

Klieme, Eckhard; Funke, Joachim; Leutner, Detlev; Reimann, Peter; Wirth, Joachim  
**Problemlösen als fächerübergreifende Kompetenz. Konzeption und erste  
Resultate aus einer Schulleistungsstudie**

*Zeitschrift für Pädagogik 47 (2001) 2, S. 179-200*



Quellenangabe/ Reference:

Klieme, Eckhard; Funke, Joachim; Leutner, Detlev; Reimann, Peter; Wirth, Joachim: Problemlösen als fächerübergreifende Kompetenz. Konzeption und erste Resultate aus einer Schulleistungsstudie - In: Zeitschrift für Pädagogik 47 (2001) 2, S. 179-200 - URN: urn:nbn:de:0111-opus-52723 - DOI: 10.25656/01:5272

<https://nbn-resolving.org/urn:nbn:de:0111-opus-52723>

<https://doi.org/10.25656/01:5272>

in Kooperation mit / in cooperation with:

**BELTZ**

<http://www.beltz.de>

#### Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

#### Kontakt / Contact:

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Digitalisiert

# Zeitschrift für Pädagogik

Jahrgang 47 – Heft 2 – März/April 2001

## *Thema: Internationale Schulleistungsvergleichsforschung*

- 155 ANDREAS HELMKE  
Internationale Schulleistungsvergleichsforschung.  
Schlüsselprobleme und Perspektiven  
*Einleitung in den Thementeil*
- 161 KARL-HEINZ ARNOLD  
Schulleistungsstudien und soziale Gerechtigkeit
- 179 ECKHARD KLIEME/JOACHIM FUNKE/DETLEV LEUTNER/PETER REIMANN/  
JOACHIM WIRTH  
Problemlösen als fächerübergreifende Kompetenz. Konzeption und erste  
Resultate aus einer Schulleistungsstudie
- 201 PETER MARTIN ROEDER  
Vergleichende ethnographische Studien zu Bildungssystemen:  
USA, Japan, Deutschland

## *Weiterer Beitrag*

- 217 MANFRED LÜDERS  
Dispositionsspielräume im Bereich der Schülerbeurteilung.  
Auch ein Beitrag zur Professions- und Organisationsforschung

## *Diskussion*

- 235 DIRK RUSTEMEYER  
Wie besonders ist das Allgemeine?
- 251 WERNER HELSPER/LEONIE HERWARTZ-EMDEN/EWALD TERHART  
Qualität qualitativer Forschung in der Erziehungswissenschaft.  
Ein Tagungsbericht
- 271 ROSWITHA LEHMANN-ROMMEL  
Neuere Veröffentlichungen über John Dewey. Eine Sammelrezension

## *Besprechungen*

- 285 KLAUS PRANGE  
*Martina Koch*: Performative Pädagogik.  
Über die welterzeugende Wirksamkeit pädagogischer Reflexivität
- 288 SABINE ANDRESEN  
*Dorle Klika*: Herman Nohl. Sein „Pädagogischer Bezug“ in Theorie,  
Biographie und Handlungspraxis
- 291 JÜRGEN ZINNECKER  
*Fritz Haselbeck*: Lebenswelt Schule. Der Schulalltag im Blickwinkel  
jugendlicher Hauptschülerinnen und Hauptschüler. Einstellungen,  
Wahrnehmungen und Deutungen
- Fritz Haselbeck*: Wie Schüler Schule sehen. Hilferufe an Lehrer,  
Eltern und Politiker? Originale, sehr aufschlußreiche Schüleraussagen  
aus Gruppendiskussionen, Einzelinterviews und Schultagebuch-  
aufzeichnungen

## *Dokumentation*

- 295 Pädagogische Neuerscheinungen

*Beilagenhinweis*: Dieser Ausgabe der Z.f.Päd. liegen Prospekte des  
VNR-Verlag für die deutsche Wirtschaft AG, Bonn, und des Verlags der  
Österreichischen Akademie, Wien, bei.

## *Content*

### *Topic: International Comparative Research On School Achievement*

- 155 ANDREAS HELMKE  
International Comparative Research On School Achievement.  
Key Issues and Perspectives  
An Introduction
- 161 KARL-HEINZ ARNOLD  
Studies On School Achievement and Social Justice
- 179 ECKHARD KLIEME/JOACHIM FUNKE/DETLEV LEUTNER/PETER REIMANN/  
JOACHIM WIRTH  
Problem Solving As Transdisciplinary Competence – Conception and  
first results of a study on school achievement
- 201 PETER MARTIN ROEDER  
Comparative Ethnographic Studies On Educational Systems:  
United States, Japan, Germany

### *Further Contributions*

- 217 MANFRED LÜDERS  
Teachers' Discretionary Powers Regarding Student Assessment –  
A contribution to research on professions and organizations

### *Discussion*

- 235 DIRK RUSTEMEYER  
How Specific Is the General?
- 251 WERNER HELSPER/LEONIE HERWARTZ-EMDEN/EWALD TERHART  
The Quality of Qualitative Research In Educational Science –  
A report on a conference
- 271 ROSWITHA LEHMANN-ROMMEL  
Recent Publications on John Dewey. A Complete Review
- 285 BOOK REVIEWS
- 295 NEW BOOKS

# Problemlösen als fächerübergreifende Kompetenz

## *Konzeption und erste Resultate aus einer Schulleistungsstudie*

### *Zusammenfassung*

Neuere Schulleistungsstudien wie das OECD-Programm PISA zielen nicht nur auf curriculare Leistungsbereiche, sondern auch auf fächerübergreifende Kompetenzen, wie z. B. Problemlösefähigkeit. Das Konzept einer fächerübergreifenden Kompetenz führt zu grundsätzlichen theoretischen und methodischen Fragen: Auswahl und Operationalisierung von Kompetenzen in Schulstudien, Abgrenzung zu fachgebundenen Leistungen und kognitiven Grundfähigkeiten sowie die Frage nach der Bedeutung des schulischen Kontextes für die Kompetenzentwicklung. Empirisch werden diese Fragen im Rahmen einer Konstruktvalidierungsstudie mit mehreren schriftlichen und computergestützten Verfahren behandelt. Erwartungsgemäß erwies sich Problemlösekompetenz als mehrdimensionales Konstrukt. Die schriftlich erfasste analytische Problemlösefähigkeit hing eng mit kognitiven Grundfähigkeiten zusammen, während computergestützte Strategiemasse einen spezifischen Faktor zu bilden scheinen.

## *1. Fächerübergreifende Kompetenzen: Konzeptualisierung und Grundlagen ihrer Erfassung*

### *1.1 Fächerübergreifende Kompetenzen als Gegenstand von Large Scale Assessments*

Der dritten internationalen Mathematik- und Naturwissenschaftsstudie (vgl. BAUMERT/LEHMANN u. a. 1997; BAUMERT/BOS/LEHMANN in Druck) wird nachgesagt, eine „empirische Wende“ in der deutschen Bildungspolitik und möglicherweise auch in der Erziehungswissenschaft eingeleitet zu haben. TIMSS, die Hamburger Studie zur Lernausgangslage in der Sekundarstufe (vgl. LEHMANN/PEEK 1995) und andere Projekte einzelner Bundesländer haben das Thema „Schulleistungskontrolle“ wieder in die fachwissenschaftliche und öffentliche Debatte gebracht, aus der es nach den heftigen, teils methodenkritischen und teils ideologischen Debatten um die Schulsystem-Evaluationen der siebziger Jahre verbannt war. Flächendeckende oder zumindest repräsentative Erhebungen zum Leistungsstand von Schülern sollen der Qualitätskontrolle und somit mittelbar der Qualitätsentwicklung auf verschiedenen Ebenen des Bildungssystems dienen – von der Evaluation der Einzelschule bis zum Systemmonitoring im internationalen Vergleich. Vom Jahr 2000 an beteiligt sich die Bundesrepublik Deutschland an dem von der OECD getragenen „Program for International Student Assessment“ (PISA), in dessen Rahmen alle drei Jahre weltweit die Lesekompetenz sowie Mathematik- und Naturwissenschaftsleistungen von fünfzehnjährigen Schülerinnen und Schülern untersucht werden.

Von früheren Schulleistungsstudien unterscheidet sich diese neue Welle von Erhebungen nicht nur durch die Einbeziehung eines breiten Spektrums von personalen, sozialen und bildungsbezogenen Variablen, mit denen Erklärungshypothesen für unterschiedliche Leistungsniveaus überprüft und Interventionsansätze erschlossen werden. Auch das Konzept von Schulleistung hat sich wesentlich verändert. Die von der „International Association for the Assessment of Educational Achievement“ (IEA) verantworteten Studien wie TIMSS und deren Vorläufer waren noch darum bemüht, Leistungskriterien zu verwenden, die das jeweilige Curriculum und die unterrichtliche Praxis möglichst gut abbildeten. Zu diesem Zweck wurde beispielsweise TIMSS von umfangreichen vergleichenden Curriculumanalysen in den beteiligten Ländern begleitet. Durch Befragung von Lehrplanentwicklern und Fachlehrern wurde das „intendierte Curriculum“ wie auch das „implementierte Curriculum“ spezifiziert, und schließlich wurde die Studie ergänzt durch repräsentative Videomitschnitte aus dem Unterricht in einem Teil der Vergleichsländer. Gerade in Deutschland haben die Verknüpfung von TIMSS mit der Diskussion um die Praxis des Mathematikunterrichts und die Besonderheiten der deutschen Curricula zu breiten, fruchtbaren und häufig kontroversen Diskussionen geführt.

Schon in Teilen der TIMS-Studie war man jedoch bewusst von der Anbindung an Curriculum und Unterrichtspraxis abgewichen. Die mathematisch-naturwissenschaftliche Leistungsfähigkeit breiter Schülerpopulationen am Ende der Sekundarstufe II konnte – nicht zuletzt wegen der äußerst heterogenen Bildungslandschaft mit einer Vielzahl von allgemein bildenden und berufsbildenden Bildungsgängen – nicht mehr in curricular validen Leistungstests abgebildet werden. Die TIMSS-Oberstufentests zur mathematischen und naturwissenschaftlichen Grundbildung (*literacy*) orientierten sich nicht mehr an den curricularen Vorgaben, sondern an normativen Vorstellungen einer breiten mathematisch-naturwissenschaftlichen Allgemeinbildung. Sie stellten einen Kompromiss dar zwischen fachbezogenen, in der jeweiligen internationalen didaktischen Diskussion verankerten Konzepten der Grundbildung im Sinne des Verständnisses fundamentaler Konzepte und Verfahren der Bezugsdisziplinen einerseits und einem lebenspraktisch orientierten Konzept von *literacy*, bei dem die Nutzung mathematisch-naturwissenschaftlicher Konzepte und Fertigkeiten für die Bewältigung alltagsnaher Probleme im Vordergrund stand, andererseits (vgl. KLIEME/BAUMERT/KÖLLER/BOS in Druck). Die Rahmenkonzeptionen des PISA-Programms setzen diese Entwicklung konsequent fort. So wird etwa naturwissenschaftliche Kompetenz (*Scientific Literacy*) im Rahmen von PISA definiert als „die Fähigkeit, naturwissenschaftliches Wissen anzuwenden, naturwissenschaftliche Fragen zu erkennen und aus Belegen Schlussfolgerungen zu ziehen, um Entscheidungen zu verstehen und zu treffen, die die natürliche Welt und die durch menschliches Handeln an ihr vorgenommenen Veränderungen betreffen“ (OECD 1999, S. 66).

Mittelfristiges Ziel der OECD-Bildungsexperten ist es, den Kranz dieser sehr breit, von ihrer Funktionalität für die Lösung von Anwendungsproblemen her gedachten Leistungsbereiche zu ergänzen um Messbereiche, die vollständig von jeder fachlichen und curricularen Anbindung gelöst sind. Die Leitfrage bei der Erweiterung des Systems der Bildungsindikatoren lautete: Welche Kompetenzen brauchen die jungen Leute, die ihre Schullaufbahn abgeschlossen ha-

ben, um in die Lage versetzt zu werden, eine positive Rolle als Bürger in der Gesellschaft zu spielen? Solche „Non Curriculum-bound Outcomes“ bzw. „Cross-curricular Competencies (CCC)“ (OECD 1997) sollten das System der funktional definierten, aber immer noch auf bestimmte Lernbereiche bezogenen Leistungsmessungen ergänzen. Auch erste Vorschläge zur Erfassung fächerübergreifender Kompetenzen und Pilotstudien wurden bereits 1997 von der OECD vorgelegt, allerdings mit unbefriedigenden Resultaten.

Im PISA-Programm wird nun mit besserer theoretischer Fundierung ein neuer Anlauf zur Erfassung von *Cross-curricular Competencies* unternommen. Schon im ersten Zyklus wurden auf der Basis von Vorarbeiten, die J. BAUMERT, H. FEND, H. O'NEIL und J. PESCHAR (1998) geleistet haben, zentrale Komponenten eines Modells des selbst regulierten Lernens in Fragebogenskalen abgebildet. Für die internationalen PISA-Erhebungen des Jahres 2003 ist dann die Erfassung fächerübergreifender Problemlösekompetenzen geplant. Das wissenschaftliche Konsortium, das unter Leitung von JÜRGEN BAUMERT PISA in Deutschland durchführt, bezog jedoch schon für den ersten Zyklus Problemlösekompetenzen und soziale Kompetenzen in das erweiterte nationale Design ein. Im Hauptteil dieses Aufsatzes (Abschnitte 2 bis 4) berichten wir über Vorstudien zur PISA-Problemlösekomponente im Rahmen dieses Designs. Parallele Entwicklungen finden im Rahmen von internationalen Vergleichsstudien statt, die alltags- und berufsbezogene „Life skills“ von Erwachsenen untersuchen (vgl. BINKLEY/STERNBERG/JONES/NOHARA 1999).

## 1.2 Der Kompetenzbegriff aus pädagogischer und psychologischer Sicht

Die ersten Ad-hoc-Vorschläge zur Erfassung fächerübergreifender Kompetenzen, die von der OECD (1997) vorgelegt wurden, gaben Anlass zu der Befürchtung, dass, einem modischen Trend folgend, relativ willkürlich und theoretisch Kompetenzkonstrukte definiert und operationalisiert werden. Um dieser Gefahr zu entgehen, ist es unerlässlich, die neue Diskussion über fächerübergreifende Kompetenzen auf bildungs- und qualifikationstheoretische Theorieansätze und psychologische Konzeptualisierungen des Kompetenzbegriffs zu beziehen. Ohne eine solche theoretische Verankerung läuft man Gefahr, dass Konstrukte, die ursprünglich sozial definiert sind, d.h. von Vertretern gesellschaftlicher Institutionen und Organisationen für wichtig erachtet werden, unter der Hand als wissenschaftliche Konstrukte eingeführt und behandelt werden, wie es in der Pädagogik mit den zahlreichen Aufzählungen von so genannten Schlüsselqualifikationen und in der Psychologie mit der neuerdings inflationären Vermehrung von „Intelligenzen“ (emotionale Intelligenz, soziale Intelligenz usw.) geschieht (zur Kritik vgl. DUBS 1995; WEBER/WESTMEYER im Druck).

In einem Gutachten zur Definition und Auswahl von Kompetenzen für internationale Schulleistungsstudien hat F.E. WEINERT (1999) eine Übersicht über sozialwissenschaftliche Kompetenzbegriffe gegeben und Empfehlungen für die Vorgehensweise in Leistungsstudien formuliert. Nach WEINERT (1999) lassen sich prinzipiell folgende Konzeptualisierungen des Kompetenzbegriffs unterscheiden:

- 1) Kompetenzen als allgemeine intellektuelle Fähigkeiten im Sinne von Dispositionen, die eine Person befähigen, in sehr unterschiedlichen Situationen anspruchsvolle Aufgaben zu meistern.
- 2) Kompetenzen als funktional bestimmte, auf bestimmte Klassen von Situationen und Anforderungen bezogene kognitive Leistungsdispositionen, die sich psychologisch als Kenntnisse, Fertigkeiten, Strategien, Routinen oder auch bereichsspezifische Fähigkeiten beschreiben lassen.
- 3) Kompetenz im Sinne motivationaler Orientierungen, die Voraussetzungen sind für die Bewältigung anspruchsvoller Aufgaben.
- 4) Handlungskompetenz als Begriff, der die ersten drei genannten Kompetenzkonzepte umschließt und sich jeweils auf die Anforderungen und Aufgaben eines bestimmten Handlungsfeldes, zum Beispiel eines Berufes, bezieht.
- 5) Metakompetenzen als Wissen, Strategien oder auch Motivationen, die Erwerb und Anwendung von Kompetenzen in verschiedenen Inhaltsbereichen erleichtern.

In Abwägung unterschiedlicher theoretischer Standpunkte und empirischer Befunde der Kognitions- und Entwicklungspsychologie gelangt WEINERT zu der Empfehlung, den zweiten der hier skizzierten Kompetenzbegriffe zu wählen. In Anlehnung an seinen Vorschlag verwenden wir hier folgende Arbeitsdefinition: Kompetenzen sind Systeme aus spezifischen, prinzipiell erlernbaren Fertigkeiten, Kenntnissen und metakognitivem Wissen, die es erlauben, eine Klasse von Anforderungen in bestimmten Alltags-, Schul- oder Arbeitsumgebungen zu bewältigen. Diese Arbeitsdefinition zeichnet sich vor allem durch vier Merkmale aus:

- a) Kompetenzen sind funktional definiert, d.h. Indikator einer Kompetenz ist die Bewältigung bestimmter Anforderungen.
- b) Der Begriff „Kompetenz“ wird für kognitive Fähigkeiten, Fertigkeiten usw. reserviert. Motivationale Orientierungen werden getrennt davon erfasst.
- c) Kompetenzen sind prinzipiell bereichsspezifisch, d.h. auf einen begrenzten Sektor von Kontexten und Situationen bezogen.
- d) Kompetenzen sind gleichwohl als – begrenzt – verallgemeinerbare gedacht, d.h. als Dispositionen. Die Zuschreibung einer Kompetenz geht über die Feststellung einzelner konkreter Leistungen („Performanz“) hinaus.

Insbesondere lassen sich fächerübergreifende Kompetenzen oder – in WEINERTS Terminologie – „Schlüsselkompetenzen“, die über eine vergleichsweise breite Spanne von Situationen und Aufgabenstellungen hinweg einsetzbar sind, dem allgemeinen Kompetenzbegriff subsumieren. WEINERT zählt hierzu unter anderem muttersprachliche und mathematische Kenntnisse und Fertigkeiten sowie die Inhalte einer basalen Allgemeinbildung. Einerseits muss mit WEINERT (1998, 1999) vor zu großen Erwartungen an die breite Anwendbarkeit bzw. Transferierbarkeit gewarnt werden. Die Vorstellung, bereichsspezifische Kompetenzen könnten durch einen generischen Satz von hoch transferierbaren Schlüsselkompetenzen ersetzt werden, ist nach psychologischen Befunden illusionär (vgl. WEINERT 1999, S. 58f.). Andererseits ist Bildung ohne die Idee des Transfers und der Generalisierung von Kompetenzen auf breite Situationsklas-

sen, d.h., ohne die Idee der formalen Bildung, die materiale Bildung begleitet, nicht denkbar. Aus unserer Sicht ist die Frage, wie weit eine Disposition reicht, d.h. wie breit eine Kompetenz gefasst werden kann, letztendlich nur jeweils für bestimmte Kompetenzkonstrukte und bestimmte Zielpopulationen empirisch zu beantworten. Im Kontext dieser pädagogisch-psychologischen Diskussion dient die Erfassung fächerübergreifender Kompetenzen, ihre Einordnung in komplexe Kompetenzmodelle und die Überprüfung schulischer wie außerschulischer Einflüsse nicht zuletzt dem Ziel, (zu) anspruchsvolle Erwartungen an Schule einer Realitätskontrolle zu unterziehen.

Der funktionale Zugang zum Kompetenzbegriff ist in unserem Zusammenhang besonders attraktiv, weil er die ebenfalls im Kern funktionale Intention des Kompetenzbegriffs der OECD-Experten deckt. Wichtig ist jedoch zu verstehen, dass diese Konzeptualisierung letztendlich im philosophischen Pragmatismus verankert ist und in exponierter Weise auf zwei sozialwissenschaftlichen Paradigmen der siebziger Jahre aufbaut, deren Stärken und Schwächen sie im Wesentlichen teilt:

Die bildungstheoretische Diskussion war gerade in den siebziger Jahren zu großen Teilen geprägt durch die Hoffnung, Ziele und Inhalte von Bildungsprozessen nach der allgemeinen Maßgabe bestimmen zu können, „den Einzelnen zur Bewältigung von Lebenssituationen auszustatten“ (ROBINSON 1967, S. 79). In der funktionalen Auffassung von *Cross-curricular Competencies* lebt gewissermaßen das Prinzip der situationsorientierten Curriculumentwicklung wieder auf – mitsamt der grundlegenden Frage, wie denn die relevanten Klassen von Lebenssituationen auszuwählen und das sie Verbindende zu beschreiben seien. Eine befriedigende Antwort auf diese Fragen konnte offensichtlich im Rahmen der Curriculumentwicklung für allgemein bildende Schulen gar nicht und für berufliche Bildungsgänge nur partiell gegeben werden (vgl. REETZ/SEYD 1983). Gerade auf diese Schwierigkeiten versuchte ja D. MERTENS (1970) mit seinem Konzept „übergeordneter Bildungselemente“ oder „situationsübergreifend transferierbarer Qualifikationen“ zu reagieren, die als so genannte Schlüsselqualifikationen auch vielen neueren Vorstellungen über fächerübergreifende Kompetenzen Pate gestanden haben.

Ebenfalls in den siebziger Jahren wurde in der Psychologie eine neue Form der Beschreibung von menschlicher Leistungsfähigkeit eingefordert, welche die Bewältigung anspruchsvoller Lebenssituationen selbst untersuchen und nicht auf kontextunabhängige Fähigkeitskonstrukte zurückführen sollte. Hierfür erschien gerade der Begriff der Kompetenz geeignet (vgl. den einflussreichen Beitrag von McCLELLAND 1973). A. BANDURA (1990, S. 315) fasst diese Konzeption folgendermaßen zusammen: „There is a marked difference between possessing knowledge and skills and being able to use them well under diverse circumstances, many of which contain ambiguous, unpredictable, and stressful elements.“ Als Alternative zu situationsunabhängig verstandenen Wissens- und Fähigkeitskonstrukten führt BANDURA Kompetenzen ein, die funktional und bereichsspezifisch interpretiert werden.

### 1.3 Auswahl und Konzeptualisierung fächerübergreifender Kompetenzen in der deutschen PISA-Studie

Ohne Verankerung in pädagogischer und psychologischer Theorie würde die Messung von Kompetenzen willkürlich und unsystematisch sein. Wir stützen daher unsere Auswahl und Operationalisierung fächerübergreifender Kompetenzen auf die folgenden drei Schritte:

In einem ersten Schritt werden aufgrund von normativen pädagogischen und bildungspolitischen Vorgaben Funktionsbereiche ausgewählt, von denen angenommen wird, dass sie für eine erfolgreiche persönliche und berufliche Entwicklung zentral sind. Dieser Auswahlprozess stützt sich auf einen Konsens. Im Fall des Programms PISA wurde zum einen der Funktionsbereich des selbst regulierten Lernens ausgewählt in der Annahme, dass der selbstständige Erwerb von Wissen fundamentale Bedeutung hat für eine erfolgreiche personale und berufliche Entwicklung über die Lebensspanne (vgl. BAUMERT u.a. 1998). Zum anderen wurden Problemlösekompetenz und soziale Kompetenz ausgewählt als Funktionsbereiche, die in der pädagogischen Diskussion über Schlüsselqualifikationen, in Arbeitsplatzanalysen (vgl. BINKLEY u.a. 1999) sowie auch als Leitziele schulischer Bildung zumeist an erster Stelle genannt werden.

In einem zweiten Spezifikationsschritt werden dann psychologische Paradigmen betrachtet, die es ermöglichen, den ausgewählten Kompetenzbereich mit Bezug auf anerkannte wissenschaftliche Konstrukte systematisch zu beschreiben und bewährte Operationalisierungsansätze zu verwenden. So stützt sich etwa die von BAUMERT u.a. (1998) entwickelte Rahmenkonzeption für den Bereich des selbst regulierten Lernens auf ein vornehmlich in der Motivationspsychologie verankertes dynamisches Modell der Selbstregulation, während die Erfassung fächerübergreifender Problemlösekompetenz (wie unten in Abschnitt 2 ausführlich dargelegt) auf entsprechenden Paradigmen der kognitiven Psychologie aufbaut.

Die konkrete Entwicklung von Fragebögen und Leistungstests auf der Basis psychologischer Paradigmen erfordert schließlich Einzelentscheidungen über die Auswahl von Kontexten und Aufgabenstellungen. Im Sinne des funktionalen Ansatzes sollte versucht werden, eine möglichst breite Spanne von Situationen abzubilden, die von fachnahen schulischen Situationen über außerfachliche schulische Situationen (d.h. Anforderungen, die das soziale Umfeld Schule, die Arbeit in fächerübergreifenden Projekten usw. betreffen) und Situationen des außerschulischen Alltags bis hin zu fiktiven, jedoch Jugendlichen sehr vertrauten Spielsituationen führt. Der Prozess der Auswahl und Konkretisierung solcher Situationen bedarf wiederum der rationalen und konsensuellen Legitimation.

## 2. Zum Konzept des Problemlösens

Die Arbeiten zum Problemlösen als fächerübergreifender Kompetenz in PISA<sup>1</sup> greifen die einschlägigen psychologischen Forschungstraditionen auf, adaptieren verschiedene Untersuchungsinstrumente aus der Problemlöseforschung für die Zielgruppe der 15-Jährigen und nutzen den bundesweiten Feldtest im Jahr 1999, um die Aussagekraft (Konstruktvalidität) der Verfahren zu prüfen. Wir gehen dabei von folgendem Verständnis des Problemlösens aus:

Problemlösen ist zielorientiertes Denken und Handeln in Situationen, für deren Bewältigung keine routinierten Vorgehensweisen verfügbar sind. Der Problemlöser hat ein mehr oder weniger gut definiertes Ziel, weiß aber nicht unmittelbar, wie es zu erreichen ist. Die Inkongruenz von Zielen und verfügbaren Mitteln ist konstitutiv für ein Problem. Das Verstehen der Problemsituation und deren schrittweise Veränderung, gestützt auf planendes und schlussfolgerndes Denken, sind konstitutiv für den Prozess des Problemlösens (vgl. MAYER/WITTRICK 1996).

Die Definition ist relativ breit, macht aber immerhin deutlich, dass Problemlösen als kognitiver Prozess verstanden wird. Sie grenzt sich damit ab von Alltagsvorstellungen oder auch klinisch-psychologischen Konzepten, in denen „Problemlösen“ mit der Klärung emotionaler und sozialer Konflikte verbunden ist. Wir beziehen uns ausschließlich auf kognitiv-analytisches Problemlösen. Auch hierfür spielt der soziale Kontext eine Rolle – etwa wenn Problemstellungen interaktiv geklärt und kooperativ bearbeitet werden –, und motivationale Faktoren sowie thematisches Interesse und Zielorientierung beeinflussen den Problemlöseprozess. Die Qualität des Problemlösens ist aber primär bestimmt durch das Verständnis der Problemsituation, die Denkprozesse bei der Problembearbeitung und die Angemessenheit der erreichten Lösung. Auch wie systematisch jemand vorgeht – ob er beispielsweise einzelne Komponenten der Problemsituation in kontrollierter Weise untersucht, ob er den Lösungsprozess „global“ oder „lokal“ plant, ob er alternative Lösungsschritte gezielt ausprobiert, ob er Feedback sucht und nutzt –, ist ein wichtiges Unterscheidungsmerkmal zwischen Personen, ein wesentlicher Teil ihrer Problemlösekompetenz.

Für unser Ziel, Problemlösekompetenzen in fächerübergreifenden und außerfachlichen Bereichen zu identifizieren, sind einige Ergebnisse der psychologischen Problemlöseforschung von besonderer Bedeutung, die mit den Stichworten „Intelligenz“, „Komplexes Problemlösen“ und „Bereichsspezifität“ verknüpft sind. Wir referieren im Folgenden kurz diese Ergebnisse und formulieren am Schluss jedes Abschnitts Konsequenzen für das Design der PISA-Problemlösekomponente.

1 Getragen werden diese Arbeiten von einer Expertengruppe, der E. KLIEME (als Sprecher), J. FUNKE, D. LEUTNER und P. REIMANN sowie inzwischen P. FRENCH angehören.

## 2.1 Problemlösen und Intelligenz

Selbstverständlich fällt die Bearbeitung mathematischer Sachaufgaben, das Erkennen von Gesetzmäßigkeiten in Zahlenfolgen (z.B. 1, 2, 4, 7, 11 ...) oder Figurenfolgen, das analoge Schließen (z.B. „/ verhält sich zu \ wie # zu ...?“) unter unseren Begriff des Problemlösens – außer in den seltenen Fällen, wo hoch trainierte Personen solche Aufgaben mit speziellen Algorithmen lösen. Diese Aufgaben gelten aber in der Tradition der Intelligenzforschung als Indikatoren für den Kernbereich der Intelligenz, der „schlussfolgerndes Denken“, „Reasoning“ oder „Verarbeitungskapazität“ genannt wird. Insbesondere Aufgaben zum analogen Denken gelten als ideale, weitgehend von erworbenem (schulischen) Wissen unabhängige Indikatoren der allgemeinen „fluiden“ Intelligenz.

„Problemlösekompetenz“, wie auch immer sie operationalisiert wird, hängt demnach sicherlich mit Maßen der Intelligenz im Sinne des „Reasoning“ zusammen. Wie stark dieser Zusammenhang bei so genannten komplexen Problemlöseleistungen (s. Punkt 2.2) ausfällt, inwieweit also „Problemlösekompetenz“ von „Intelligenz“ abgegrenzt werden kann, ist in der psychologischen Forschung umstritten (vgl. FRENCH/FUNKE 1995; SÜSS 1999).

Eine andere Sichtweise dieser Kontroverse zeigt sich in der amerikanischen Diskussion. R. J. STERNBERG und J. C. KAUFMANN (1998, S. 494) etwa vertreten ein sehr breites Konzept von „Intelligenz“, das diese im Kern gleichsetzt mit Problemlösefähigkeit. Sie unterscheiden drei Teilkomponenten der Intelligenz: a) analytische Fähigkeiten wie z. B. „identifying the existence of a problem, defining the nature of a problem, setting up a strategy for solving the problem, and monitoring one's solution process“, b) kreative Fähigkeiten, „required to generate problem-solving options“, und c) praktische Fähigkeiten, die benötigt werden, um Problemlösungen in realen Kontexten anzuwenden. STERNBERG geht davon aus, dass praktische Intelligenz von klassisch-psychometrisch erfasseter Intelligenz (dem IQ) deutlich unterscheidbar ist. Allerdings ist die „reine“ Messung praktischer Aspekte der Intelligenz bzw. Problemlösefähigkeit eine bislang ungelöste Aufgabe. Im „komplexen Problemlösen“ (siehe unten) mischen sich vermutlich praktische und analytische Anteile. STERNBERG selbst schlägt als Maße der praktischen Intelligenz Verfahren vor, die nicht mehr als Leistungstests verstanden werden können: Er legt den Probanden Beschreibungen von Problemsituationen aus dem Alltag oder Beruf vor und bittet sie, unterschiedliche Handlungsalternativen zu bewerten. Stimmen die Bewertungen mit jenen einer Referenzgruppe („Experten“ im Berufsfeld bzw. repräsentativ ausgewählte Vergleichsgruppen bei Alltagsproblemen) überein, so wird dem Probanden „implizites Wissen“ attestiert, das STERNBERG als Kern der praktischen Intelligenz ansieht.

Ähnlich schwierig erscheint nach dem Stand der Forschung die Erfassung des zweiten Aspekts in STERNBERGS triarchischem Konzept von Intelligenz, der Kreativität. Insofern Problemlösen auf neuartige Situationen abhebt, die nicht mit Standardroutinen zu bewältigen sind, erfordert es stets ein gewisses Maß an Kreativität. Versuche, Kreativität explizit als Qualitätsmerkmal von Problemlöseleistungen zu erfassen (vgl. MUMFORD/SUPINSKI/BAUGHMAN/COSTANZA/THRELFALL 1997) sind bislang jedoch nicht überzeugend.

Das Konzept der PISA-Problemlösekomponente ist darauf ausgerichtet,

Problemlöseindikatoren zu entwickeln, die durch Alltagsbezug und Anwendungsorientierung gekennzeichnet sind. Diese Indikatoren sollen aber Leistungen erfassen, nicht „implizites Wissen“. Die Problemlöseaufgaben werden somit Aspekte von praktischer Intelligenz erfassen, sicherlich aber auch analytische Komponenten. Die Zusammenhänge mit „Reasoning“ werden substantiell sein; die Problemlöseindikatoren sollten dennoch einen eigenständigen Beitrag zum kognitiven Profil der Schüler bzw. zum Leistungsprofil von Schulen liefern.

## *2.2 Problemtypen: Reichweite, Kontext, Komplexität und Dynamik*

Unsere Definition des Problemlösens umschließt einen sehr breiten Bereich von Problemlöseprozessen:

- Die Reichweite der Problemstellung kann von begrenzten Teilen einer Aufgabenbearbeitung über die Handlungssteuerung bis zur Planung und Ausführung umfangreicher Projekte führen.
- Der inhaltliche Kontext kann in unterschiedlichen Domänen (Fächern, Stoffgebieten, Handlungszusammenhängen, Erfahrungsbereichen) angesiedelt sein, die eher theoretischer oder eher praktischer Art, eher akademisch/schulisch oder eher alltagsbezogen sind. In Bezug auf diese Domänen kann das Problem mehr oder weniger authentisch sein.
- Das Problem kann unterschiedlich geschlossen oder offen sein, mit wohl definiertem oder schlecht definiertem Ziel, mit transparenten (explizit angegebenen) oder intransparenten Rahmenbedingungen, mit wenigen, isolierten oder zahlreichen, vernetzten Elementen. Diese verschiedenen Merkmale bestimmen insgesamt die Komplexität des Problems.

In den vergangenen Jahren hat sich die Psychologie zunehmend komplexen, authentischen Problemen mit größerer Reichweite zugewandt. Es geht nicht mehr nur darum, wie wohl definierte „Puzzles“ (im Extremfall reine Reasoning-Aufgaben) durch Anordnung geeigneter Operationen gelöst werden, sondern um das Denken von Experten in wissenschaftlichen und beruflichen Domänen (REIMANN 1998), um Planen und Problemlösen in Alltagskontexten (LAVE 1988; FUNKE/FRITZ 1995), um das Verstehen und Steuern von komplexen ökologischen, ökonomischen oder technischen Systemen (vgl. FRENSCH/FUNKE 1995).

Als Mittel zur Untersuchung komplexer Problemlöseleistungen haben sich Simulationen bewährt. In Interaktion mit dem Computer exploriert der Bearbeiter das simulierte System, entwickelt und testet (mehr oder weniger systematisch) Hypothesen über Zusammenhänge und Gesetzmäßigkeiten, erwirbt Wissen und kann schließlich durch gezielte Eingriffe den Zustand des Systems steuern. Unter anderem wird mit realistischen Simulationen hoch vernetzter ökologisch-ökonomischer Systeme gearbeitet (vgl. LEUTNER/SCHRETENBRUNNER 1989), mit systematisch konstruierten, kleineren diskreten Systemen („finite Automaten“; vgl. BUCHNER 1999) und mit virtuellen Experimentierumgebungen. Aus pädagogisch-psychologischer Sicht lassen sich solche Verfahren

als Umgebungen für entdeckendes Lernen verstehen (vgl. van Someren/REIMANN/BOSHUIZEN/DE JONG 1998). Aus Sicht der Problemlöseforschung ist mit diesen Instrumenten eine neue Qualität von Problemstellungen gegeben, die durch hohe Komplexität und insbesondere durch Dynamik charakterisiert ist (vgl. Frensch/FUNKE 1995). Der dynamische Charakter hat drei Vorteile, die mit statischen Aufgaben (Papier-und-Bleistift-Verfahren) nicht realisiert werden können:

- a) Die Anforderung wird erweitert um das aktive Suchen und kontinuierliche Verarbeiten von externer Information, insbesondere von Feedback zu Wirkungen eigener Problemlöseschritte. Auch beim schriftlichem Problemlösen können verschiedene Strategien eingesetzt, evaluiert und gegebenenfalls modifiziert werden; die Interaktion mit dem Computer zwingt jedoch unausweichlich dazu.
- b) Die Problemsituation kann mit diesem Medium wesentlich authentischer gestaltet werden als bei einem schriftlichen Verfahren.
- c) Es können nicht nur Resultate, sondern auch Verlaufsmerkmale des Problemlöseprozesses erfasst werden, z.B. Art, Häufigkeit, Dauer und Abfolge der Eingriffe in das System. Hieraus lassen sich Indikatoren über Problemlösestrategien, z.B. die Systematik des Problemlöseverhaltens, gewinnen.

Die drei Vorteile sind unabhängig voneinander nutzbar und lassen es insgesamt angeraten erscheinen, Problemlöseleistungen am Computer zu erfassen. Hinsichtlich der Verwendung von Strategiemaßen bestehen allerdings große methodische Probleme. Wie solche Maße definiert werden können, welche Rolle dabei Aspekte der Arbeitsgeschwindigkeit haben, wie verlässlich solche Maße sind und inwieweit sie über verschiedene Simulationen hinweg vergleichbar sind, ist in der aktuellen Forschung ein ungelöstes Problem.

Das PISA-Messkonzept zum Problemlösen bezieht – entsprechend der Breite des psychologischen Problemlösebegriffs – Problemlöseprozesse ein, die hinsichtlich Reichweite, Kontext, Komplexität und Dynamik unterschiedliche Ausprägungen besitzen. Es nutzt die Möglichkeiten computergestützter Simulationen, insbesondere zur Erfassung von Strategiemerkmalen, und bezieht unterschiedliche Formen von Simulationsszenarien ein. Daneben werden aber auch schriftliche Verfahren – bis hin zu Reasoning-Aufgaben – eingesetzt, a) um die theoretische Breite des Problemlösebegriffs abzudecken, b) um die (wissenschaftlich noch nicht hinreichend geklärte) Frage des Zusammenhangs zwischen komplexem Problemlösen und schlussfolgerndem Denken sowie die methodischen Probleme der Interpretation von Strategieparametern für den Kontext einer Schulleistungstudie zu klären sowie schließlich c) um eine breitere Stichprobe von Schulen einbeziehen zu können.

### 2.3 Bereichsspezifität von Problemlöseleistungen

Eine der wichtigsten Erkenntnisse der kognitionspsychologischen Forschungen der letzten Jahre ist der Befund, dass anspruchsvolle Problemlöseleistungen nicht ohne einen Fundus von Wissen in der jeweiligen Domäne erbracht wer-

den können. Das Konzept des Problemraums, in dem sich ein *General Problem Solver* mittels allgemeiner Suchstrategien bewegt (vgl. NEWELL/SIMON 1972), erwies sich als zu einfach, um das Verstehen von Problemsituationen und den Prozess der Lösungsfindung zu beschreiben. Auch die Bemühungen, im Rahmen der Forschungen zum komplexen Problemlösen so etwas wie eine allgemeine, bereichsunabhängige Kompetenz zur Steuerung dynamischer Systeme („operative Intelligenz“) zu bestimmen, waren nicht erfolgreich; die Leistungen im Umgang mit einem solchen System sind nur bedingt auf andere Systeme übertragbar (vgl. KLUWE/HAIDER 1990).

Problemlösen setzt Wissen über Konzepte und Sachverhalte (deklaratives Wissen) und Wissen über Regeln und Strategien (prozedurales Wissen) im jeweiligen Gegenstandsbereich voraus. Empirische Belege für die Bedeutung von Wissen stammen aus dem Vergleich der Problemlösestrategien, die Experten und Novizen einsetzen (vgl. REIMANN 1998), aus Analysen zu Bedingungsfaktoren der schulischen Leistungsentwicklung (vgl. HELMKE/WEINERT 1985), aus Laboruntersuchungen zum naturwissenschaftlichen Lernen (vgl. GLASER/SCHAUBLE/RAGHAVAN/ZEITZ 1992) ebenso wie aus Studien zu alltäglichen Problemlöseleistungen – etwa von brasilianischen Straßenverkäufern, die sich bei der mathematisch durchaus anspruchsvollen Kalkulation von Preisen auf komplexe, von der Schulmathematik abgelöste Regeln stützen (vgl. LAVE 1988).

Die zentrale kognitive Kompetenz, die Problemlöseleistungen in unterschiedlichen Inhaltsbereichen gemeinsam zugrunde liegt, ist vermutlich gerade die Fähigkeit zum schlussfolgernden Denken. Der Erfolg einzelner Problemlöseprozesse ist dann durch das Zusammenspiel von bereichsspezifischem Wissen und Reasoning erklärbar. Aufgabenstellungen, die mit erlernten Routinen lösbar sind, oder Aufgaben, welche die Schüler überfordern und bloßes Versuchs- und Irrtum-Verhalten auslösen, können nach diesem Verständnis nicht als Problemlöseaufgaben gelten. Testaufgaben müssen daher sorgfältig auf die Zielpopulation ausgerichtet werden, und das relevante Vorwissen der Schüler sollte nach Möglichkeit empirisch kontrolliert werden.

Für die Diagnose von Problemlösekompetenz sind die obigen Überlegungen von kritischer Bedeutung. Die Idee einer messbaren universellen, bereichsübergreifenden Problemlösefähigkeit ist nicht haltbar. Ziel der Erfassung von Problemlösekompetenz in PISA ist vielmehr ein Profil von Maßen, die verschiedene Arten des Problemlösens in unterschiedlichen inhaltlichen Bereichen ansprechen. Der Auswahl dieser Bereiche und insbesondere der Balance zwischen hoher Spezifität einerseits, Nähe zur allgemeinen Intelligenz im Sinne von Reasoning andererseits kommt eine große Bedeutung zu.

### 3. Design zur Messung von Problemlösefähigkeiten in der deutschen PISA-Studie

Die Überlegungen des vorangehenden Abschnitts lassen es notwendig erscheinen, in die Problemlösedagnostik, vor allem bei der Konstruktvalidierung im Rahmen des PISA-Feldtests, mehrere Instrumente einzubeziehen, die sich unterscheiden hinsichtlich

- a) des inhaltlichen Kontextes (alltagsbezogene, spielerische oder fächerverbindende schulische Kontexte),
- b) der Komplexität und Dynamik (u.a. bestimmt durch schriftliche vs. computergestützte Vorgabe) sowie
- c) des Paradigmas der psychologischen Problemlöseforschung, das für die Aufgabenstellung adaptiert wird.

Insgesamt kamen im PISA-Feldtest drei computergestützte und drei schriftliche Problemlöseverfahren zum Einsatz, deren Bearbeitung jeweils ca. 45 Minuten erfordert. Einige Verfahren konnten aus der Problemlöseforschung adaptiert werden, andere wurden völlig neu entwickelt.

### *Test 1: Virtuelles Labor*

(Autoren: REIMANN/SCHMITT)

Der Schüler soll bestimmte Gesetzmäßigkeiten induktiv in einer Serie von simulierten Experimenten erschließen. Er wählt dazu jeweils die experimentellen Bedingungen und lässt sich das Experiment in einem Videoclip vorführen. Die Abfolge der Experimente und ihrer Resultate wird in einer Tabelle festgehalten. Auf der Basis dieser Informationen kann der Bearbeiter Hypothesen formulieren, in nachfolgenden Experimenten prüfen und als richtig oder falsch bewerten. Erfasst wird, ob die im Gegenstandsbereich zentralen Hypothesen aufgestellt werden und wie systematisch der Bearbeiter vorgeht.

### *Test 2: Ökologisches Planspiel*

(Autor: SCHRETTENBRUNNER, Adaptation: LEUTNER)

Mit diesem Planspiel, das die Situation eines Agrarbetriebes simuliert, werden Themen unterschiedlicher Fächer angesprochen. Das System ist zudem relativ komplex und kann daher die Forschungstradition des „komplexen Problemlösens“ gut abbilden (vgl. LEUTNER/SCHRETTENBRUNNER 1989). Das System ist von Fachexperten als realitätsadäquat bewertet worden. Erfasst wird hier, wie effizient der Bearbeiter das System exploriert, welches Wissen er dabei erwirbt und wie gut er in einem abschließenden Testdurchgang „seinen Betrieb“ bewirtschaften kann.

### *Test 3: Raumfahrtspiel*

(Autoren: FUNKE/WAGENER)

In der Forschung zum „komplexen Problemlösen“ wurden nicht nur dynamische Systeme mit quantitativen Variablen untersucht, wie beispielsweise die zuvor beschriebene ökologische Simulation, sondern auch diskrete Systeme. Sie bestehen aus einer Reihe von Zuständen, die durch bestimmte Eingriffe (darstellbar als Betätigen von „Schaltern“) verändert werden. Alltagsbeispiele sind Fahrkartenautomaten oder elektronische Geräte. Es können aber auch technische, ökonomische oder soziale Prozesse mit derartigen Modellen (so genannten „finiten Automaten“) dargestellt werden (BUCHNER 1999; FRENSCH/FUNKE 1995). Für PISA wurde eine Variante entwickelt, die in ein Raumfahrtszenario eingebettet ist. Erfasst wird auch hier, wie effizient der Bearbeiter das System exploriert, welches Wissen er dabei erwirbt und mit welchem Erfolg er abschließende Steuerungsaufgaben bewältigt.

*Test 4: Projektaufgaben*

(Autoren: KLIEME, EBACH u.a.)

Dieser Aufgabentyp wurde ursprünglich für die Hamburger Schulleistungsstudie entwickelt (vgl. KLIEME/EBACH u.a. im Druck) und soll parallel zu PISA auch in einer internationalen Studie zu Basiskompetenzen von Erwachsenen eingesetzt werden (vgl. BINKLEY u.a. 2000). Das Konzept entstand in Analogie zu berufspädagogischen Konzepten einer „handlungsorientierten Diagnostik“. Ein „Projekt“, wie z.B. die Organisation eines Klassenfestes, wird in mehrere Arbeitsschritte aufgegliedert (Ziele klären – Informationen beschaffen – Planen – Entscheiden und Ausführen – Bewerten), zu denen jeweils einzelne Aufgaben gestellt werden. Der Test erfasst, wie Re-Analysen der Hamburger Erprobungsdaten zeigten, im Wesentlichen die Fähigkeit, Informationen zu erfassen, zu ordnen und zu kombinieren, aus unterschiedlichen Perspektiven zu bewerten und – unter Berücksichtigung multipler Zielsetzungen – Entscheidungen abzuleiten. Der PISA-Test verwendete zwei derartige „Projekte“ mit unterschiedlichen Kontexten.

*Test 5: Technisches Problemlösen*

(Autoren: O'NEIL, BAKER u.a.; Adaptation: KLIEME/WIRTH)

Eine Arbeitsgruppe unter Leitung von H. O'NEIL hat an der University of California, Los Angeles, einen Aufgabentyp zur Diagnose von Problemlösekompetenz entwickelt, der im Rahmen des Feldtests mit erprobt wurde, damit die deutsche Studie international anschlussfähig bleibt (vgl. BAKER 1998). Den Probanden wird die Skizze eines einfachen technischen Gerätes vorgelegt. Der Test erfordert den Aufbau eines mentalen Modells seiner Funktionsweise und Schlussfolgerungen in diesem Modell, z.B. über Fehlerursachen.

*Test 6: Analoges Problemlösen*

(Autoren: WIRTH/KLIEME)

Es werden mehrschrittige Planungs- und Ordnungsaufgaben gestellt, die jeweils in eine Alltagsgeschichte eingebunden sind. Zu jeder Aufgabe wird sodann eine (mehr oder weniger) analoge Bezugsaufgabe mit fertiger Lösung vorgegeben. Der Bearbeiter muss die Beziehungen zwischen seiner Aufgabe und der analogen Bezugsaufgabe erkennen und den Lösungsweg übertragen. Erfasst wird, wie gut der analoge Transfer beim Problemlösen gelingt.

Im Rahmen der PISA-Felderprobung, die bundesweit im Sommer 1999 in allen Schulformen der Sekundarstufe I durchgeführt wurde, wurden auch die sechs Problemlöseverfahren eingesetzt. Für den Problemlöseteil der PISA-Studie wurde ein vergleichsweise aufwendiges Konstruktvalidierungsdesign entworfen. Einbezogen waren hier insgesamt 650 Schüler aus 18 Schulen in vier Bundesländern. Pro Schule waren 35 Schüler ausgewählt, die überwiegend der internationalen Zielgruppe der Fünfzehnjährigen, zum geringeren Teil der deutschen Ergänzungspopulation von Schülern der 9. Jahrgangsstufe angehörten. Alle Schüler unserer Stichprobe bearbeiteten – wie die nahezu 4000 übrigen Teilnehmer der PISA-Felderprobung – an einem ersten Testtag die internationalen Aufgaben zum Leseverständnis, zur Mathematik und zu den Naturwissenschaften sowie einen Begleitfragebogen. An einem zweiten Testtag

wurden nationale Ergänzungsaufgaben und Fragebögen durchgeführt; in diesem Fall bearbeitete jeweils die Hälfte der Schüler Test 4 (Projektaufgaben) und Test 6 (Analoges Problemlösen); hinzu kamen als Maße der kognitiven Grundfähigkeit die Subtests „Figurale Analogien“ und „Wortanalogien“ des kognitiven Fähigkeitstests (vgl. HELLER/GAEDIKE/WEINLÄDER 1985) sowie ein Maß der *mental speed*. Den Schülerinnen und Schülern wurde zudem ein nationaler Ergänzungsfragebogen vorgelegt. Ausschließlich für die 18 Schulen der Problemlöse-Stichprobe wurde schließlich ein dritter Testtag organisiert zu einem mit der jeweiligen Schule individuell vereinbarten Termin. Mitarbeiter des Max-Planck-Instituts für Bildungsforschung bereisten, ausgerüstet mit einem Satz von identisch konfigurierten Laptops, die Schulen und führten jeweils in einer Doppelstunde Zusatzerhebungen durch.

Insgesamt bearbeiteten nach einem ausbalancierten Design die Schüler die Verfahren in wechselnden Kombinationen, sodass jedes einzelne Verfahren von etwa 180 Schülern und jede Kombination von ca. 60 Schülern bearbeitet wurde. Die Daten aller Einzelverfahren wurden am Max-Planck-Institut für Bildungsforschung zusammengespielt und analysiert. Im Folgenden berichten wir über ausgewählte Befunde dieser ersten Analysen.

#### 4. Ergebnisse der Konstruktvalidierungsstudie

Unsere Darstellung und Diskussion der Befunde orientiert sich an den folgenden Fragen:

- a) Einsetzbarkeit und Zuverlässigkeit der ausgewählten Instrumente,
- b) Identifizierbarkeit von Problemlösekompetenzen und Abgrenzung gegen fachgebundene Leistungen einerseits, kognitive Grundfähigkeiten andererseits (konvergente und diskriminante Validität),
- c) Sensitivität für Einflüsse der schulischen Bildung.

##### 4.1 Einsetzbarkeit und Zuverlässigkeit der Messinstrumente

Tabelle 1 fasst Ergebnisse zu Schwierigkeit und Zuverlässigkeit (Reliabilität) der Tests zusammen. Bei den drei computergestützten Verfahren sind jeweils Kennwerte für den Wissenstest aufgeführt, mit dem überprüft wurde, welche Kenntnisse die Probanden über das simulierte System erworben haben; dieser Test ist am ehesten mit den schriftlichen Verfahren vergleichbar. Der Schwierigkeitsindex liegt bei allen Verfahren zwischen einem Drittel und zwei Drittel. Alle sechs Verfahren vermögen somit gut zwischen den Bearbeitern zu differenzieren; lediglich beim Technischen Problemlösen und beim Analogem Problemlösen gibt es leichte Bodeneffekte.

Die Reliabilitäten der schriftlichen Tests sind insgesamt befriedigend bis gut. Dies gilt auch für die Wissenstests zu den drei computergestützten Verfahren. Der zu niedrige Wert beim Technischen Problemlösen resultiert aus der geringen Zahl von nur fünf Items. Die niedrige Zuverlässigkeit des Wissenstests beim ökologischen Planspiel resultiert vermutlich aus der Tatsache, dass

Tab. 1: Schwierigkeit und Reliabilität der eingesetzten Testverfahren		
Verfahren	Schwierigkeit (Mittlere Lösungshäufigkeit) Ziel: um 0,50	Reliabilität Cronbach's Alpha Ziel: > 0,70
Virtuelles Labor (Wissen)	0,50	0,78
Ökologisches Planspiel (Wissen)	0,62	0,63
Raumfahrtspiel (Wissen)	0,53	0,82
Projektaufgaben	0,68	0,74
Technisches Problemlösen	0,36	0,63
Analoges Problemlösen	0,33	0,79

diese Aufgaben in unterschiedlicher Weise ökologisches, ökonomisches und geographisches Hintergrundwissen mit ansprechen. (Dieser Test wurde inzwischen neu entwickelt.)

Von besonderer Bedeutung für die Fairness eines Leistungsvergleichs ist die Kontrolle des möglicherweise verzerrenden Einflusses konstrukt fremder Variablen. Da alle Aufgaben in spezifische situative Kontexte eingebettet sind, Ziel der Diagnose aber eine fächerübergreifende Problemkompetenz ist, sollten das spezielle Interesse am thematischen Kontext der Aufgabe und das Vorwissen möglichst niedrig mit dem Leistungsergebnis korrelieren. Diese beiden Einflussgrößen wurden durch Selbsteinschätzungen der Probanden erfasst. Wie Tabelle 2 zeigt, bleibt in der Tat das thematische Interesse ohne Bedeutung für das Leistungsergebnis. Der Einfluss des selbst eingeschätzten Vorwissens ist etwas größer, bleibt aber durchweg unter der Schwelle von  $r = .30$ , bei der man in der psychologischen Literatur von „schwacher Bedeutsamkeit“ spricht. Ein Vorteil der computergestützten Verfahren scheint zu sein, dass in den Ergebnismaßen dieser Instrumente (sowohl in Maßen des Wissenserwerbs als auch bei Strategiemmaßen, hier exemplarisch für das Raumfahrtspiel wiedergegeben) Vorwissen eine geringere Rolle spielt als bei schriftlichen Aufgaben. Die Ausnahme bildet erwartungsgemäß das virtuelle Labor, in dem geringfügig besser abschneidet, wer von dem Gegenstand des simulierten Experiments bereits gehört hat.

Ein relativer Nachteil der computergestützten Verfahren ist der durchweg signifikante Zusammenhang mit der Computererfahrung. Allerdings ist die Computererfahrung konfundiert mit anderen Personenmerkmalen, insbesondere den kognitiven Grundfähigkeiten, woraus sich auch die als signifikant bewerteten Korrelationen zwischen schriftlichen Testleistungen und Computererfahrung erklären. Kontrolliert man die intellektuellen Grundfähigkeiten, so sinkt die (Partial-)Korrelation mit der Computererfahrung für alle schriftlichen Problemlöseleistungen unter  $.15$  und für die computerbasierten Maße auf Werte zwischen  $.15$  und  $.30$ .

Tab. 2: Konstruktremde Einflüsse (Korrelationen mit Ergebnissen der Problemlöseverfahren)					
Verfahren		Thematisches Interesse	Selbst eingeschätztes Vorwissen	Geschlecht <sup>a</sup>	Computererfahrung
(Ziel: Korrelation von $r = < 0,3$ )					
Virtuelles Labor	Wissen	0,14	0,28 (0,24) <sup>b</sup>	n.s.	0,21
Ökologisches Planspiel	Wissen	n.s.	n.s.	0,28 (0,25) <sup>b</sup>	0,28
Raumfahrtspiel	Exploration	n.s.	0,18 (n.s.) <sup>b</sup>	0,32 (0,25) <sup>b</sup>	0,26
	Wissen	n.s.	n.s.	n.s.	0,18
	Steuerung	n.s.	0,16 (n.s.) <sup>b</sup>	0,35 (0,24) <sup>b</sup>	0,36
Projektaufgaben		n.s.	0,19	n.s.	0,16
Technisches Problemlösen		n.s.	0,22	0,16	0,22
Analoges Problemlösen		n.s.	0,31	n.s.	0,14
<sup>a</sup> Kodierung: 1 = männlich, 0 = weiblich <sup>b</sup> In Klammern angegeben ist der Zusammenhang bei Kontrolle der Computererfahrung					

Akzeptabel ist auch das Ausmaß der geschlechtsspezifischen Leistungsunterschiede. Zwar wird im Raumfahrtspiel den Jungen ein systematischeres Vorgehen attestiert; der Effekt ist jedoch nicht mehr bedeutsam, wenn man die Computererfahrung kontrolliert.

Das Gesamtfazit lautet, dass die Adaptation von Verfahren der Problemlöseforschung für die Untersuchung einer großen Stichprobe in der Zielpopulation der fünfzehnjährigen Schüler als gelungen bezeichnet werden kann. Die Verfahren haben einen angemessenen Schwierigkeitsgrad und eine weitgehend gute Reliabilität. Zusammenhänge mit Vorwissen, Computererfahrung und Geschlecht sind mit wenigen Ausnahmen nicht als bedeutsam anzusehen und dürften daher die Fairness des Tests nicht gefährden.

#### 4.2 Konvergente und diskriminante Validität

Die acht Indikatoren für Problemlösekompetenzen, die in Tabelle 2 aufgeführt sind, konnten in eine Strukturanalyse einbezogen werden. Ihre Interkorrelationen sind hoch signifikant; die Kennwerte liegen zwischen .32 und .80 bei einem Median von .55. Werte dieser Größenordnung sind allerdings nur ein unzulänglicher Beleg für die konvergente Validität der Problemlösetests. Unbefriedi-

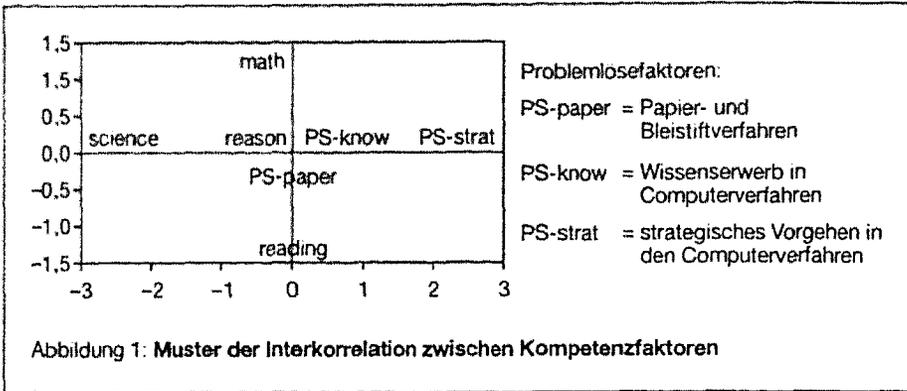
gend sind auch die Ergebnisse zur diskriminanten Validität: Die Korrelationen der Problemlösemaße mit dem Indikator der kognitiven Grundfähigkeit (.45–.72, Median = .58) und mit dem Ergebnis des internationalen Lesetests (.39–.77, Median = .61) besitzen ähnliche Größenordnungen wie die konvergenten Korrelationen. Es gelingt also nicht, eine allgemeine Problemlösekompetenz zu identifizieren und von anderen Konstrukten zu trennen. Wie in Abschnitt 2 erörtert, müssen wir die Problemlösefähigkeit in verschiedene Teilkonstrukte aufgliedern.

Eine systematische Antwort auf die Frage, welche Teilfaktoren der Problemlösekompetenz identifiziert und von anderen Kompetenzen unterschieden werden können, lässt sich mittels Strukturgleichungsmodellen geben. Neben den acht Problemlöseindikatoren gingen hier jeweils mehrere Indikatoren (Subtests) für die in PISA erfassten *Literacy*-Bereiche und die kognitive Grundfähigkeit ein. Die beste Anpassung an die empirischen Daten zeigte ein Modell, in dem sieben Kompetenzfaktoren unterschieden werden, die untereinander mehr oder weniger eng korrelieren. Diese Faktoren sind kognitive Grundfähigkeit (figurales und verbales schlussfolgerndes Denken), Lesekompetenz, mathematische Kompetenz, naturwissenschaftliche Kompetenz sowie drei getrennte Problemlösefaktoren: Problemlösen bei schriftlichen Aufgaben (Projektaufgaben, Technisches Problemlösen und Analoges Problemlösen), Wissenserwerb in computergestützten Systemen (Tests 1 bis 3) sowie Strategien beim Explorieren und Steuern eines komplexen Systems.<sup>2</sup> Dieses Modell mit sieben Faktoren wurde verglichen zum einen mit einem Modell, das lediglich einen Generalfaktor der kognitiven Leistungsfähigkeit vorsieht, sowie einem zweiten Vergleichsmodell, bei dem die drei Problemlösekompetenzen in einem Faktor zusammenfallen. Beide Alternativmodelle zeigten deutlich schlechtere Anpassung als das ursprüngliche Modell.

Von besonderem Interesse ist nun, die Korrelationen der drei Problemlösefaktoren untereinander bzw. mit den *Literacy*-Dimensionen zu untersuchen. Im Strukturgleichungsmodell werden diese Zusammenhänge messfehlerbereinigt geschätzt und sind daher – anders als die oben berichteten Korrelationen zwischen einzelnen Testresultaten – eindeutig als Aussagen über die Struktur der kognitiven Kompetenzen interpretierbar. Wir haben sie in Abbildung 1 grafisch veranschaulicht.

Die Abbildung zeigt das Ergebnis einer nichtmetrischen mehrdimensionalen Skalierung, basierend auf dem Muster der Interkorrelationen der sieben Faktoren. Je näher in dieser Darstellung zwei Faktoren angeordnet sind, umso höher ist ihr (im Strukturgleichungsmodell geschätzter) korrelativer Zusammenhang. Die zweidimensionale Lösung vermag die Daten sehr

2 Das entsprechende LISREL-Modell hat folgende Güteparameter:  $\chi^2 = 169.89$  ( $df = 83$ ;  $p < .001$ ),  $RMR = .056$ ,  $AGFI = .71$ . Diese Anpassungsgüte ist nicht voll befriedigend, lässt sich aber durch Modifikation des Modells nicht steigern. Vermutlich liegt dies daran, dass wir aufgrund unseres Designs gezwungen sind, die Matrix der Interkorrelationen zwischen allen Indikatorvariablen mit paarweisem Ausschluss zu schätzen. Ohne ein vollständiges Design, bei dem alle Versuchspersonen alle Tests bearbeiten – was jedoch bei der Vielzahl der Instrumente nicht realisierbar ist –, sind Analysen der Kovarianzstruktur relativ instabil. Für unsere Diskussion an dieser Stelle ist jedoch entscheidend, dass das Modell mit sieben korrelierten Faktoren den beiden Alternativmodellen ausweislich des signifikant niedrigeren  $\chi^2$ -Wertes deutlich überlegen ist.



gut wiederzugeben.<sup>3</sup> Das auffälligste Merkmal der Konfiguration ist die Positionierung des Faktors „Reasoning“ (d.h. der allgemeinen kognitiven Grundfähigkeit, gemessen durch KFT-Aufgaben) im Zentrum. Diesem sehr nahe liegt der Faktor „Problemlösekompetenz, erfasst in schriftlichen Verfahren (PS-paper)“. Damit ist nachgewiesen, dass unsere schriftlichen Problemlösetests – wie aufgrund der in Abschnitt 2 referierten theoretischen Überlegungen zu erwarten war – zu einem erheblichen Teil Problemlösekompetenz im Sinne von schlussfolgerndem Denken erfassen.

Die beiden Faktoren, die auf Computerverfahren basieren, sind jedoch weniger mit dem allgemeinen schlussfolgernden Denken konfundiert. Außerdem sind sie von den drei bereichsspezifischen PISA-Leistungsmaßen gut unterscheidbar: Die Faktoren für Mathematik, Lesekompetenz, naturwissenschaftliche Kompetenz und strategisches Vorgehen beim computergestützten Problemlösen liegen gleichmäßig um das Zentrum (Reasoning) herum verteilt; es handelt sich also hier um vier Kompetenzen, die sowohl untereinander als auch von der kognitiven Grundfähigkeit abgrenzbar sind.

Dieses Ergebnis stützt unserer Erwartung, dass Problemlösekompetenz kein einheitliches Konstrukt ist, sondern aus mehreren Teilkompetenzen besteht, die unterschiedliche Nähe zu kognitiven Grundfähigkeiten besitzen und prinzipiell von bereichsspezifischen Kompetenzen (Lesen, Mathematik, Naturwissenschaften) abgegrenzt werden können. Weitere Analysen, die auch bislang nicht ausgewertete Problemlöseverfahren berücksichtigen, müssen die genaue Struktur der Kompetenzen untersuchen.

#### 4.3 Zur Beeinflussung von Problemlösekompetenzen durch Effekte institutionalisierter Bildung

In einem ersten Schritt zur Untersuchung der Sensitivität von Problemlösemaßen für Effekte schulischer Bildung (in Abgrenzung etwa zu familiären und anderen psychosozialen Einflüssen) untersuchen wir, welcher Teil der Leistungsvarianz zwischen den Schülerinnen und Schülern überhaupt rechnerisch auf

3 Stress = .07, RSQ = .97.

Schuleffekte zurückzuführen ist. Hierbei ist es sinnvoll, die allgemeine kognitive Grundfähigkeit auszupartialisieren.<sup>4</sup>

Die Rechnungen ergaben zunächst, dass bei den schriftlichen Problemlöseaufgaben sich gut die Hälfte der Leistungsvarianz auf Schulunterschiede zurückführen lässt. Beispielsweise sind dies bei den Projektaufgaben 52 Prozent. Dieser Anteil entspricht recht genau der Varianzaufklärung durch schulbezogene Unterschiede bei den Tests zur Lesekompetenz (57 Prozent), zu Naturwissenschaften (54 Prozent) und zur Mathematik (64 Prozent). Die Computerverfahren sind hingegen mit ca. einem Drittel (für die Maße des Wissenserwerbs) bzw. einem Fünftel (für strategiebezogene Maße) erklärter Varianz weniger sensitiv für institutionelle Einflüsse.

Wenn man die kognitiven Grundfähigkeiten der Schüler auspartialisiert, bleiben für die Problemlösemaße nur noch 10 bis 20 Prozent institutionell zu erklärender Varianz übrig. Damit werden reine Selektionseffekte, die darauf beruhen, dass verschiedene Schulen bzw. Schulformen unterschiedliche Teilpopulationen mit je spezifischen intellektuellen Voraussetzungen erreichen, gleichsam aus den Daten herausgefiltert. Der verbleibende Varianzanteil ist nach wie vor statistisch signifikant und auch praktisch bedeutsam, sodass weitere Analysen, in denen Problemlösefähigkeiten in Abhängigkeit von schulbezogenen Merkmalen untersucht werden, sinnvoll erscheinen. Interessant (und theoretisch den Erwartungen entsprechend) ist der Befund, dass für die curricular bezogenen Bereiche auch nach Auspartialisierung der kognitiven Grundfähigkeit ein höherer institutionell zu erklärender Varianzanteil verbleibt: Bei der Lesekompetenz sind es 29 Prozent, bei den Naturwissenschaftstests 33 Prozent und bei den Mathematiktests sogar 43 Prozent.

Fazit: Jene Kompetenzen, die unterrichtsnah sind, d.h. durch das Curriculum und seine Implementierung an der Einzelschule stark beeinflusst werden, sind zu großen Teilen durch Schuleffekte erklärbar. Problemlösekompetenzen hingegen werden vermutlich weitestgehend durch außerschulische Faktoren determiniert. Allerdings lassen sich in einer Querschnittserhebung nur jene Bedingungsfaktoren identifizieren, die innerhalb der untersuchten Stichprobe von Schulen und Schülern variieren und Varianz in den gemessenen Kompetenzen erzeugen. Einflüsse, die innerhalb des Bildungssystems – etwa aufgrund von strukturellen Rahmenbedingungen – gleichförmig wirksam sind, lassen sich lediglich in Längsschnittstudien, bei gezielter quasi-experimenteller Kontrolle oder im internationalen Vergleich erkennen.

4 Die Methode der Wahl ist hier eine Mehrebenen-Analyse, bei der die Zusammenhänge zwischen Problemlöseleistungen und schlussfolgerndem Denken innerhalb jeder Schule ermittelt und getrennt davon auf der zweiten Datenebene, der Ebene der Schulen, die hier interessierenden Effekte geschätzt werden. Explorative Analysen mit dem Ansatz des *Hierarchical linear Modeling* (HLM-III) zeigten jedoch, dass Art und Enge des Zusammenhangs zwischen Problemlöseleistung und schlussfolgerndem Denken in den untersuchten Schulen nicht variierten. Es war daher möglich, die Varianzzerlegung unmittelbar auf der Ebene der Individualdaten durchzuführen. Diese Art der Varianzschätzung gibt angesichts der relativ geringen Zahl von Schulen robustere Schätzwerte.

## 5. Ausblick

Die hier berichteten ersten Erfahrungen im Rahmen von PISA-Deutschland können als Beleg dafür dienen, dass mit Einbeziehung von fächerübergreifenden und außerschulischen Problemlösekompetenzen das Spektrum der Schulleistungsdiagnostik im Rahmen eines *Large Scale Assessments* sinnvoll erweitert wird. Erst das so ergänzte Profil von fachbezogenen und außerfachlichen Kompetenzen wird es erlauben, die kognitive Entwicklung von Schülerinnen und Schülern im Zusammenspiel von schulischen und außerschulischen Entwicklungsbedingungen differenziert zu untersuchen.

Allerdings muss man sich von zu einfachen Vorstellungen über „Schlüsselqualifikationen“ lösen. Sowohl die theoretische Diskussion als auch die Daten unserer Konstruktvalidierungsstudie weisen darauf hin, dass „Problemlösekompetenz“ nicht als eindimensionales Konstrukt identifizierbar ist. Jenes Teilkonstrukt, das in unserer Studie durch verschiedene schriftliche Testverfahren operationalisiert wurde (Analoges Problemlösen, Technisches Problemlösen und Planungsaufgaben zu „Projekten“), ist zudem nicht von kognitiven Grundfähigkeiten („Reasoning“) abgrenzbar. Unsere durchaus komplexen, realitätsnahen, in außerfachliche Kontexte eingebundenen Testaufgaben messen – jedenfalls in der deutschen Stichprobe – offenbar nahezu dasselbe wie klassische psychometrische Instrumente zum „schlussfolgernden Denken“. Damit werden Problemlösetests nicht unbedingt überflüssig: Aus inhaltlichen Überlegungen und aus Akzeptanzgründen könnte eine kontextualisierte Messung des schlussfolgernden Denkens gewünscht sein. Außerdem bieten internationale Vergleichsuntersuchungen zusätzlich die Möglichkeit zu prüfen, ob die Dimensionalität fachübergreifender Kompetenzen und die Enge ihrer Beziehung zu curricular gebundenen Leistungen von Land zu Land unterschiedlich ausgeprägt ist. Möglicherweise lassen sich Bildungssysteme identifizieren, denen die Koppelung von fachgebundenem und fachübergreifendem Kompetenzerwerb unterschiedlich gut gelingt. „Projektaufgaben“ und ähnliche schriftliche Aufgabentypen sind daher in neueren Konzeptpapieren für eine Problemlösedagnostik auf internationaler Ebene enthalten (vgl. KLIEME/EBACH u. a. 2000; entsprechende Vorüberlegungen bestehen für den zweiten PISA-Zyklus im Jahr 2003). Die mit solchen schriftlichen Verfahren erfasste Teilkomponente der Problemlösekompetenz wird allerdings korrekterweise einschränkend als „analytisches Problemlösen“ gekennzeichnet.

Andere, weniger durch allgemeine Intelligenz bestimmte Aspekte von Problemlösekompetenz können möglicherweise durch computergestützte Bewertung des Explorations- und Steuerungsverhaltens in komplexen Problemlöse-szenarien identifiziert werden. Der entsprechende Befund aus unserer PISA-Erprobung muss nun durch weitere Analysen und insbesondere durch die PISA-Hauptuntersuchung im Jahr 2000, bei der das ökologische Planspiel und das Raumfahrtspiel nochmals leicht überarbeitet eingesetzt wurden, überprüft werden. Sollte der Befund sich als stabil erweisen, spräche dies für einen breiteren Einsatz von computersimulierten Szenarien zur Erfassung von Problemlösestrategien auch in *large scale assessments*. Ein wichtiges Forschungsthema bleibt die Frage, welche Rolle in solchen Lern- und Arbeitsumgebungen fach- bzw. bereichsspezifische Strategien in Abgrenzung zu fachübergreifenden Stra-

tegien spielen. Hierzu sind Untersuchungen erforderlich, welche die Nähe bzw. Distanz zu schulischen Lernbereichen – bei Konstanz sonstiger Anforderungsmerkmale – systematischer variieren, als es in unserer Studie möglich war.

## Literatur

- BAKER, E.L.: Final report for validation of problem-solving measures. Los Angeles: National center for Research on Evaluation, Standards, and Student Testing (CRESST) 1998.
- BANDURA, A.: Conclusion: Reflections on nonability determinants of competence. In: R. STERNBERG/J. KOLLIGIAN JR. (Hrsg.): *Competence considered*. New Haven/London (Yale University Press) 1990, S. 315–362.
- BAUMERT, J./BOS, W./LEHMANN, R. (Hrsg.): *TIMSS: Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn* (in 2 Bänden) Opladen (im Druck).
- BAUMERT, J./FEND, H./O'NEIL, H./PESCHAR, J.: Components of self-regulated learning. Unpublished Paper prepared for OECD-Network A. Paris 1998.
- BAUMERT, J./LEHMAN, R. u.a.: *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Opladen 1997.
- BINKLEY, M.R./STERNBERG, R./JONES, S./NOHARA, D.: An overarching framework for understanding and assessing life skills. In: *International Life Skills Survey (ILSS) Frameworks*. Washington: National Center of Education Statistics 1999, S. 1–23.
- BUCHNER, A.: Komplexes Problemlösen vor dem Hintergrund der Theorie finiter Automaten. In: *Psychologische Rundschau* 50 (1999), S. 206–212.
- DUBS, R.: Entwicklung von Schlüsselqualifikationen in der Berufsschule. *Handbuch der Berufsbildung*. Opladen 1995, S. 171–182.
- FRENSCH, P.A./FUNKE, J. (Hrsg.): *Complex problem solving. The european perspective*. Hillsdale/NJ 1995.
- FUNKE, J./FRITZ, A. (Hrsg.): *Neue Konzepte und Instrumente zur Planungsdiagnostik*. Bonn 1995.
- GLASER, R./SCHAUBLE, L./RAGHAVAN, K./ZEITZ, C.: Scientific Reasoning Across Different Domains. In: E. DE CORTE/M. C. LINN/H. MANDL/L. VERSCHAFFEL (Hrsg.): *Computer-based learning environments and problem solving*. Vol. 84 (1992), S. 345–371.
- HELLER, K./GAEDTKE, A.-K./WEINLÄDER, H.: *Kognitiver Fähigkeits-Test KFT 4-13*. Weinheim 1985.
- HELMKE, A./WEINERT, F.E.: Bedingungsfaktoren schulischer Leistungen. In: F.E. WEINERT (Hrsg.): *Enzyklopädie der Psychologie D/I/3. Psychologie des Unterrichts und der Schule*. Göttingen 1985, S. 71–176.
- KLIEME, E./BAUMERT, J./KÖLLER, O./BOS, W.: Mathematisch-naturwissenschaftliche Grundbildung: Erfassung und Skalierung von Kompetenzen. In: J. BAUMERT/W. BOS/R. LEHMANN (Hrsg.): *TIMSS – Mathematisch-naturwissenschaftliche Bildung am Ende der Sekundarstufe II*. Bd. 1. Opladen (im Druck).
- KLIEME, E./EBACH, J./DIDI, H.-J./HENSGEN, A./HEILMANN, K./MEISTER, K.-H.: Problemlösetest für die 7. Jahrgangsstufe. In: R. LEHMANN/I. STEINBACH (Hrsg.): *Hamburger Schulleistungstest für sechste und siebte Klassen*. Göttingen (im Druck).
- KLUWE, R./HAIDER, H.: Modelle zur internen Repräsentation komplexer technischer Systeme. In: *Sprache & Kognition* 9 (1990), S. 173–192.
- LAVE, J.: *Cognition in practice*. Boston/MA (Cambridge University Press) 1988.
- LEHMANN, R./PEEK, R.: Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Jahrgangsstufe an Hamburger Schulen (Forschungsbericht über die Voruntersuchungen). Berlin 1995.
- LEUTNER, D./SCHRETTENBRUNNER, H.: Entdeckendes Lernen in komplexen Realitätsbereichen: Evaluation des Computer-Simulationsspiels „Hunger in Nordafrika“. In: *Unterrichtswissenschaft* 17 (1989), S. 327–341.
- MAYER, R.E./WITTRICK, M.C.: Problem-solving transfer. In: D.C. BERLINER/R.C. CALFEE (Hrsg.): *Handbook of educational psychology*. New York (Simon & Schuster Macmillan) 1996, S. 47–62.
- MCCLELLAND, D.C.: Testing for competence rather than for „intelligence“. *American Psychologist* 28 (1973), S. 1–14.
- MERTENS, D.: Berufliche Flexibilität und adaptive Ausbildung in einer dynamischen Gesellschaft. In: R. JOCHIMSEN/U.E. SIMONIS (Hrsg.): *Theorie und Praxis der Infrastrukturpolitik*. Berlin: Schriften des Vereins für Sozialpolitik. Bd. 34 (1970).
- MUMFORD, M.D./SUPINSKI, E.P./BAUGHMAN, W.A./COSTANZA, D.P./THRELFALL, K.V.: *Process-based*

- measures of creative problem-solving skills: V. overall prediction. In: *Creativity Research Journal* 10 (1997), S. 73–85.
- NEWELL, A./SIMON, H.A.: *Human Problem Solving*. Englewood Cliffs (Prentice Hall) 1972.
- OECD: *Prepared for life? How to measure cross-curricular competencies*. Paris (Organization for Economic Co-operation and Development) 1997.
- OECD: *Measuring student knowledge and skills. A new framework for assessment*. Paris 1999.
- REETZ, L./SEYD, W.: *Curriculumtheorien im Bereich der Berufsbildung*. In: U. HAMEYER/K. FREY/H. HAFT (Hrsg.): *Handbuch der Curriculumforschung*. Weinheim/Basel 1983, S. 171–192.
- REIMANN, P.: *Novizen- und Expertenwissen*. In: F. KLIX/H. SPADA (Hrsg.): *Wissen 6*. Göttingen (1998), S. 335–367.
- ROBINSON, S.B.: *Bildungsreform als Revision des Curriculums*. Neuwied 1967.
- SOMEREN, M. W. VAN/REIMANN, P./BOSHUIZEN, H.P.A./DE JONG, T. (Hrsg.): *Learning with multiple representations*. Amsterdam (Pergamon) 1998.
- STERNBERG, R.J./KAUFMAN, J.C.: *Human Abilities*. *Annual Review of Psychology* 49 (1998), S. 479–502.
- SÜSS, H.-M.: *Intelligenz und komplexes Problemlösen: Perspektiven für eine Kooperation zwischen differentiell-psychometrischer und kognitionspsychologischer Forschung*. In: *Psychologische Rundschau* 50 (1999), S. 220–228.
- WEBER, H./WESTMEYER, H.: *Die Inflation der Intelligenzen*. In: E. STERN/J. GUTHKE (Hrsg.): *Perspektiven der Intelligenzforschung*. Lengerich (im Druck).
- WEINERT, F.E.: *Vermittlung von Schlüsselqualifikationen*. In: S. MATALIK/D. SCHADE (Hrsg.): *Entwicklungen in Aus- und Weiterbildung. Anforderungen, Ziele, Konzepte*. Baden-Baden 1998, S. 23–43.
- WEINERT, F.E.: *Konzepte der Kompetenz. Unveröffentlichtes Gutachten zum OECD-Projekt „Definition and Selection of Competencies“: Theoretical and Conceptual Foundations (DeSeCo)*. Manuskript: Max-Planck-Institut für psychologische Forschung. München 1999.

#### *Abstract*

Recent studies on school achievement such as the OECD-program PISA do not aim at curricular fields of achievement alone, but also at transdisciplinary competences such as problem-solving. The concept of a transdisciplinary competence leads to basic theoretical and methodological questions: selection and operationalization of competences in school studies, delimitation from specific disciplinary achievements and basic cognitive competences as well as the question of the significance of the school context for the development of competences. These questions are dealt with empirically within the framework of a construct-validation study by employing several paper-and-pencil and computer-supported procedures. As was to be expected, the problem-solving competence proved to be a multidimensional construct. The analytic ability to solve problems, analyzed by means of a paper-and-pencil test, is closely connected with basic cognitive competences, while computer-supported strategy measures seem to constitute a specific factor.

#### *Anschrift der Autoren*

PD Dr. Eckhard Klieme; Dipl.-Psych. Joachim Wirth, Max-Planck-Institut für Bildungsforschung, Lentzeallee 93, 14195 Berlin,  
 Prof. Dr. Joachim Funke, Prof. Dr. Peter Reimann, Universität Heidelberg,  
 Akademiestr. 3, 69117 Heidelberg,  
 Prof. Dr. Detlev Leutner, Universität Erfurt, Nordhäuser Str. 63, 99089 Erfurt