitenden, verbalen Wissenstest: Probanden, die den Test zu Beginn des Experiments absolviert hatten, steuerten das System „Schneiderwerkstatt“ nachfolgend zunächst erfolgreicher als Kontrollpersonen, denen man keinen Test vorgab. Der Test könne die Teilnehmer „zu Reflexion und Hypothesenbildung“ angeregt haben, vermuten die Autoren (Süß et al., 1993, S. 196).

## Fragestellung und Hypothesen

Es stellte sich also die Frage, ob und inwieweit praxisrelevante Reaktivitätseffekte auch beim Einsatz von Kausaldiagramm-Analysen in DYNAMIS-Experimenten auftreten. Ausgehend von Kluwes (1988) Kritik nahmen wir an, dass Kausaldiagramm-Analysen ebenso wie der verbale Wissenstest bei Süß et al. (1993) die Reflexion und Hypothesenbildung im Problemlöseprozess folgendermaßen unterstützen: Für Versuchspersonen, die bereits nach un-

vollständiger Systemexploration und -identifikation abstraktes Systemwissen artikulieren sollen (wie es bei Kausaldiagramm-Analysen typischerweise der Fall ist), dürfte zumindest ein Teil der Angaben im Kausaldiagramm hypothetischen Charakter haben. Es handelt sich um zwar unsichere, aber explizite und mental leicht zugängliche Festlegungen. Diese könnten das nachfolgende Explorationsverhalten lenken, indem sie zur Grundlage systematischen Testens werden. Nach der Terminologie von Klahr und Dunbar (1988) wäre eine verstärkte Suche im Hypothesenraum anzunehmen.

Dass hypothesengeleitetes Vorgehen im Gegensatz zu experimentegeleiteten Trial-and-error-Explorationsstrategien mit größerem Erfolg in komplexen Problemsituationen einhergeht, zeigt ein DYNAMIS-Experiment von Vollmeyer und Burns (1996) zum Einfluss expliziter Hypotheseninstruktionen: Die Autoren fanden heraus, dass Versuchspersonen in Wissenserwerb und Systemsteuerung dann besonders gut abschnitten, wenn sie parallel zum Erkunden und Steuern eines Szenarios überprüfen sollten, ob eine vorgegebene Hypothese zutrifft oder nicht. Die Hypothese enthielt Angaben über einen spezifischen, kausalen Systemzusammenhang, ähnelte also den Angaben einer Versuchsperson in einem Kausaldiagramm.

Vor diesem Hintergrund ist zu erwarten, dass mögliche interne, selbstgenerierte Hypothesen beim Erstellen von Kausaldiagrammen einen vergleichbar leistungssteigernden Einfluss haben wie fremdgenerierte, extern vorgegebene Hypothesen in der Untersuchung von Vollmeyer und Burns (1996). Als primären Reaktivitätseffekt prozessbegleitender Kausaldiagramme vermuten wir daher einen gesteigerten Erwerb von kausalem Strukturwissen, begleitet durch zwei Prozesscharakteristika, die auf eine tiefe, hypothesengeleitete Verarbeitung schließen lassen: erhöhte Bearbeitungsdauer und systematischeren Strategieinsatz. In Übereinstimmung mit einer Mehrzahl früherer Studien, die positive Zusammenhänge zwischen dem erworbenen Strukturwissen und dem Steuern komplexer, dynamischer Systeme belegen (z. B. Beckmann & Guthke, 1995; Funke & Müller, 1988; Kröner, 2001; Müller, 1993; Vollmeyer & Rheinberg, 1998), interpretieren wir die Steuerung eines Systems in erster Linie als Anwendung erworbenen Wissens. So ist als Folgeeffekt („sekundärer Reaktivitätseffekt“) zu erwarten, dass auch Steuerziele exakter erreicht werden, wenn Problemlöser Kausaldiagramme bearbeiten.

Eine weitere Fragestellung betrifft den Einfluss von Systemvernetztheit auf den Problemlöseerfolg. Nach Arbeiten von Funke (1985) bzw. Kluge (2004) ist davon auszugehen, dass die formale, technische Komplexität eines Systems mit der subjektiv empfundenen, psychologischen Komplexität zusammenfällt. Im Sinne einer Replikation prüfen wir die Hypothese, dass stärker vernetzte Systeme mit erhöhtem kognitivem Aufwand, erschwerterem Wissenserwerb und eingeschränkter Wissensanwendung einhergehen. Auch ein eventueller, moderierender Einfluss der Systemvernetztheit auf die vermuteten Reaktivitätseffekte ist zu prüfen.

## Methode

*Design:* Zwei Hauptanforderungen stellten sich an das Versuchsdesign: Zum einen sollten generelle Reaktivitätseffekte aufgedeckt werden können, zum anderen sollte das Design Rückschlüsse darauf zulassen, ob auftretende Effekte spezifisch für die Methode der Kausaldiagramm-Analysen sind. Möglich wäre es etwa, dass *jegliche* Art prozessbegleitender Wissensdiagnostik reaktiv wirkt (vgl. Beckmann, 1994, S. 73) und im selben Maß zu hypothesengeleiteten Problemlösen anregt. Im äußersten Fall könnte schon die bloße Unterbrechung eines Problemlöseprozesses ein abstrahierendes, hypothesengeleitetes Vorgehen erzwingen, wenn konkrete Lernepisoden nach der Unterbrechung nicht mehr hinreichend verfügbar sind (vgl. Mandler, 1990) – auch diese Interpretation wäre, wenngleich weniger plausibel, nicht prinzipiell auszuschließen.

Neben einer Kausaldiagramme bearbeitenden Experimentalgruppe (Kausalwissen-EG) und einer Kontrollgruppe ohne wissensdiagnostische Tests (Baseline-KG) wurde daher eine Experimentalgruppe mit alternativem, auf spezifisches Eingriffswissen ausgerichteten Wissenstest (Oberflächenwissen-EG) und eine Kontrollgruppe mit szenario-irrelevanter Distraktoraufgabe (Unterbrechungs-KG) ins Design aufgenommen. Spezifisches Eingriffswissen umfasst assoziative Verknüpfungen von niedrigem Abstraktionsgrad und entspricht somit weitgehend den neueren Auffassungen von impliziten Wissensrepräsentationen (vgl. Frensch & Rünger, 2003; Shanks & Johnstone, 1998). Der vierstufige Faktor *Wissensdiagnose* wurde im vollständig gekreuzten 4×2-ANOVA-Design mit einem zweifach gestuften Faktor *Systemvernetztheit* kombiniert.

*Materialien:* Bei dem von uns verwendeten Szenario handelte es sich um eine überarbeitete, Java-programmierte Version des DYNAMIS-Szenarios „Ökosystem“ von Funke (1985). Das System umfasst drei exogene Variablen (Gift, Schädlingsfresser, Dünger) und drei endogene Variablen (Käfer, Wasserverschmutzung, Blätter), die semantisch in einen ökologischen Kontext eingebunden sind. Lineare autoregressive Abhängigkeiten erster Ordnung bestehen zwischen den Variablen, wie im Strukturdiagramm (Abb. 1) angegeben. Abhängig vom Grad der Vernetztheit variiert die Anzahl der Systemrelationen: vier Teilprozesse (einschließlich einer Eigendynamik) liegen unter der Bedingung geringer Vernetztheit vor, sechs Teilprozesse (einschließlich einer Eigendynamik und einer Nebenwirkung) finden sich unter der Bedingung stärkerer Vernetztheit. Die Zustände exogener und endogener Variablen werden den Versuchspersonen als Zahlenwerte in tabellarischer Darstellung präsentiert, wobei der aktuelle und maximal vier von insgesamt sieben Vorgängerzuständen simultan am Bildschirm sichtbar sind. Versuchspersonen nehmen Eingriffe im System per Tastatur oder Scrollbalken durch Eingabe ganzzahliger Werte der Input-Variablen vor.

Die Abfrage von Strukturwissen erfolgte nach der Methode der Kausaldiagramm-Analysen (Funke, 1985)

wie oben beschrieben. Die Namen der sechs inhaltlichen Systemvariablen wurden auf einem Vorlagebogen vorgegeben, die Systemrelationen waren handschriftlich zu vervollständigen. Zur Diagnose spezifischen Eingriffswissens wurde ein computerbasierter Rekognitionstest (vgl. Preußler, 2001) konstruiert. Der Rekognitionstest legt den Probanden eine zufällige Auswahl von Systemzuständen des zuvor bearbeiteten Szenariodurchgangs vor und verlangt die Entscheidung, ob es sich um eine bekannte, tatsächlich aufgetretene Szenariosituation (Target) oder um eine unbekannte, nicht aufgetretene Situation (Distraktor) handelt. Targets und Distraktoren werden in zufälliger Abfolge mit gleicher Wahrscheinlichkeit dargeboten. Ähnlichkeiten zwischen den beiden Itemklassen entstehen dadurch, dass in Distraktorsituationen lediglich einzelne Zustandselemente durch frühere oder spätere Elemente derselben Variable ausgetauscht werden. Zusätzlich zur dichotomen Entscheidung „Target vs. Distraktor“ geben die Probanden auf einer zehnstufigen Skala von 1 (sehr unsicher) bis 10 (sehr sicher) ihre Antwortsicherheit an.

Als nicht-szenariobezogene Distraktoraufgabe wurde ein kurzer Emotionsfragebogen zur aktuellen, persönlichen Befindlichkeit gewählt. Der Bogen enthält 20 Adjektiv-Items nach der Eigenschaftswörterliste von Janke und Debus (1978), bei wiederholter Vorgabe in der Unterbrechungs-EG in variierender Reihenfolge.

*Stichprobe:* 64 (16 männliche, 48 weibliche) Versuchspersonen nahmen an dem Experiment teil (durchschnittliches Alter: 25.7 Jahre), acht Probanden pro Bedingung bei gleichem Geschlechterverhältnis randomisiert in jeder Versuchsbedingung. Für die beiden Haupteffekte ergeben sich damit bei Erwartung starker Effekte ( $f = 0.40$ ) und bei einem zu Grunde gelegten Alpha-Niveau von 0.05 akzeptable Teststärken von 0.74 bzw. 0.88 für den F-Test (ermittelt mit dem Programm „G\*Power“, s. Erdfelder, Faul & Buchner, 1996). Bei den Teilnehmenden handelte es sich zum überwiegenden Teil um Psychologiestudierende der Universität Heidelberg, die durch Aushänge im Psychologischen Institut auf die Untersuchung aufmerksam gemacht wurden und bei Bedarf eine Bescheinigung über Versuchspersonenstunden erhielten. Sie verfügten über keine Vorerfahrung hinsichtlich des Szenarios.

*Durchführung:* Alle Probanden bearbeiteten das Szenario „Ökosystem“ für fünf Durchgänge zu je sieben Takten. Im Gegensatz zur üblichen Praxis bei DYNAMIS-Experimenten wurden den Versuchspersonen bereits im ersten Durchgang konkrete Zielwerte mitgeteilt, d. h. es wurde auf eine zeitlich-strukturelle Trennung von Explorations- und Anwendungsphasen verzichtet. Dies geschah aus der Überlegung heraus, dass die Anforderung „Exploriere das System“ ohne Nennung von Erfolgskriterien für Kontrollprobanden allzu unspezifisch ausfallen könnte. Mögliche Einbußen im Wissenserwerb auf Grund spezifischer Zielvorgaben (vgl. Vollmeyer & Burns, 1996) sollten, wenn wirksam, alle Treatment-Gruppen gleichermaßen betreffen; im Fall der Kausalwissen-EG sollte die Beschäftigung mit Kausaldiagrammen dem Effekt entgegenwirken, wie wir vermuteten.

Mit der Bekanntgabe von Zielwerten verknüpft war die Anforderung, das System möglichst geschickt zu steuern. Zudem verwies der Instruktionstext auf die Notwendigkeit explorativer Eingriffe: „Nehmen Sie sich Zeit zu beobachten, wie sich Gift, Schädlingsfresser und Dünger auf die 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 den Probanden überlassen.<sup>1</sup>

Die fünf kombinierten Explorations- und Steuerdurchgänge wurden unterbrochen durch ein je nach Bedingung des Faktors Wissensdiagnose variabel gestaltetes Intervall: Teilnehmer der Kausalwissen-EG bearbeiten nach jedem Durchgang ein Kausaldiagramm, Teilnehmer der Oberflächenwissen-EG einen Rekognitionstest, Teilnehmer der Unterbrechungs-KG einen Emotionsfragebogen; Teilnehmer der Baseline-KG übersprangen das Intervall, um sofort mit dem nächsten Durchgang fortzufahren. Im Anschluss an den fünften Durchgang folgte für alle Versuchspersonen ein Rekognitionstest und ein Kausaldia-

*Tabelle 1.* Überblick über den Versuchsablauf für die vier Gruppen Kausalwissen-EG (Kausalwissen erfragt), Oberflächenwissen-EG (Oberflächenwissen erfragt), Unterbrechungs-KG (Unterbrechung durch Emotionsabfrage) und Baseline-KG (keinerlei prozessbegleitende Diagnostik).

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

<sup>1</sup>Die vollständigen Instruktionstexte und das Versuchssteuerprogramm können bei den Autoren angefordert werden.

gramm. Schulische Leistung und soziodemografische Variablen wurden abschließend zur Kontrolle erfasst. Danach erhielten die Versuchspersonen Gelegenheit, eine individuelle Leistungsrückmeldung anzufordern. Sie wurde ihnen ca. drei Wochen nach der Erhebung per E-Mail zugesandt. In Tabelle 1 ist der Versuchsablauf im Überblick dargestellt.

*Abhängige Variablen:* Zur Prüfung der vorhergesagten Effekte wurden fünf abhängige Variablen herangezogen.

(1) *Güte der Kausaldiagramme:* Abstraktes Strukturwissen wurde anhand der bearbeiteten Kausaldiagramme über den Index „Güte der Kausaldiagramme“ (GdK) nach Müller (1993) ermittelt. GdK verrechnet die Häufigkeiten richtiger und falscher Angaben (Treffer bzw. falscher Alarme) gewichtet mit den Ratewahrscheinlichkeiten als Differenzenmaß. Treffer- und Fehlerzahlen werden zudem an der jeweiligen maximal möglichen Anzahl relativiert, so dass trotz unterschiedlicher Vernetztheit weitgehend Vergleichbarkeit für die hier eingesetzten „Ökosystem“-Versionen gegeben ist. Die folgenden Befunde nehmen fast ausschließlich auf einen GdK-Globalwert Bezug. Dieser entspricht einer gewichteten Summe der Strukturwerte für Relationswissen (Gewichtung: 1), Vorzeichenwissen (Gewichtung: 1) und numerisches Wissen (Gewichtung: 0.5) und ist umso höher, je umfassender und präziser das Strukturwissen ausfällt. Grundlage der ANOVA-Zwischen-Gruppen-Vergleiche ist der GdK-Wert der Kausaldiagramme nach dem fünften Szenario-Durchgang.

(2) *Güte des Oberflächenwissens:* Auf der Grundlage des Rekognitionstests bestimmten wir „oberflächliche“ Wiedererkennungslösungen über den Index „Güte des Oberflächenwissens“: Dieser Index entspricht der Summe der Antwortsicherheiten gegeben eine richtige Antwort abzüglich der Summe der Antwortsicherheiten gegeben eine falsche Antwort. Bei maximal vier richtigen oder vier falschen Antworten mit einer maximalen Antwortsicherheit von 10 ergibt sich für die Güte des Oberflächenwissens ein ganzzahliger Wertebereich von -40 bis +40. Werte von null oder weniger sind als Nullsensitivitäten aufzufassen.

(3) *Güte der Systemsteuerung:* Das anwendungsbezogene Maß „Güte der Systemsteuerung“ (GdS) berechneten wir als Summe der logarithmierten Istwert-Sollwert-Abweichungen über alle Durchgänge und alle Takte nach der von Müller (1993) entwickelten Formel. Um die Problematik nicht-definierter Logarithmen von Null zu umgehen, übernahmen wir das Vorgehen von Vollmeyer, Burns und Holyoak (1996, S. 85), vor der logarithmischen Transformation zu jeder Differenz pauschal den Wert 1 zu addieren.

(4) *Güte der Strategieanwendung:* Strategieeinsatz als Prozessmaß des Problemlösens wurde über einen Strategiepunktswert („Güte der Strategieanwendung“) operationalisiert (vgl. Vollmeyer et al., 1996). Das Konstruktionsprinzip trägt dem empirischen Befund Rechnung, dass beim Umgang mit komplexen, dynamischen Systemen „weniger Manipulation oft mehr ist“. Insbesondere die

Strategie, pro Takt nicht mehr als eine einzelne Inputvariable zu verändern, auf die die nachfolgend beobachteten Effekte dann eindeutig rückführbar sind, steht offensichtlich in einem deutlich positiven Zusammenhang zur Problemlöseleistung (vgl. Putz-Osterloh, 1993b; Rollett, 2003; Vollmeyer & Rheinberg, 1998; Vollmeyer, Burns & Holyoak, 1996). Die Güte der Strategieanwendung in der vorliegenden Untersuchung ist der Mittelwert aller Strategiepunkte, die für jeden Interventionstakt vergeben werden: Zwei Punkte für einen strategisch besonders wirkungsvollen Null- oder Einfacheingriff, einen Punkt für einen Zweifach-Eingriff, null Punkte für einen oft wenig systematischen Dreifach-Eingriff. Hohe Strategiepunktswerte (Maximum: 2 Punkte  $\times$  7 Takte  $\times$  5 Durchgänge = 70) gehen nach dieser Klassifikation mit systematischem Strategieeinsatz einher.

(5) *Bearbeitungsdauer:* Die Bearbeitungsdauer wurde aus den Protokollaufzeichnungen der Computersimulation als Netto-Bearbeitungsdauer berechnet; d. h. es ging ausschließlich der Zeitbedarf für Exploration und Steuerung des Systems ein, nicht aber der Zeitbedarf für intermittierende Aufgaben wie z. B. Ausfüllen der Kausaldiagramme oder Lesen zusätzlicher Instruktionen.

## Ergebnisse

### Zusammenhänge zwischen Strukturwissen und Systemsteuerung

Ein lineares Strukturgleichungsmodell sollte prüfen, inwieweit eine latente Variable Systemidentifikation (I) die Steuerleistung (S), ebenfalls auf der Ebene einer latenten Konstruktvariable, vorhersagen kann. Als manifeste Variablen wurden separate Strukturwerte für Relationswissen (GdK<sub>rel</sub>), Vorzeichenwissen (GdK<sub>vor</sub>) und numerisches Wissen (GdK<sub>num</sub>) herangezogen sowie die Steuergüte GdS für alle fünf Durchgänge (GdS<sub>1</sub>, GdS<sub>2</sub>, GdS<sub>3</sub>, GdS<sub>4</sub>, GdS<sub>5</sub>). Damit inhaltlich positive Zusammenhänge zwischen Wissen und Steuern auch rechnerisch positiv erscheinen, gingen umgepolte GdS-Werte in die Analyse ein.

Mit Ausnahme von GdS<sub>1</sub> stellten sich alle manifesten Variablen als substantielle Prädiktoren mit einer Indikatorreliabilität von mindestens 0.4 heraus. Kritisch hingegen ist die globale Anpassungsgüte des Modells zu beurteilen: Mit einem beinahe signifikanten  $\chi^2$ -Kriterium ( $\chi^2 = 21.53$ ,  $df = 13$ ,  $p = .063$ ) und einem RMSEA von .102 ist eine Diskrepanz zwischen Modell und Datenstruktur relativ wahrscheinlich. Der Goodness-of-Fit-Index und die inkrementellen Maße sprechen wiederum für ein angemessenes und sparsames Modell ( $GFI = .920$ ,  $NFI = .927$ ,  $TLI = .950$ ;  $\chi^2/df = 1.656$ ). Die inhaltliche Interpretation der Pfadkoeffizienten erfolgt also unter Vorbehalt.

Mit einem Pfadkoeffizienten von 12.72 in der unstandardisierten Lösung (s. Abb. 2) ist ein substantiell positiver Prädiktionswert der latenten Variable Identifikation (I) auf die Steuerung (S) abzulesen. Die Höhe des Zusam-

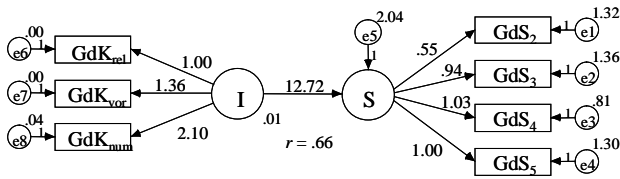


Abbildung 2. Pfadmodell zur Vorhersage von Steuerleistung (S) aus Identifikationsleistung (I). Anpassungsgüte des Modells bei  $N = 64$  Versuchspersonen:  $\chi^2 = 21.53$ ,  $df = 13$ ,  $p = .063$ ;  $RMSEA = .102$ ;  $GFI = .920$ ,  $NFI = .927$ ,  $TLI = .950$ ;  $\chi^2/df = 1.656$ .

menhangs entspricht einem Korrelationskoeffizienten von  $r = .66$ , beinahe übereinstimmend mit der zusätzlich ermittelten Korrelation zwischen dem manifesten GdK-Globalwert und der mittleren GdS über alle Durchgänge ( $r = .64$ ,  $p < .05$ ). Wie zu erwarten ist, werden die Steuerdurchgänge aus den drei letzten Durchgängen im gegebenen Modell am stärksten gewichtet, um die Identifikationsleistung nach dem fünften Übungsdurchgang vorherzusagen.

### Effekte der Systemvernetztheit

Wie bereits bei Funke (1985) und Kluge (2004) experimentell aufgezeigt, gingen wir davon aus, dass Problemlösern das Identifizieren und Steuern dynamischer Systeme mit zunehmender Systemvernetztheit schwerer fällt. Im Mittelwertsvergleich beider Vernetztheitsbedingungen lag die mittlere Steuergüte GdS unter der Bedingung geringer Vernetztheit signifikant niedriger als unter der Bedingung hoher Vernetztheit ( $t(62) = -5.49$ ,  $p < .05$ ,  $d = 1.39$ ); d. h. die Steuerleistung fiel bei geringer Vernetztheit besser aus. Auch die mittlere GdK (Strukturwissen) nach dem fünften Übungsdurchgang war signifikant höher bei Probanden, die das einfache anstelle des komplexeren Szenarios bearbeitet hatten ( $t(62) = 2.28$ ,  $p = .028$ ,  $d = .57$ ). Lediglich für das abhängige Maß Güte des Oberflächenwissens (spezifisches Eingriffswissen) war kein vergleichbarer Vernetztheitseffekt festzustellen ( $t(59) = .72$ ,  $p < .474$ ). Die Mittelwerte für Steuerleistung und Strukturwissen (GdS und GdK) auf den beiden Vernetztheitsstufen erscheinen in Abb. 3 in Verbindung mit den Effekten für den Faktor Wissensdiagnose.

### Kausaldiagramm-Analysen, Problemlösegüte und Problemlöseverhalten

**Strukturwissen:** Als zentralen Indikator für den vermuteten Reaktivitätseffekt von Kausaldiagramm-Analysen zogen wir das nach dem fünften Durchgang erfasste Strukturwissen im Kausaldiagramm heran. Eine globale  $4 \times 2$ -ANOVA mit den Faktoren Wissensdiagnose und Vernetztheit erbrachte auf dem 10%-Niveau signifikante Mittelwertsunterschiede für den Faktor Wissensdiagnose ( $F(3, 56) = 2.41$ ,  $p = .077$ ,  $f = .33$ ) bei einer nicht signifikanten Interaktion mit dem Vernetztheitsfaktor ( $F(3, 56) = .79$ ,

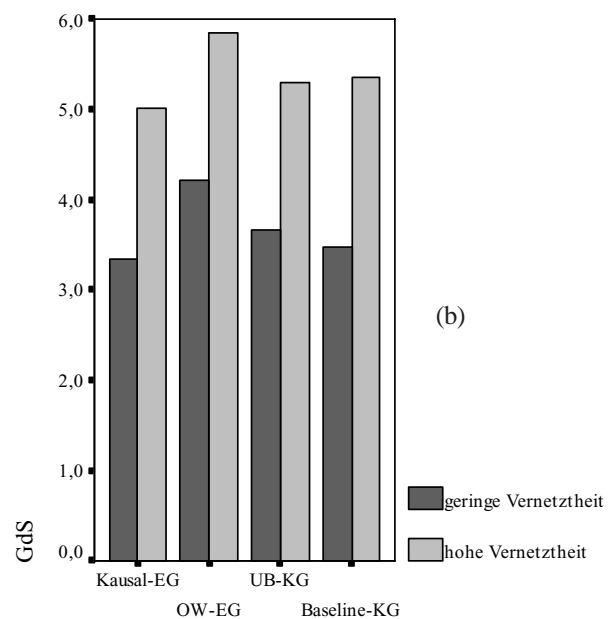
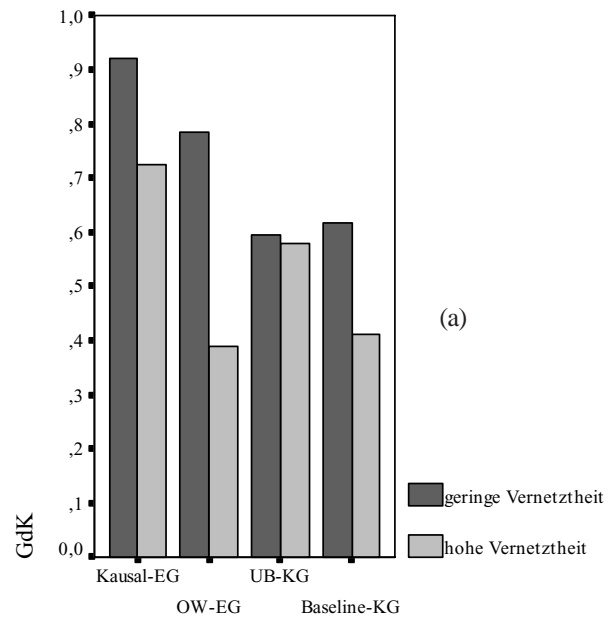


Abbildung 3. Strukturwissen (a) und Steuerleistung (b) unter den vier wissensdiagnostischen Bedingungen gekreuzt mit dem Vernetztheitsfaktor. Hohe Werte für GdK kennzeichnen hohe Identifikationsleistungen, hohe Werte für GdS bedeuten eine geringe Steuerleistung. Abkürzungen: Kausal-EG = Kausalwissen-EG, OW-EG = Oberflächenwissen-EG, UB-KG = Unterbrechungs-KG.

$p = .507$ ). Ein a priori Kontrast testete die Hypothese, dass das Kausalwissen bei Teilnehmern der Kausalwissen-EG gegenüber dem Mittel der drei Vergleichsgruppen erhöht ist. Dieser Kontrast fiel signifikant aus ( $p = .012$ ). Probanden der Kausalwissen-EG erzielten mit einer mittleren GdK von  $M = .82$  ( $SD = .39$ ) den höchsten Strukturwert, Proban-

den der Baseline-KG mit  $M = .51$  ( $SD = .25$ ) den niedrigsten Wert (vgl. Abb. 3 a). Die Analyse der abhängigen Variable Güte des Oberflächenwissens erbrachte demgegenüber weder signifikante Gruppenunterschiede für den Faktor Wissensdiagnose ( $F(3, 53) = .44, p = .727$ ) noch für die Interaktion mit dem ebenfalls nicht signifikanten Komplexitätsfaktor ( $F(3, 53) = .77, p = .516$ ).

**Steuerleistung:** Die Analyse von Reaktivitätseffekten auf die Steuerleistung GdS erfolgte analog zum Vorgehen bei den Variablen GdK und Güte des Oberflächenwissens. Entgegen der Erwartung verfehlte der Faktor Wissensdiagnose die Signifikanz auf einem aussagekräftigen Niveau ( $F(3, 56) = 1.28, p = .291$ ). Eine Interaktion mit dem Faktor Vernetztheit war definitiv zu vernachlässigen ( $F(3,56) = .037, p = .990$ ). Die Irrtumswahrscheinlichkeit für einen vorhandenen Mittelwertsunterschied zwischen der Kausalwissen-EG und dem Mittel der drei Vergleichsgruppen lag im a priori Kontrast bei  $p = .218$  mit einem negativen Kontrastgewicht, d. h. kleinerem mittleren GdS für Versuchspersonen der Kausalwissen-EG. Deskriptiv betrachtet, war die Kausalwissen-EG mit der geringsten GdS allen anderen Versuchsbedingungen überlegen (s. Abb. 3 b), jedoch jenseits jeder interpretierbaren Signifikanzgrenze.

**Strategieinsatz:** Da sich die Annahme einer Normalverteilung für die Variable Güte der Strategieanwendung im Kolmogorov-Smirnov-Anpassungstest als unsicher herausstellte, prüften wir Mittelwertsunterschiede zwischen den Wissensdiagnosebedingungen zunächst mit dem nicht-parametrischen Kruskal-Wallis-Test. Signifikante Bedingungsunterschiede 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 die Strategiegüte in der Kausalwissen-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.

**Bearbeitungsdauer:** Hinsichtlich der Bearbeitungsdauer vermuten wir, dass Teilnehmer der Kausalwissen-EG auf Grund eines hypothesengeleiteten Problemlösestils mehr Zeit benötigen als Vergleichspersonen. Mittelwertsvergleiche der Netto-Bearbeitungszeiten wiesen allerdings ein uneindeutiges Muster auf. Zwar fand sich neben einem schwachen Haupteffekt für Vernetztheit ( $F(1, 56) = 2.81, p = .099$ ) ein signifikanter Haupteffekt für den Faktor Wissensdiagnose ( $F(3, 56) = 2.84, p < .05$ ); dieser war jedoch keineswegs auf eine höhere Bearbeitungsdauer der Kausalwissen-EG gegenüber den anderen drei Gruppen zurückzuführen (a priori Kontrast:  $p = .14$ ). Auffällig war, dass die Unterbrechungs-KG als einzige

Gruppe bei hoher Vernetztheit deutlich mehr Zeit benötigte verglichen mit der Bedingung niedriger Vernetztheit, wie auch eine schwache Interaktion zwischen Wissensdiagnose und Vernetztheit ( $F(3, 56) = 2.33, p = .084$ ) bestätigt. Eine plausible Erklärung kann an dieser Stelle nicht gegeben werden. Bei einer Stichprobengröße von nur acht Personen pro experimenteller Bedingung sind auch zufällige Abweichungen der Bearbeitungszeiten vom Gesamtmittel noch wahrscheinlich.

## Diskussion

*Zur Frage der Reaktivität.* Hauptziel des Experiments war es zu untersuchen, ob und in welcher Weise die Bearbeitung von Kausaldiagramm-Analysen parallel zum Steuern eines DYNAMIS-Szenarios Einfluss auf Problemlösequalität und -verhalten nimmt. Ausgehend von Kluwes (1988) kritischen Anmerkungen und empirischen Befunden bei Süß et al. (1993) vermuteten wir, dass prozessbegleitende Kausaldiagramm-Analysen zu einem überwiegend hypothesentestenden Problemlösestil nach Klahr und Dunbar (1988) anregen. Wir nahmen an, dass Problemlöser beim Verbalisieren von kausalen Systemzusammenhängen systematisch eigene Hypothesen generieren, die einen vergleichbar leistungssteigernden Effekt haben wie explizite, vom Versuchsleiter vorgegebene Hypothesen (vgl. Vollmeyer & Burns, 1996). Als primäres Ergebnis erwarteten wir daher einen erhöhten Wissenserwerb bei Probanden, die wiederholt Kausaldiagramme anfertigten.

Reaktivitätseffekte auf den Wissenserwerb zeigten sich in der erwarteten Weise: Teilnehmer der Kausalwissen-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 dem System – erstmals nach dem fünften Durchgang ein Kausaldiagramm erstellt hatten. Dass sich speziell die Probanden der Kausalwissen-EG von den drei relativ leistungshomogenen Vergleichsgruppen abhoben, spricht für einen Reaktivitätseffekt, der auf den besonderen Aufforderungscharakter von Kausaldiagramm-Abfragen zurückgeht. Denkbare und von uns geprüfte Alternativinterpretationen derart, dass ein oberflächlicher Rekognitionstest oder eine Unterbrechung des Problemlöseprozesses für eine Distraktoraufgabe vergleichbare förderliche Auswirkungen hat, erscheinen kaum plausibel.

Ebenfalls im Zusammenhang mit dem Erwerb von Strukturwissen war angenommen worden, dass sich ein systematisches, kognitiv aufwändiges, hypothesengeleitetes Vorgehen in charakteristischen Prozessmaßen manifestiert: Für Probanden, die Kausaldiagramme bearbeiteten, sollte die Netto-Bearbeitungsdauer erhöht sein. Unter den verwendeten Explorationsstrategien sollten gezielte Manipulationen einzelner Systemvariablen oder das Konstanthalten aller Variablen gegenüber weniger systematischen Doppel- oder Dreifach-Eingriffen überwiegen. Keine der beiden Hypothesen konnte jedoch statistisch be-



kräftigt werden. Hinsichtlich des Strategieeinsatzes fand sich in der Gesamtstichprobe unabhängig von der experimentellen Bedingung ein hohes Maß an Systematik. Die Auswahl überdurchschnittlich leistungsfähiger, studentischer Versuchspersonen mag zu dem hohen Leistungs-niveau bei gleichzeitig geringer Varianz beigetragen haben.

Überraschend und zugleich unklar bleibt der Befund, dass gerade Probanden, die zwischen den einzelnen Durchgängen des Szenarios nach ihrer emotionalen Befindlichkeit befragt wurden, auffallend lange am Szenario arbeiteten, jedoch nur, wenn ein hoher Grad an Systemvernetztheit vorlag. Sollte die wiederholte Reflexion über emotionale Zustände hier tatsächlich die Aufmerksamkeit auf die eigene emotionale Befindlichkeit hin- und von der Problemstellung ablenken, wie man den erhöhten Zeitbedarf begründen könnte, stünde eine Erklärung für die Interaktion mit dem Vernetztheitsfaktor aus. Für die hier durchgeführte Untersuchung legen Einblicke in Individualdaten eine Verzerrung durch Ausreißerwerte näher. 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.

*Zum Vernetztheitseffekt:* Für beide Leistungsindikatoren, die Güte des Kausalwissens und die Güte der Systemsteuerung, gilt, dass sie vom Grad der Systemvernetztheit abhängen. Unter der Bedingung eines eher einfachen, schwach vernetzten Szenarios bewältigten Probanden die Problemanforderungen in Wissenserwerb und Systemsteuerung erwartungskonform signifikant besser als bei höherer Systemvernetztheit. Frühere Ergebnisse von Funke (1985) konnten damit klar repliziert werden, was als erfolgreiche Validitätsprüfung und günstige Voraussetzung weiterer Analysen interpretiert werden kann.

*Zum Zusammenhang von Systemidentifikation und Systemsteuerung:* In einem weiterführenden Schritt prüfen wir die Annahme, dass Kausaldiagramm-Analysen nicht nur den Erwerb von explizitem Strukturwissen fördern, sondern – vermittelt über den Wissenserwerb – auch die Steuerleistung. Die Hypothese impliziert einen positiven Zusammenhang zwischen Systemidentifikation und -steuerung, dem wir in separater Analyse nachgingen. Sowohl die Korrelationen zwischen manifesten Güteindikatoren als auch der pfadanalytisch ermittelte Zusammenhang zwischen latenten Variablen sprechen für deutlich positive Zusammenhänge zwischen Systemidentifikation und -steuerung wie etwa bei Funke und Müller (1988), Kröner (2001) und Müller (1993), nicht aber bei Putz-Osterloh (1993 a, 1993 b) aufgezeigt. Zum Einsatz der Pfadanalyse in dieser Untersuchung bleibt trotz der insgesamt zufrieden stellenden Ergebnisse methodenkritisch anzumerken: Mit einem Stichprobenumfang von nur 64 Versuchspersonen wurden diesbezüglich empfohlene Richtwerte (z. B. Hair et al., 1998) ohne Zweifel deutlich unterschritten. Die Vorhersage der Steuerleistung durch erworbenes Strukturwissen stellt hier zudem keine Vorhersage im zeit-

lich strikten Sinn dar, 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 Modellanpassung 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. Es fand sich aber keine statistische Bestätigung dieser Hypothese. 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* mit Hilfe „systemun-spezifischer Heuristiken“ steuern. Gemeint sind allgemeine, aus der Systemerfahrung abgeleitete Problemlöseregeln und -strategien, die zu effizienterer Steuerung beitragen ohne sich auf erworbenes, inhaltliches Wissen zu stützen. Beckmann bringt dieses zweite Steuerprinzip mit semantischer Einkleidung, einem formalen Oberflächenmerkmal dynamischer Systeme, in Verbindung. Er verglich die Steuerleistung und den Wissenserwerb von Personen, die ein semantisch eingekleidetes System bearbeiteten, mit der Problemlöse-güte von Personen, die ein abstraktes, inhaltsfreies System von formal identischer Struktur bearbeiteten. Während sich die Steuerleistung nicht unterschied, fiel der Wissenserwerb unter der Semantik-Bedingung geringer aus. Ein inhaltlicher, vorwissensassoziierter Kontext, wenn er auch ein stark vereinfachtes Abbild realer Problemsituationen darstellt, scheint einem systematischen Wissenserwerb demnach entgegenzuwirken. Andererseits kann er zu verhältnismäßig ökonomischen, erfolgreichen Steuereingriffen anregen, sofern Vorerwartungen über Variablenzusammenhänge (Weltwissen) der Systemstruktur weitgehend entsprechen. Für das von uns verwendete „Ökosystem“ vermuten wir, dass Vorwissenskompatibilität nicht für alle, aber für eine Mehrzahl der Teilrelationen gegeben ist. Genauere Voruntersuchungen diesbezüglich wären ohne Zweifel aufschlussreich gewesen, um abzuschätzen, inwieweit die Nutzung von Weltwissen für Steuervorgänge von Bedeutung ist.

Da Problemlöse-güte hier nicht absolut, sondern nur relativ beurteilt wurde, lässt sich die Wissens-Steuer-Problematik aus einer zweiten Perspektive untersuchen: Die Frage, weshalb Kontrollprobanden das System trotz oberflächlicherem Wissen über Kausalzusammenhänge erfolgreich steuern, impliziert die Frage, weshalb Experimentalprobanden das System trotz vertieftem Wissen nicht noch

erfolgreicher steuern konnten – ein Befund, wie er auch in frühen Arbeiten von Berry und Broadbent (1984) berichtet wurde. Handelt es sich beim erworbenen Strukturwissen um anwendungsfernes, „träges“ Wissen, das Problemlöser grundsätzlich nicht zur Systemsteuerung nutzen können? Ein Experiment von Preußler (1996) konnte aufzeigen, dass abstraktes Strukturwissen den Steuerprozessen nur unter zwei Bedingungen zugute kommt: Das Strukturwissen sollte erstens in einem anwendungsnahen Kontext und zweitens vor der Wissensanwendung erworben werden. Zumindest die zweite Bedingung war in den parallelen Explorations- und Anwendungsphasen der vorliegenden Untersuchung nicht gegeben. Begrenzte kognitive Kapazitäten könnten eine Entkopplung der beiden kognitiven Teilprozesse (vgl. Sweller, 1988) begründen: Personen richten ihr Problemlöseverhalten entweder auf Wissenserwerb oder auf Steuerung aus, nur in seltenen Fällen gelingt ihnen beides.

Welche Anforderung eine Person als wesentlich ansieht, hängt sicherlich nicht allein von expliziten Instruktionen ab. Vielmehr sollte die Strukturierung der Problemstellung, der mit Teilproblemen verbundene, kognitive und zeitliche Aufwand eine Rolle spielen. Probanden der Kontrollbedingungen, deren nahezu ausschließliche Aufgabe im Steuern des Systems besteht, sollten diesem Ziel verständlicherweise eine größere Bedeutung beimessen und darin vergleichsweise mehr Expertise entwickeln als Experimentalprobanden, die einen Großteil der Bearbeitungszeit den Kausaldiagrammabfragen widmen (einen ähnlichen Redefinitionseffekt vermuten Funke & Müller, 1988). Zusammenfassend könnte man Unterschiede in den wahrgenommenen Anforderungen als Auslöser differenzieller kognitiver Prozesse (hypothesengeleitet vs. experimentegeleitet) annehmen, die zu Unterschieden in der Problemlösequalität führen.

*Implikationen für die Forschungspraxis:* Verschiedene Implikationen für die Forschungspraxis sind denkbar: Zum einen bleibt die Möglichkeit, so genannten nonreaktiven Verfahren (z.B. Blickbewegungsanalysen, Reaktionszeitanalysen) einen höheren Stellenwert einzuräumen. Problematisch an diesem Ansatz ist, dass mit der Analyse basaler Wahrnehmungs- oder einfacher kognitiver Verarbeitungsprozesse lediglich Vorstufen artikulierbarer, mentaler Repräsentationen erfasst werden. Die interessierenden, vollständigen Repräsentationen bleiben der Beobachtung durch sprachlich oder symbolisch vermittelte Methoden wie Kausaldiagramm-Analysen vorbehalten. Aus diesem Grund plädieren wir für eine alternative, möglichst konstruktive Interpretation der Befunde: 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 (vgl. Frensch & Funke, 1995). Im analogen Sinn lässt der Reaktivitätseffekt von Kausaldiagramm-Analysen auf den Wissenserwerb neben der Kritik auch eine kritische Würdigung zu: Kausaldiagramm-Analysen sind *mehr* als nur wissensdiagnostische Instrumente; sie sind ein integrativer Bestandteil der Problemstellung mit hoher diagnostischer Aussagekraft.

## Literatur

- 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.
- 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., Faul, F. & Erdfelder, E. (1997). *G•Power: A priori, post-hoc, and compromise power analyses for the Macintosh* (Version 2.1.2) [Computer program]. Trier: Universität Trier.
- Erdfelder, E., Faul, F. & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments & Computers*, 28, 1–11.
- Frensch, P. A. & Funke, J. (Eds.). (1995). *Complex problem solving: The European perspective*. Hillsdale, NJ: Erlbaum.
- Frensch, P. A. & Rüniger, D. (2003). Implicit learning. *Current Directions in Psychological Science*, 12, 13–17.
- Funke, J. (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle. *Zeitschrift für Psychologie*, 193, 435–457.
- 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.
- Funke, J. & Müller, H. (1988). Eingreifen und Prognostizieren als Determinanten von Systemidentifikation und Systemsteuerung. *Sprache & Kognition*, 7, 176–186.
- Hair, J. F., Anderson, R. E., Tatham, R. L. & Black, W. (1998). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice Hall.
- Janke, W. & Debus, G. (1978). *Die Eigenschaftswörterliste EWL. Eine mehrdimensionale 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. (2004). *Wissenserwerb für das Steuern komplexer Systeme*. Lengerich: Pabst.
- 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.
- Mandler, G. (1990). Interruption (discrepancy) theory: Review and extensions. In S. Fisher & C. L. Cooper (Eds.), *On the move: The psychology of change and transition* (pp. 13–31). New York: Wiley.

- Müller, H. (1993). *Komplexes Problemlösen: Reliabilität und Wissen*. Bonn: Holos.
- 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, 214–225.
- 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.
- 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.
- Rollett, W. (2003). *Strategieinsatz, Informationsnutzung und Motivation beim Explorieren und Steuern komplexer Systeme*. Dissertation, Humanwissenschaftliche Fakultät der Universität Potsdam, Potsdam.
- Schoppek, W. (2002). Examples, rules, and strategies in the control of dynamic systems. *Cognitive Science Quarterly*, 2, 63–92.
- Shanks, D. R. & Johnstone, T. (1998). Implicit knowledge in sequential learning tasks. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 533–572). Thousand Oaks, CA: Sage Publications.
- 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.
- 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.

Prof. Dr. Joachim Funke

---

Psychologisches Institut  
Universität Heidelberg  
Hauptstraße 47–51  
69117 Heidelberg  
E-Mail: Joachim.Funke@psychologie.uni-heidelberg.de