

MARTIN KERSTING, ANNETTE KLUGE

INNOVATIVE EIGNUNGSDIAGNOSTISCHE INSTRUMENTE UND TRAININGSTOOLS - WARUM SIE HÄUFIG SCHEITERN

Das Beispiel computergestützter Problemlöseszenarien

„Die wohl faszinierendste Innovation im Bereich der computergestützten Diagnostik hat die Gruppe um Dörner mit dem Paradigma des »komplexen Problemlösens« und der Simulation Lohhausen (...) geschaffen. Durch die wertvollen Untersuchungen der Gruppe wurde die theoretische Basis geschaffen, die es überhaupt ermöglicht, Managementfähigkeiten in komplexen Problemsituationen eignungsdiagnostisch greifbar zu machen.“ Dieses Zitat von Obermann (1992, S. 302) verdeutlicht die Begeisterung, mit der neue Instrumente in der Praxis begrüßt werden. Computergestützte Problemlöseszenarien fanden nicht nur in der Eignungsdiagnostik, sondern auch im Bereich des Personaltrainings zunächst eine positive Resonanz. Mittlerweile ist die Euphorie verflogen, zumindest die diagnostische Verwendbarkeit der Szenarien wird von zahlreichen Autoren¹ kritisch eingestuft (z.B. Kersting, 1999; Süß, 1996).

Der Artikel zeigt zunächst problematische Kennzeichen computergestützter Problemlöseszenarien auf und verdeutlicht im nächsten Schritt, warum sich Szenarien mit solchen Defiziten für bestimmte Einsatzformen in der Diagnostik und im Training disqualifizieren. Anschließend werden Mindestanforderungen² aufgestellt, die in der Diagnostik und im Training angewandte Szenarien erfüllen müssen. Das Beispiel der computergestützten Problemlöseszenarien wird abschließend genutzt, um über den Forschungs-Praxis Transfer zu reflektieren und Empfehlungen für eine gelungene Zusammenarbeit zwischen Wissenschaft und Anwendung abzuleiten.

Schlüsselwörter: Diagnostik, Training, Problemlösen

EINFÜHRUNG

Mit der Publikation der frühen Arbeiten zum Verhalten bei Problemlöseaufgaben, die von den Autoren als komplex und realitätsorientiert bezeichnet wurden (z.B. Dörner, Kreuzig, Reither & Stäudel, 1983), haben sich computergestützte Problemlöseszenarien insbesondere in der deutschsprachigen psychologischen Forschung etabliert. Etwas später waren die Instrumente auch in der diagnostischen Praxis und im Trainingsbereich en vogue. Mittlerweile existieren zahlreiche Problemlöseszenarien, die überwiegend auf solche Aufgaben rekurrieren, die im Alltag von Entscheidungsträgern (z.B. Bürgermeister, Manager, Entwicklungshelfer usw.) vorkommen können (für einen Überblick siehe Funke, 1995). Die anfängliche Begeisterung für die neuen Instrumente kann man im nachhinein so deuten, daß es Begriffe und Umstände zu geben scheint, deren Inanspruchnahme auf seltsame Weise von der Beweisführung entlastet. Wer attraktive Worte wie „Problemlösen“, „vernetztes“ oder „systemisches Denken“ be- und den Computer als Medium nutzt, braucht seine Argumente offensichtlich nicht weiter zu begründen. Denn die Argumente, die für den Einsatz der Szenarien in der Praxis sprechen, werden zwar mit einiger Plausibilität behauptet, entbehren aber einer empirischen Referenz. Dies gilt beispielsweise für den angeblich hohen Realitäts- und Anforderungsbezug der Szenarien, welcher u.a. von Funke (1986) und Kluwe (1995) überzeugend in Frage gestellt wurde. Die Voraussetzungen für eine sinnvolle Benutzung des Simulationsbegriffs, z.B. Modellanalysen, Spezifikationen von Abbildungsvorschriften und Modellvalidierungen (siehe z.B. Page, 1983), sind für die meisten Szenarien nicht einmal ansatzweise erfüllt. Ähnliches gilt hinsichtlich des Anspruchs auf ökologische

¹Personen und Funktionsbezeichnungen gelten jeweils in weiblicher und männlicher Form.

Validität. Die Behauptung, computergestützte Problemlöseszenarien erfreuten sich generell einer hohen sozialen Akzeptanz, kann in dieser Pauschalität ebenfalls nur durch singuläre Erfahrungen, nicht aber durch empirische Studien abgesichert werden (siehe Kersting, 1998). Und zuletzt mußte auch die Hoffnung, die neuen Verfahren würden einen Zugang zu bislang nicht diagnostizierbaren Fähigkeiten - wie z.B. dem Problemlösen oder einer Fähigkeit zum vernetzten Denken - erschließen, aufgegeben werden. Tatsächlich läßt sich die Steuerung der Szenarien mit den herkömmlichen Fähigkeitskonstrukten Intelligenz und Wissen hinlänglich erklären (siehe Kersting, 1999, 2001; Süß, 1996). Den erhofften und behaupteten Vorzügen (Simulationscharakter, ökologische Validität, soziale Akzeptanz und Zugang zu neuen Fähigkeitskonstrukten) kommt vor allem viel exkulpative Potenz für die in der Praxis schwer handhabbaren Eigenheiten der neuen Instrumente zu. Im folgenden werden nur einige der problematischen Kennzeichen kurz angesprochen, eine ausführliche Darstellung findet sich bei Kersting (1999).

PROBLEMATISCHE KENNZEICHEN COMPUTERGESTÜTZTER PROBLEMLÖSESZENARIEN

- 1) Ein zentrales Problem stellt die sogenannte „Dynamik“ komplexer Szenarien dar, die dazu führt, daß sich der zu steuernde Systemzustand aufgrund vorhergehender Problembearbeitungen oder - im Falle der Eigendynamik - auch ohne eine Aktion des Problemlösers verändert. Infolge dieser „Dynamik“ sind die einzelnen Schritte der Problembearbeitung nicht voneinander unabhängig. Die Szenarien entbinden eine Eskalationsdynamik, in deren Folge es unklar ist, ob die Szenarieneignisse der steuernden Person / Gruppe und / oder eben der Systemdynamik selbst zugeschrieben werden müssen.
- 2) Die formalen Aufgabenmerkmale und Systemkennwerte etlicher Problemlöseszenarien sind nur unzureichend analysiert und beschrieben, Optimallösungen können häufig nicht bestimmt werden (siehe Funke, 1990; Kluwe, Misiak &

Haider, 1991). Die im Training und in der Diagnostik wünschenswerte (vergleichende) Interpretation der Leistungen einzelner Personen oder Gruppen ist dadurch schwer beeinträchtigt, die Bestimmung von Problemlösegütemaßen (siehe nächster Punkt) enorm erschwert.

- 3) Einer der nach wie vor ungelösten neuralgischen Punkte bei der Verwendung von computergestützten Problemlöseszenarien liegt in der Operationalisierung der Problemlösegüte (Hussy, 1985, S. 59). Allein die Vielzahl der existierenden Operationalisierungen / Indikatoren des Problemlöseerfolgs weist auf den Interpretationsfreiraum in der Zuordnungs- und Bewertungsfrage hin. Die - gelegentlich auch ex post facto vorgenommene - Ableitung der Maße des Problemlöseerfolgs erscheint nicht selten arbiträr, was insbesondere für die Bildung sogenannter „Verhaltensmaße“ und „Strategien“ gilt. Demgegenüber erfolgen automatisierte Verhaltensbewertungen zwar mit dem ausreichenden Maß an Standardisierung, lassen dafür aber hinsichtlich der Nachvollziehbarkeit der Auswertung zu wünschen übrig. Insgesamt ist festzuhalten, daß die Bewertung der Steuerungsleistung in zahlreichen Szenarien weder kohärent noch widerspruchsfrei erfolgt. Diagnosen und Trainingseffekte werden insbesondere vereitelt, wenn solche Indikatoren der Steuerungsleistung verwendet werden, die sich nur „indirekt“ an der Zielvorgabe orientieren oder sogar nachweislich nicht mit einer instruktionsgemäßen Problemlösung einhergehen. (Für einen Überblick über die Vielfalt von Problemlösegütemaßen und für eine Sammlung von Fällen, in denen die herangezogenen Problemlösegütemaße der Instruktion an die Problemlöser widersprechen, siehe Kersting, 1999)
- 4) Die Zuverlässigkeit der mit computergestützten Problemlöseszenarien erzielten Messungen ist fraglich, die adäquate Bestimmung der Reliabilität ist problematisch (für einen Überblick siehe Müller, 1993).
- 5) In den bisher durchgeführten Untersuchungen konnte kein hinreichender Zusammenhang zwischen dem Problemlöseverhalten bei unterschiedlichen Systeme-

men nachgewiesen werden. Solange dieser Nachweis aussteht, gilt, daß die Steuerungsleistungen überwiegend szenarienspezifisch sind und über ihren Intelligenz- und Wissensanteil hinaus nicht generalisiert werden können. Dies bedeutet, daß Diagnosen und/oder Trainingseffekte in Abhängigkeit davon variieren, welches Szenario eingesetzt wird.

- 6) Ungeklärt ist der Zusammenhang zwischen dem (Vor-)Wissen der Systemsteuerer und der Steuerungsleistung. Daß dem Wissen der Systemsteuerer über den Inhaltsbereich des Szenarios eine zentrale Bedeutung zukommt, ist theoretisch weitgehend konsensuabel. So sollte z.B. eine Person beim Steuern eines Szenarios, bei dem sie die Rolle eines Managers einer Fabrik übernimmt, im Vorteil sein, wenn sie über ein hohes Wirtschaftswissen verfügt. De facto ist die Befundlage zum Zusammenhang zwischen Wissen und Steuerungsleistungen disparat und es gibt sowohl Studien, in denen ein Zusammenhang aufgezeigt werden konnte als auch Studien, in denen dieser Zusammenhang ausblieb. Zur Erklärung der disparaten Befundlage können sowohl inhaltliche Moderatoren des Zusammenhangs (z.B. unterschiedliche Wissensarten) als auch methodische Probleme (z.B. psychometrische Defizite der in den Studien eingesetzten Wissensdiagnosen) berücksichtigt werden.

ZUM DIAGNOSTISCHEN EINSATZ DER SZENARIEN

Die Frage, ob computergestützte Problemlöseszenarien sinnvoll zur Diagnostik, z.B. zur sogenannten Management Diagnostik, eingesetzt werden können, hängt von den Gütekriterien dieser Instrumente ab. Insbesondere die Validität markiert die Grenze der Sinnhaftigkeit der aufgrund der Szenarien getroffenen diagnostischen Aussage. Die Bestimmung der instrumentinternen Gütekriterien der Szenarien (wie z.B. Reliabilität) ist beeinträchtigt, da sich aufgrund der Dynamisierung die Modelle der klassischen und probabilistischen Testtheorie nur bedingt anwenden lassen (Kubinger, 1993).

Die Dynamik der Szenarien erschwert die Bewertung der Szenarien nach Maßgabe

der Testgütekriterien ebenso wie sie den in der Diagnostik gewünschten Vergleich der Ergebnisse verschiedener Diagnostikanden erschwert. Werden computergestützte Problemlöseszenarien zu Diagnosezwecken eingesetzt, so sind sie - ebenso wie Tests - als diagnostische Instrumente zu klassifizieren und den differentiellen Methoden zuzuordnen. Die folgenden Ausführungen zu Testverfahren von Stapf (1995, S. 233) gelten somit auch für diagnostisch eingesetzte Problemlöseszenarien: *„Aus Gründen der Vergleichbarkeit der Testresultate muß ein Testverfahren allen zu untersuchenden Personen in gleicher, standardisierter Weise appliziert werden. D.h. jegliche Variationen der Situations-Variablen sind untersagt, da ja das Testergebnis allein von der individuellen Merkmalsstruktur der Testperson und keinesfalls von fördernden oder hemmenden Bedingungen der Untersuchungssituation abhängen soll.“* Dieser Forderung werden computergestützte Problemlöseszenarien aufgrund der Dynamik und der Abhängigkeit der einzelnen Steuerungstakte nicht gerecht. Umso bedeutsamer ist es nachzuweisen, daß trotz dieser diagnostisch ungünstigen Voraussetzungen ein Zusammenhang der Steuerungsleistungen mit den Ergebnissen anderer Meßinstrumente sowie vor allem mit externen Kriterien wie z.B. dem Berufserfolg, besteht. Gerade die Kriteriumsvalidierung wurde aber sträflich vernachlässigt, bislang liegen für den eignungsdiagnostischen Bereich nur vier veröffentlichte Studien vor, die z.T. mit methodischen Mängeln behaftet sind. Die Studie von Obermann (1991) zur Kriteriumsvalidität des Szenarios „Airport“ ist in ihrer Aussagekraft eingeschränkt, da sie nur 23 Personen umfaßt und als Kriterium lediglich ein single-Item zur Ausprägung des Problemlöseverhaltens zur Verfügung stand. Demgegenüber wurde in der Studie von U. Funke (Schuler, Funke, Moser & Donat, 1995) methodisch gründlicher gearbeitet, die metrische Qualität des inhaltlich sorgfältig konstruierten Kriteriumsmaßes wurde geprüft, und es wurde eine größere Stichprobe aus dem Bereich Forschung und Entwicklung berücksichtigt. Für das Szenario „DISKO“ konnte nur dann der Nachweis der Kriteriumsvalidität erbracht werden, wenn sich die Analyse auf eine Subgruppe von solchen Personen beschränkte, die (1) im Berufsall-

tag hohen kognitiven Anforderungen ausgesetzt waren und die (2) von solchen Vorgesetzten beurteilt wurden, die aufgrund längerer Zusammenarbeit zu einem validen Urteil befähigt waren. Die erste Studie zur *prädiktiven* Kriteriumsvalidität wurde von Hasselmann (1993) präsentiert. Auch hier gelang der Nachweis der Kriteriumsvalidität (für das Szenario „Textilfabrik“), allerdings untersuchte der Autor eine Gruppe von nur 21 Personen. In einer späteren Darstellung der Untersuchung berichtet Hasselmann (1995), daß in der Studie nicht nur die „Textilfabrik“ eingesetzt wurde, sondern daß 25 weitere Personen das Szenario „Brennstoffvertrieb“ bearbeiteten. Auch für dieses Szenario ergaben sich Korrelationen zu Außenkriterien, allerdings war das Muster der Korrelationen ein anderes als bei der „Textilfabrik“. Dieser Befund weckt erneut Zweifel an der Generalität der mit den Szenarien gemessenen Leistungen.

Insgesamt liegen nur für wenige Szenarien überhaupt Studien zur Kriteriumsvalidität vor. Aufgrund der Systemspezifität der Steuerungsleistungen können die Nachweise der Kriteriumsvalidität nicht auf andere Szenarien übertragen werden, d.h. für zahlreiche, kommerziell für die Diagnostik angebotene Szenarien liegen keine veröffentlichten und nachvollziehbaren Nachweise der Kriteriumsvalidität vor. Allerdings finden sich in einigen Informations- und Werbebrochüren zu Szenarien gelegentlich Hinweise auf angeblich durchgeführte Studien zur Kriteriumsvalidität und deren positive Ergebnisse. Die Angaben bleiben aber vage, die vorgeblich durchgeführten Studien sind nicht hinreichend dokumentiert und auch auf Nachfragen erhält man häufig keine weiteren schriftlichen aussagekräftigen Informationen. Eine solche Vorgehensweise ist ein Indikator für die Unseriosität eines Anbieters. Es gibt keinen Grund, Studien zur Kriteriumsvalidität der Instrumente nicht zu veröffentlichen bzw. aussagekräftige und nachvollziehbare Dokumentationen zu verweigern. Eine solche Praxis widerspricht allen Qualitätsstandards, z.B. auch der DIN-Norm zur Personalauswahl, die voraussichtlich Ende 2001 verabschiedet werden wird. In den bislang genannten Studien konnte für einige wenige Szenarien ein (teilweise aufgrund des geringen Umfangs der Untersuchungsgruppe zweifelhafter) Hinweis auf die

Kriteriumsvalidität erbracht werden, der nicht auf andere Szenarien übertragen werden kann. Der Nachweis, daß ein Instrument kriteriumsvalid ist, ist für die Eignungsdiagnostik allerdings längst nur noch eine notwendige, keinesfalls aber eine hinreichende Voraussetzung zur Legitimation des Instrumenteneinsatzes. Daß die Psychologie über kriteriumsvalid diagnostische Instrumente verfügt, ist ein empirisch genügend abgesicherter Befund (siehe z.B. die metaanalytisch gewonnene Zusammenfassung der Erkenntnisse aus 85 Jahren diesbezüglicher Forschung bei Schmidt und Hunter, 1998). Die eignungsdiagnostische Praxis braucht keine neuen Instrumente, sofern diese gegenüber den vorhandenen Instrumenten keine Vorteile aufweisen. Nachzuweisen ist daher nicht nur, daß ein neues Instrument kriteriumsvalid ist, sondern auch, daß es gegenüber vorhandenen Instrumenten eine *zusätzliche* Treffsicherheit der Diagnosen ermöglicht, also eine *inkrementelle Validität* aufweist. Um diese Frage prüfen zu können, muß die Kriteriumsvalidität neuer Instrumente im direkten Vergleich mit der Kriteriumsvalidität vorhandener Instrumente untersucht werden. Diese Fragestellung wurde in Hinsicht auf die computergestützten Problemlöseszenarien bislang nur von Kersting (1999, 2001) untersucht, der in seiner Studie zwei Szenarien („DISKo“ und die „Schneiderwerkstatt“) gemeinsam mit Intelligenz- und Wissenstests einsetzte und die Vorhersagbarkeit der über Vorgesetztenurteile operationalisierten beruflichen Bewährung von 73 Führungskräften der Polizei vergleichend analysierte. Dabei zeigte sich, daß die beiden Szenarien zwar kriteriumsvalid waren, die Berücksichtigung der Steuerungsleistung aber keinen gegenüber der Intelligenz und dem Wissen signifikant-inkrementellen Beitrag zur Vorhersage der Vorgesetztenurteile leisten konnte. Beim Szenario „DISKo“ wies das standardmäßig vorgesehene und in der Praxis verwendete Problemlösegutemaß „Endergebnis des Gesamtvermögens“ keinerlei Validität auf und erst durch die Bestimmung eines neuen Problemlösegutemaßes konnte die prädiktive Kraft der bei dem Szenario „DISKo“ erzielten Steuerungsleistungen aufgezeigt werden.

EINIGE ANFORDERUNGEN AN SZENARIEN, DIE IN DER EIGNUNGS- DIAGNOSTIK EINGESETZT WERDEN

Die geschilderten Probleme der Szenarien und die bislang vorliegenden Befunde zur Kriteriumsvalidität, bzw. insbesondere das Fehlen entsprechender Studien, gebieten Zurückhaltung beim diagnostischen Einsatz von Problemlöseszenarien. Eingesetzt werden sollten nur Szenarien, für die veröffentlichte Nachweise der Kriteriumsvalidität vorliegen und die mindestens die folgenden Anforderungen erfüllen (nach Kersting, 1999, S. 260):

- Analyse und Dokumentation der formalen Aufgabenmerkmale des Szenarios
- weitgehende Freiheit von Zufallseinflüssen
- Analysen zur Verfälschbarkeit der Steuerungsleistung
- Analysen zu geschlechts- und altersspezifischen Steuerungsleistungen sowie zur Frage der Abhängigkeit der Steuerungsleistung von der Computereinführung
- Nachweis der angemessenen Schwierigkeit der Steuerungsaufgabe für die diagnostische Zielgruppe
- weitgehende Freiheit von Zufallseinflüssen
- direkte und einfache Steuerung des Systems durch die Diagnostikanden
- Lernphase zur Computerbedienung
- Standardisierung der Durchführungszeiten und Bearbeitungstakte
- eindeutige Zielvorgabe der Steuerungsaufgabe
- Transparenz hinsichtlich der Handlungsmöglichkeiten und Bewertungskriterien
- standardisierte Bestimmung der Steuerungsleistung, die sich unmittelbar an den Diagnostikanden genannten Zielvorgabe orientiert
- Im Falle mehrerer Leistungsindikatoren: Vorgabe von konfiguralen oder integralen Bewertungsrichtlinien
- Szenarienspezifischer Nachweis der Reliabilität und der Validität

Darüber hinaus sollten die Szenarien gemeinsam mit Instrumenten eingesetzt werden, die vergleichbare Konstrukte erfassen und psychometrisch ausgereifter sind (z.B. Intelligenztests).

ZUM EINSATZ DER SZENARIEN IM TRAININGSBEREICH

Jeder Interventionsmaßnahme sollte nach Sonntag (1998) eine Anforderungs- und Bedarfsanalyse vorausgehen, denn die Beschreibung menschlicher Verhaltensweisen bei der Arbeitstätigkeit stellt eine wesentliche Informationsquelle für die Gestaltung personaler Fördermaßnahmen dar. Hier zeigt sich jedoch im organisationalen Kontext, daß "(...) *one of the most critical objectives of the requirement analysis is to identify precisely the job being analyzed*" (Goldstein, 1993, S. 47). Dies gilt vor allem in Zeiten rapiden Wandels. Da sich die Tätigkeiten und Anforderungen ständig rapide zu ändern scheinen (häufig mit "Megatrends" bezeichnet wie Globalisierung, technischer Wandel, Wissen als die wichtigste Ressource), ergibt sich die Schwierigkeit, daß definierte Anforderungen schnell „überholt“ sind. Diesem raschen Wechsel der Anforderungen könnte man begegnen – so die Annahme –, indem generelle Fähigkeiten wie eben z.B. das „Problemlösen“, das „vernetzte“ oder „systemische Denken“ trainiert werden. Ein Argument für den Einsatz von computergestützten Szenarien im Training begründet sich demzufolge in der Annahme, daß mit diesem Lernmedium das *Lernziel* des Erwerbs *genereller* kognitiver Fähigkeiten – wie z.B. die Nutzung von Metakognitionen zur Selbststeuerung des Lernprozesses – erreicht werden kann. Die Annahme basiert auf der generellen Erfahrung mit Trainingsstudien im Bereich der kognitiven Psychologie, die als Trainingsziel meist die Vermittlung von generalisierbarem und auf neue Probleme transferierbarem Wissen verfolgten (Funke 1995, S. 196).

Eine zweite Annahme bezieht sich auf die als motivationsförderlich erachtete *Lernmethode* des erfahrungsbasierten Lernens (Experiential Learning sensu Kolb, 1984). Laut Funke (1995, S. 194 f.) wurde die Empfehlung, Szenarien im Training einzusetzen u.a. damit begründet, daß die Szenarien Gelegenheit zum „learning by doing“ sowie besondere Feedbackqualitäten (Verdeutlichung von Zeitverläufen, Systemdynamik und Nebenwirkungen) bieten und eine hohe Motivation der Trainingsteilnehmer auslösen könnten. Funke berichtet, daß in den Trainings mit computergestützten Problemlö-

seszenarien als Trainingsmethoden die Prinzipien des „Learning-by-Doing“, des gestuften Trainings, tutorielle Hilfen zum Aufbau von Wissen sowie die Methoden Strategievermittlung und Selbstreflexion angewandt werden.

Beim Lernen mit Szenarien wird implizit davon ausgegangen, daß die Hypothesenformulierung, das Prüfen der Hypothesen in der Simulation, das Auswerten der Beobachtungen und das Reformulieren der eigenen Annahmen den Lernprozeß im Sinne des erfahrungsbasierten Lernens nach Kolb (1984) fördert.

Mit dem Einsatz von Szenarien wird zusätzlich auch die Diskussion und der Austausch über eigene Wissensstrukturen ermöglicht, um die kognitive Basis für Entscheidungsprozesse zu vereinheitlichen. So ist der Einsatz von computergestützten Szenarien z.B. durch die System Dynamics-Gruppe mit dem Ziel betrieben worden, mentale Modelle aufzudecken und zu verändern, die eigenen Vorannahmen bewußt zu machen, ihre Wirkung zu erproben und mit Hilfe der Methode der Kausalitätsdiagramme zu explizieren (Cavaleri & Sterman, 1997; Isaacs & Senge, 1992). Die Teilnehmer erkennen dabei, wie sie selbst zu Verursachern ihrer eigenen Probleme werden (Sterman, 1989). Welche empirischen Ergebnisse hinsichtlich der Wirksamkeit von computergestützten Szenarien liegen vor? Learning-by-doing führt im Durchschnitt zu einer Leistungsverbesserung im gleichen Szenario und übertrifft diejenige bei traditionellem Lernen durch Instruktion (Kluwe, Misiak & Haider, 1998, zitiert nach Funke, 1995, S. 199). Flexibilitätstraining weist die Übersicht von Funke zufolge keine zusätzlichen Trainingseffekte auf. Nach de Jong und van Joolingen (1998) trägt sog. „scientific discovery learning“ am meisten dazu bei, Fertigkeiten in bezug auf „scientific discovery skills“ zu erwerben.

Die szenariengestützte Vermittlung von Methoden des systemischen Denkens führt nach dem Stand der vorliegenden Studien (Cavaleri & Sterman, 1997; Huz, Andersen, Richardson & Boothroyd, 1997) vor allem zu Verschiebungen in Richtung einer eher strategischen Informationsverarbeitung, zeigt sich aber nicht in verbesserten betriebswirtschaftlichen Kennzahlen.

De Jong und van Joolingen (1998) konnten bei den bisherigen Studien zum Lernen mit computergestützten Szenarien und learning-by-doing keinen klar erkennbaren generellen Vorteil dieser Instrumente aufzeigen. Denn beim Lernen mit Szenarien treten nach de Jong und van Joolingen (1998) u.a. die folgenden Probleme auf, die die Wirksamkeit des erfahrungsbasierten Lernens beeinträchtigen:

Die Lerner

- wissen nicht genau, wie eine Hypothese formuliert wird;
- sind nicht in der Lage, ihre Hypothesen aufgrund der gewonnenen Daten zu verändern, da sie widersprechende Ergebnisse ignorieren, zu eigenen Gunsten reinterpretieren, oder weil sie unfähig sind, alternative Hypothesen zu entwickeln;
- führen im Sinne des Konfirmationsbias überwiegend solche Experimente durch, die lediglich bestätigende Beweise erzeugen und vermeiden solche Hypothesen, die sich mit größerer Wahrscheinlichkeit nicht bestätigen lassen;
- steuern das System derart unsystematisch (z.B. verändern alle Variablen gleichzeitig), daß keine Schlussfolgerungen mehr möglich sind;
- testen nicht das gesamte Spektrum der Handlungsmöglichkeiten aus.

Außerdem fällt es den Personen schwer, die bei den Experimenten gesammelten Erfahrungen auszuwerten. Häufig werden die Ergebnisse der Steuerung falsch encodiert und mißinterpretiert.

Die Nutzung von computergestützten Szenarien im Training sollte ebenso wie die Nutzung in der Managementdiagnostik mit Vorsicht betrieben werden. So erweist sich der Einsatz von computergestützten Szenarien nicht per se, sondern nur dann als sinnvoll, wenn durch die angenommene Authentizität der Lernumgebung Verhaltenspotentiale verändert werden, d.h. wenn gelernt und transferiert wird. Hier zeigt sich das auch sonst bestehende Problem des Trainings generalisierbarer und transferierbarer kognitiver Skills. Um Transfer zu unterstützen, muß im Lernprozeß ein flexibler Umgang mit der Information oder Fähigkeit erworben werden, indem z.B. eine Regel zunächst präsentiert und anschließend in

einer Anzahl variierender Kontexte angewandt wird. Vertreter der Cognitive Flexibility Theorie (Jacobson & Spiro, 1992) fordern dementsprechend, dasselbe Konzept zu verschiedenen Zeiten in verschiedenen Kontexten unter veränderter Zielsetzung und aus verschiedenen Perspektiven beleuchtet zu trainieren. Auf diese Weise soll das Lernen multidirektional und multiperspektivisch erfolgen, um ein facettenreiches Wissen und dessen flexible Anwendung sicherzustellen. Das ist zeitaufwendig und bedarf der Nutzung diverser unterschiedlicher Trainings szenarien. Dies setzt natürlich voraus, daß die zur Steuerung notwendigen Fähigkeiten und Kenntnisse von einem Szenario auf ein anderes transferiert werden können. Gerade diese Generalisierbarkeit könnte bei der Steuerungsleistung aber bislang nicht nachgewiesen werden (s.o.), die Steuerungsleistungen erwiesen sich vielmehr als szenarienspezifisch. Dies führt u.U. dazu, daß ein Transfer der Erfahrungen zu Mißerfolgen im nächsten Szenario führt, ein entsprechendes Trainingskonzept somit u.U. sogar negative Konsequenzen zeitigt.

Um Transfer zu ermöglichen, ist des weiteren ein desorganisiertes Präsentieren von Informationen zu vermeiden (Sternberg & Frensch 1993). Desorganisation kann durch eine ungünstige Art der Programmgestaltung erzeugt werden oder aber auch dadurch hervorgerufen werden, daß die Lerner die Komplexität des Systems unangemessen reduzieren, inadäquate mentale Modelle erwerben oder das System nur mit einer Handvoll Faustregeln (Steuerungswissen) steuern, ohne daß Kausalwissen erworben wurde.

Informationen sind deshalb zusätzlich zur Eigenerfahrung semantisch - z.B. in Form von Skripten - einzubetten. Dabei ist auf eine interne und externe Vernetzung zu achten: Interne, im Sinne einer inneren Logik bzw. Sinnhaftigkeit und externe, durch die Verknüpfung mit Ereignissen aus der Welt der Lerner. Auch diese Voraussetzung ist bei zahlreichen Szenarien nicht erfüllt, denen es insbesondere häufig an einer inneren Logik mangelt oder bei denen die Teilnehmer die Logik des Szenarientwicklers und /oder Programmierers erraten müssen.

EINIGE ANFORDERUNGEN AN TRAININGS MIT COMPUTERGESTÜTZTEN PROBLEMLÖSESZENARIOEN

Die geschilderten problematischen Kennzeichen des Einsatzes von Szenarien in der Personalauswahl treffen ebenso auf die Personalentwicklung zu: vor allem das Fehlen von Studien zur Trainingsvalidität, zur Transfervalidität, zur inter- sowie zur intraorganisationalen Validität (Goldstein, 1993) gebietet auch hier Zurückhaltung beim Einsatz von Problemlöseszenarien als Methode des Wissens- und Fähigkeitserwerbs. Eingesetzt werden sollten nur Szenarien, für die veröffentlichte Nachweise der Kriteriumsvalidität vorliegen, d.h. es sollte der Nachweis vorliegen, daß mit dem Einsatz der Simulation eine Veränderung des Verhaltenspotentials erreicht werden kann. Hager und Hasselhorn (2000, S. 44) schlagen vor, nur dann von „Wirksamkeit eines Programms zu sprechen, wenn beobachtbare Performanzen oder Leistungen durch eine Interventionsmaßnahme beeinflußt werden, und zwar in der von den Vertretern des Programms behaupteten Weise.“ Und hier zeigen sich die oben genannten problematischen Kennzeichen der Szenarien erneut. Um die Wirksamkeit nachweisen zu können, wäre eine Optimallösung hilfreich, zumindest bedarf es einer mit dem Trainingsbedarf korrespondierenden sinnvollen Operationalisierung der Problemlösegüte. Anhand dieser Problemlösegüte könnte dann der Lern- und Transfererfolg bestimmt werden, um Kompetenz- oder Performanzzuwachs feststellen zu können. Um Transferleistungen unter dem Gesichtspunkt des Generalisierens und Beibehaltens zu bestimmen, sind dann Vortest-Nachtest-Follow-up Messungen von Nöten (Hager und Hasselhorn, 2000), die sich aufgrund der Schwierigkeiten der Reliabilitätsbestimmung von Steuerungsleistungen als problematisch erweisen.

Es sollten die folgenden Anforderungen erfüllt sein:

- Ableitung des Szenarien-Trainingsdesigns aus einer Anforderungsanalyse.
- Analyse und Dokumentation der Äquivalenz von Trainingsanforderungen und Anwendungssituation (Voraussetzung: kognitiv orientierte Arbeits- und Anforderungsanalysen des Arbeitsplatzes sowie die ge-

naue Kenntnis der kognitiven Anforderungen des Szenarios).

- eine vorausgehende Differenzierung und Fixierung der Trainingsziele ("Was soll erreicht werden?") sowie eine Festlegung des Ausgangszustandes (Hager & Hasselhorn, 2000): Wie ist die Ausgangslage bei Personen, die für eine bestimmte Interventionsmaßnahme in Frage kommen (gerade im Hinblick auf das Vorwissen, dem eine zentrale Bedeutung zukommt)?
- Abgeschätzt werden sollten dann sowohl Folgewirkungen ("Was kann eintreten, wenn die Ziele erreicht werden?") als auch Nebenwirkungen ("Was kann neben dem angestrebten Zielzustand noch eintreten?"). Es bedarf zusätzlich einer Effektivitätsbeurteilung im Hinblick auf den erwarteten Transfer, insbesondere dem Anforderungs- und Situationstransfer: "Sollen spezifische Testleistungen verbessert, besondere kriterial definierbare Erfolge erzielt, spezifische Fertigkeiten vermittelt oder gar Fähigkeiten verbessert werden?" (Hasselhorn & Mähler, 2000, S. 86). Denn in den bisherigen Untersuchungen konnte kein hinreichender Zusammenhang zwischen dem Problemlöseverhalten bei unterschiedlichen Szenarien nachgewiesen werden. Die Trainingseffekte variierten in Abhängigkeit vom eingesetzten Szenario.
- Angemessener Schwierigkeitsgrad der Szenarien für die Zielgruppe des Trainings. Die Steuerung der Szenarien sollte einerseits nicht zu simpel sein, damit eine Chance auf Transfer besteht, andererseits nicht zu komplex, damit aus der Bearbeitung überhaupt etwas gelernt werden kann. Das gestufte Training (z.B. stufenweiser Anstieg der Komplexität der im Training eingesetzten Szenarien) entspricht dieser Forderung.

FORSCHUNGS- PRAXIS TRANSFER

Am Beispiel der computergestützten Problemlöseszenarien wird deutlich, daß innovative Instrumente aus der Forschung nicht ohne weiteres in der diagnostischen Praxis und im Training eingesetzt werden können. Der Einsatz von Instrumenten und Methoden dient der Erreichung von Zielen, unterschiedliche Ziele erfordern unterschiedliche Gestaltungen der Instrumente. Herrmann

(1995) betrachtet Forschung als Problemlösen und typisiert Forschungsprobleme u.a. nach der Klarheit der Zielvorgabe. Bei der Entwicklung und Evaluation eines in der Diagnostik und im Training eingesetzten Instruments ist das Ziel, der Soll-Zustand, klar definiert. Es geht um die Effizienz- und Rationalitätssteigerung der nicht-forschenden Handlung des Diagnostizierens und Trainierens und somit letztendlich um ein *technologisches Problem*. Davon abgrenzbar sieht Herrmann *wissenschaftliche* oder auch *grundlagenwissenschaftliche Probleme* wie die Entwicklung und Prüfung einer kognitionswissenschaftlichen Theorie als ein Problem ohne Klarheit der Zielkriterien. Während die grundlagenwissenschaftliche Forschung das Problem u.a. rekonstruiert und das Auftreten bestimmter Ereignisse erklärt (und vorhersagt), stellen technologische Forschungsprogramme operatives Hintergrundwissen und standardisierte Techniken (im Sinne normierter Handlungsanweisungen) zur Effizienzsteigerung bereit, wobei die angestrebten Problemlösungen nach Herrmann in der Praxis funktionieren müssen, also auch verlässlich, nebenwirkungsfrei, routinisierbar und nicht zuletzt wirtschaftlich sein müssen. Während der Erfolg der Lösung (grund-)wissenschaftlicher Probleme und theoretischer Fragen von deren Fähigkeit zur Erklärung und Vorhersage bestimmt wird, hängt der Erfolg der Lösung technologischer Probleme oder pragmatischer Fragen von deren *Nutzen* ab. Technologische Regeln sind nach Bunge (1967) nicht wahr oder falsch, sie sind effektiv, unbestimmt oder aber ineffektiv. Der Einsatz der Szenarien in der Forschung erfolgte anwendungsorientiert zu erfolgen. Während der Einsatz der Szenarien in der Praxis teilweise auf den Einzel- und Sonderfall von Verhalten als Produkt einer einzigartigen Konstellation motivationaler, emotionaler und kognitiver Faktoren ausgerichtet war (siehe z.B. den Ansatz von Dörner, 1992), zielt z.B. der eignungsdiagnostische Ansatz auf die Messung (zumindest partiell) konstanter, individuell unterschiedlich ausgeprägter Persönlichkeitsmerkmale, die (zumindest partiell) konstant mit beruflichem Erfolg kovariieren. Situationspezifische Ausprägungen allgemeiner Verhaltensformen (z.B. Entscheiden, Schlußfolgern, Urteilen), welche nur unter nicht

generalisierbaren situationsspezifischen Randfaktoren (wie z.B. spezifischen Hypothesen, Planungen, extremen emotionalen Zustände usw.) auftreten, sind für die Eignungsdiagnostik und das Training im Regelfall unbedeutend.

Hinsichtlich der Praxisverwendung computergestützter Problemlösenszenarien sollte man sich mit Herrmann (ebd., S. 32) vergegenwärtigen, daß Technologie nicht dasselbe ist wie angewandte Wissenschaft. „Vielmehr werden im Kontext technologischer Problemlösungsprozesse (grundlagen-)wissenschaftliche Problemlösungsergebnisse genutzt, indem man sie aus ihrem wissenschaftsimmanenten Zusammenhang löst, sie für den technologischen Zweck selektiert und entsprechend aufbereitet.“ Genau dies ist im Fall der computergestützten Problemlösenszenarien vernachlässigt worden. Für die Diagnostik und das Training mit computergestützten Problemlösenszenarien ist es erforderlich, die Untersuchungsbedingungen, die Instruktion sowie die Auswertung und Interpretation zu standardisieren und z.B. die oben genannten Anforderungen zu erfüllen. Diese Aufbereitung ist bei der Übertragung der Instrumente von der Forschung in die Praxis zu leisten, de facto wurden die Vielzahl der für die Praxis angebotenen computergestützten Problemlösenszenarien aber einfach nur dem Vorbild der in der Forschung genutzten Szenarien nachgebaut. Dies ist umso bedauerlicher, da z.B. im eignungsdiagnostischen Bereich die wenigen Studien zur Kriteriumsvalidität durchaus das Potential der Problemlösenszenarien aufzeigen, welches bislang aber aufgrund eines vorschnellen Einsatzes letztendlich noch unausgereifter Instrumente verspielt wurde. Nutzen wie Schaden neuer Instrumente lassen sich dann abschätzen, wenn man ihnen zumindest eine Chance einräumt: die, sich in der Praxis zu bewähren. Empirische, im Anwendungskontext durchgeführte und nachvollziehbar dokumentierte Bewährungskontrollen sind die Mindestvoraussetzungen, die von Anwendern einzuklagen sind, bevor sie Angebote neuer Instrumente wahrnehmen.

LITERATUR

- Bakken, B., Gould, J., Kim, D. (1994). Experimentation in learning organizations: A management flight simulator approach. In: J.D.W. Morecroft, J.D. Sterman (Hrsg.) *Modelling for Learning Organizations* (S. 243-267). Portland: Productivity Press.
- Bunge, M. (1967). *Scientific research - Vol. 2*. Berlin: Springer.
- Cavaleri, S., Sterman, J.D. (1997). Towards evaluation of systems thinking interventions: a case study. *System Dynamics Review*, 13, 171-186.
- De Jong, T., van Joolingen, W.R. (1998). Scientific Discovery Learning with Computersimulations of Conceptual Domains. *Review of Educational Research*, 68, 179-201.
- Dörner, D. (1992). Über die Philosophie der Verwendung von „Mikrowelten“ oder „Computerszenarios“ in der psychologischen Forschung. In: H. Gundlach (Hrsg.) *Psychologische Forschung und Methode: Das Versprechen des Experiments* (S. 53-87). Passau: Passiva Universitäts-Verlag.
- Dörner, D., Kreuzig, H.W., Reither, F., Stäudel, T. (1983). *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Huber.
- Funke, J. (1986). *Komplexes Problemlösen*. Berlin: Springer.
- Funke J. (1990). Systemmerkmale als Determinanten des Umgangs mit dynamischen Systemen. *Sprache & Kognition*, 9, 143-154.
- Funke, U. (1995). Szenarien in der Eignungsdiagnostik und im Personaltraining. In: B. Strauß, M. Kleinmann (Hrsg.) *Computersimulierte Szenarien in der Personalarbeit* (S. 145-216). Göttingen: Verlag für Angewandte Psychologie.
- Goldstein, I.L. (1993). *Training in Organizations. Needs Assessment, Development and evaluation*. Pacific Grove: Brooks/Cole Publishing.
- Hager, W., Hasselhorn, M. (2000). Psychologische Interventionsmaßnahmen: Was sollen sie bewirken können? In: W. Hager, J.-L. Patry, H. Brenzing (Hrsg.) *Handbuch Evaluation psychologischer Interventionsmaßnahmen. Standards und Kriterien* (S. 41-85). Bern: Huber.
- Hasselhorn, M., Mähler, C. (2000). *Transfer: Theorien, Technologien und empirische Erfassung*. In: W. Hager, J.-L. Patry, H. Brenzing (Hrsg.) *Handbuch Evaluation psychologischer Interventionsmaßnahmen. Standards und Kriterien* (S. 86-101). Bern: Huber.
- Hasselmann, D. (1993). *Computersimulierte komplexe Problemstellungen in der Management-Diagnostik*. Hamburg: Windmühle.
- Hasselmann, D. (1995). Die Konstruktion computersimulierter Szenarien in der Personalarbeit. In: B. Strauß, M. Kleinmann (Hrsg.) *Computersimulierte Szenarien in der Personalarbeit* (S. 237-259). Göttingen: Verlag für Angewandte Psychologie.
- Herrmann, T. (1995). Methoden als Problemlösungsmittel. In: E. Roth (Hrsg.) *Sozialwissen-*

- schaftliche Methoden (S. 20-48). München: Oldenbourg.
- Hussy, W. (1985). Komplexes Problemlösen - Eine Sackgasse? *Zeitschrift für Experimentelle und Angewandte Psychologie*, 32, 55-74.
- Huz, S., Andersen, D.F., Richardson, G.P., Boothroyd, R. (1997). A framework for evaluating system thinking interventions: an experimental approach to mental health system change. *System Dynamics Review*, 13, 149-169.
- Isaacs, W.N. (1993). Taking flight: dialogue, collective thinking and organizational learning. *Organizational Dynamics*, 22, 24-39.
- Isaacs, W., Senge, P. (1992). Overcoming limits to learning in computer-based learning environments. *European Journal of Operational Research*, 59, 187-190.
- Jacobson, M.J., Spiro, R.J. (1992). Hypertext learning environments and cognitive flexibility: Characteristics promoting the transfer of complex knowledge. In: L. Birnbaum (Hrsg.) *The international conference of the learning sciences. Proceedings of the 1991 conference* (S. 240-248). Charlestonville: Association for the advancement of computing in education.
- Kersting, M. (1998). Differentielle Aspekte der sozialen Akzeptanz von Intelligenztests und Problemlösenszenarien als Personalauswahlverfahren. *Zeitschrift für Arbeits- und Organisationspsychologie*, 42, 61-75.
- Kersting, M. (1999). Diagnostik und Personalauswahl mit computergestützten Problemlösenszenarien? Zur Kriteriumsvalidität von Problemlösenszenarien und Intelligenztests. Göttingen: Hogrefe.
- Kersting, M. (2001). Zur Konstrukt- und Kriteriumsvalidität von Problemlösenszenarien anhand der Vorhersage von Vorgesetztenurteilen über die berufliche Bewährung. *Diagnostica*, 47, 67-76.
- Kluwe, R.H. (1995). Computergestützte Systemsimulationen. In: W. Sarges (Hrsg.) *Management-Diagnostik* (2. Aufl., S. 572-578). Göttingen: Hogrefe.
- Kluwe, R.H., Misiak, C., Häider, H. (1991). Modeling the process of complex system control. In: P.M. Milling, E.O.K., Zahn (Hrsg.) *Computer-based management of complex systems. Proceedings of the 1989 international conference of the system dynamics society* (S. 335-342). Berlin: Springer.
- Kolb, D.A. (1984). *Experiential learning. Experience as the source of learning and development*. Englewood Cliffs, N.J.: Prentice Hall.
- Kuhnger, K.D. (1993). Testtheoretische Probleme der Computerdiagnostik. *Zeitschrift für Arbeits- und Organisationspsychologie*, 37, 130-137.
- Müller, H. (1993). *Komplexes Problemlösen: Reliabilität und Wissen*. Bonn: HoloS.
- Obermann, C. (1991). *Airport. Problemlösensimulation V.2.2. Handbuch*. Göttingen: Hogrefe.
- Obermann, C. (1992). *Assessment Center*. Wiesbaden: Gabler.
- Page, B. (1983). Der Gültigkeitsnachweis von komplexen Simulationsmodellen. *Angewandte Informatik*, 25, 149-157.
- Schein, E.H. (1993). On dialogue, culture and organizational learning. *Organizational Dynamics*, 22, 40-51.
- Schmidt, F.L., Hunter, J.E. (1998). Meßbare Personmerkmale: Stabilität, Variabilität und Validität zur Vorhersage zukünftiger Berufsleistung und berufsbezogenen Lernens. In: M. Kleinmann, B. Strauß (Hrsg.) *Potentialfeststellung und Personalentwicklung* (S. 15-43). Göttingen: Verlag für Angewandte Psychologie.
- Schuler, H., Funke, U., Moser, K., Donat, M. (1995). *Personalauswahl in Forschung und Entwicklung. Eignung und Leistung von Wissenschaftlern und Ingenieuren*. Göttingen: Hogrefe.
- Sonntag, Kh. (1998). Ermittlung tätigkeitsbezogener Merkmale: Qualifikationsanforderungen und Voraussetzungen menschlicher Aufgabenbewältigung. In: Kh. Sonntag (Hrsg.) *Personalentwicklung in Organisationen* (S. 157-180). Göttingen: Hogrefe.
- Stapf, K.H. (1995). Laboruntersuchungen. In: E. Roth (Hrsg.) *Sozialwissenschaftliche Methoden* (S. 228-244). München: Oldenbourg.
- Sterman, J.D. (1989). Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science*, 33, 321-339.
- Sternberg, R.J., Frensch, P.A. (1993). Mechanisms of Transfer. In: Dettermann (Hrsg.) *Transfer on trial: Intelligence, cognition, and instruction* (pp. 25-38). Norwood: Ablex Publishing.
- Süß, H.M. (1996). *Intelligenz, Wissen und Problemlösen*. Göttingen: Hogrefe.



DR. MARTIN KERSTING & DR. ANNETTE KLUGE
 Institut für Psychologie der RWTH Aachen
 Jägerstr. 17/19
 D-52056 Aachen
 E-mail: Kersting@psych.rwth-aachen.de
 Annette.Kluge@psych.rwth-aachen.de