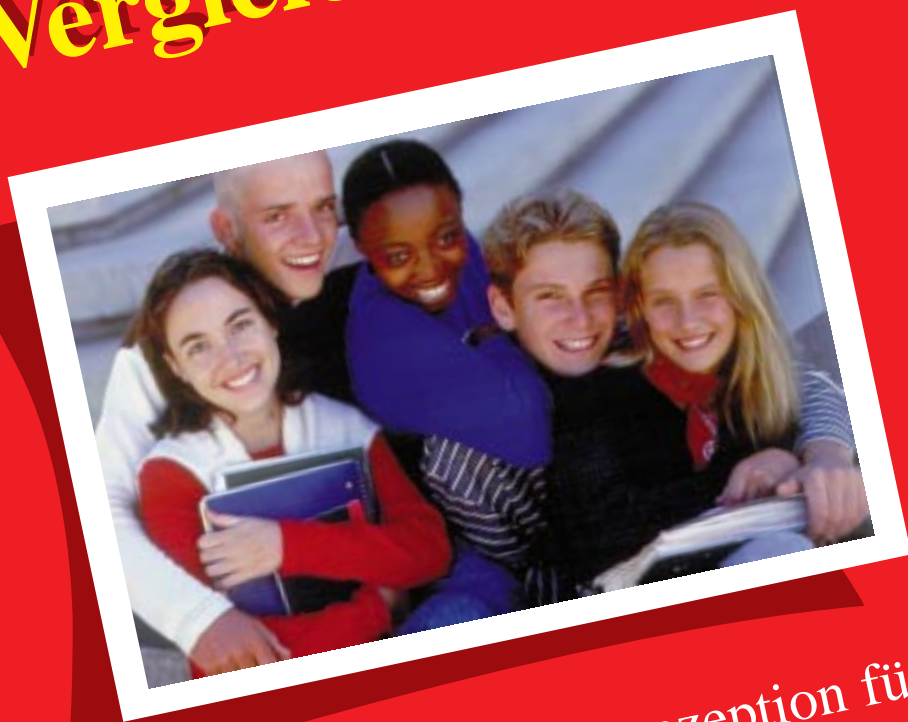


Schülerleistungen im internationalen Vergleich



Eine neue Rahmenkonzeption für
die Erfassung von Wissen
und Fähigkeiten

SCHÜLERLEISTUNGEN IM INTERNATIONALEN VERGLEICH

*Eine neue Rahmenkonzeption
für die Erfassung von Wissen und Fähigkeiten*

Herausgeber der deutschen Fassung: Deutsches PISA-Konsortium

Herausgeber der englischen und französischen Originalfassung: OECD

OECD PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT

ORGANISATION FÜR WIRTSCHAFTLICHE ZUSAMMENARBEIT UND ENTWICKLUNG

Gemäß Artikel 1 des am 14. Dezember 1960 in Paris unterzeichneten und am 30. September 1961 in Kraft getretenen Übereinkommens fördert die Organisation für Wirtschaftliche Zusammenarbeit und Entwicklung (OECD) eine Politik, die darauf ausgerichtet ist:

- in den Mitgliedsstaaten unter Wahrung der finanziellen Stabilität eine anhaltende Wirtschaftsentwicklung und Beschäftigung sowie einen steigenden Lebensstandard zu erreichen und dadurch zur Entwicklung der Weltwirtschaft beizutragen,
- in den Mitglieds- und Nichtmitgliedsstaaten, die in wirtschaftlicher Entwicklung begriffen sind, zu einem gesunden wirtschaftlichen Wachstum beizutragen und
- im Einklang mit internationalen Verpflichtungen auf multilateraler und nicht-diskriminierender Grundlage zur Ausweitung des Welthandels beizutragen.

Die Gründungsmitglieder der OECD sind: Belgien, Dänemark, Deutschland, Frankreich, Griechenland, Irland, Island, Italien, Kanada, Luxemburg, Niederlande, Norwegen, Österreich, Portugal, Schweden, Schweiz, Spanien, Türkei, Vereinigtes Königreich von Großbritannien und Nordirland und Vereinigte Staaten. Folgende Staaten wurden zu den nachstehend genannten Daten Mitglieder der OECD: Japan (28. April 1964), Finnland (28. Januar 1969), Australien (7. Juni 1971), Neuseeland (29. Mai 1973), Mexiko (18. Mai 1994), Tschechien (21. Dezember 1995) und Ungarn (7. Mai 1996). Die Kommission der Europäischen Gemeinschaften nimmt an den Tätigkeiten der OECD teil (Artikel 13 des Übereinkommens über die OECD).

Bildnachweis: PIX/Denis Boissavy

Zuerst veröffentlicht von der OECD in Englisch und Französisch unter den Titeln *Measuring Student Knowledge and Skills: A New Framework for Assessment/Mesurer les connaissances et compétences des élèves: Un nouveau cadre d'évaluation*. Copyright OECD, 1999.

Schülerleistungen im internationalen Vergleich: Eine neue Rahmenkonzeption für die Erfassung von Wissen und Fähigkeiten. Ausgabe in deutscher Sprache. Herausgegeben vom deutschen PISA-Konsortium.

Übersetzung: Hella Beister

Betreuung der Übersetzung: PISA-Projektleitung am Max-Planck-Institut für Bildungsforschung; Petra Stanat

Für die Qualität der Übersetzung und ihre Übereinstimmung mit dem Originaltext ist die OECD nicht verantwortlich.

© 2000 Max-Planck-Institut für Bildungsforschung, Lentzeallee 94, D-14195 Berlin.

Vervielfältigungen zu kommerziellen Zwecken sowie für umfangreiche nicht-kommerzielle Verbreitung bedarf der Genehmigung des Herausgebers.

ISBN 3-87985-078-X

VORWORT ZUR DEUTSCHEN AUSGABE

Mit dieser Broschüre wird die deutsche Übersetzung der internationalen Rahmenkonzeption für die Erfassung von Schülerleistungen im *Programme for International Student Assessment* (PISA) der OECD vorgelegt. Die PISA-Rahmenkonzeption wurde von internationalen Expertengruppen unter der Leitung von Raymond Adams, *Australian Council for Educational Research*, entwickelt. Daran waren auch Wissenschaftler aus Deutschland beteiligt, die sich auf die Zusammenarbeit mit deutschen Fachgruppen stützen konnten. Obwohl das vorliegende Konzept einen internationalen Abstimmungsprozess durchlaufen hat, sind die spezifischen theoretischen Orientierungen innerhalb der Erhebungsbereiche nicht verloren gegangen. Dies macht eine wichtige Stärke des Ansatzes aus. Unseres Erachtens stellt das Konzept einen internationalen Kompromiss dar, der auch in fachlicher Hinsicht tragfähig ist. Im Fortgang des Projekts soll die PISA-Rahmenkonzeption, die in Übereinkunft mit den Teilnehmerstaaten als *living framework* behandelt wird, weiterentwickelt werden.

Selbstverständlich können mit einer einzigen Studie nicht alle Ziele schulischer Bildung abgedeckt werden, und auch PISA wird nur einen Ausschnitt dessen erfassen, was deutsche Schulen anstreben und tatsächlich erreichen. In PISA haben sich die Teilnehmerstaaten für einen anwendungsorientierten Ansatz und die Erfassung grundlegender, fachbezogener und fächerübergreifender Kompetenzen entschieden. Damit betont die Studie einen funktionalen Aspekt moderner Grundbildung, ohne die Wichtigkeit anderer Bestandteile einer breiten Allgemeinbildung geringer zu bewerten.

Die internationale Rahmenkonzeption stellt das theoretische Dach dar, unter dem das deutsche PISA-Konsortium und die assoziierten Expertengruppen den nationalen Teil der PISA-Studie konzeptuell entwickelt und instrumentiert haben. In welcher Hinsicht PISA-Deutschland das gemeinsame internationale Vorhaben ergänzt, erweitert und durch spezifische Akzentsetzungen an die institutionellen Gegebenheiten in Deutschland anzupassen versucht, darüber gibt die nationale Rahmenkonzeption Auskunft, deren Komponenten auf der Homepage des deutschen PISA-Konsortiums dargestellt sind (<http://www.mpib-berlin.mpg.de/PISA/>).

Für das deutsche PISA-Konsortium und die assoziierten Expertengruppen

Jürgen Baumert

Berlin, im Januar 2000

VORWORT

Die von der OECD getragene internationale Schulleistungsstudie PISA (Programme for International Student Assessment) ist Ausdruck einer neuen Selbstverpflichtung der OECD-Mitgliedsstaaten, sich durch Messung von Schülerleistungen auf der Grundlage einer gemeinsamen internationalen Rahmenkonzeption ein Bild von der Leistungsfähigkeit ihrer Bildungssysteme zu verschaffen. PISA ist ein Kooperationsprojekt, in dem übergreifende Entscheidungen gemeinschaftlich und auf der Basis gemeinsamer politischer Interessen von den Regierungen aller Teilnehmerstaaten getroffen werden, die auch die politische Verantwortung für das Projekt tragen. Darüber hinaus wird in PISA die wissenschaftliche Kompetenz aus den beteiligten Ländern zusammengeführt. Es wurden Arbeitsgruppen mit Experten aus den Teilnehmerstaaten gebildet, die gewährleisten sollen, dass die politischen Zielsetzungen von PISA mit der größtmöglichen fachwissenschaftlichen und verfahrenstechnischen Kompetenz auf dem Gebiet des internationalen Leistungsvergleichs verknüpft werden. Durch ihre Beteiligung an diesen Expertengruppen stellen die Länder sicher, dass die im Rahmen von PISA eingesetzten Instrumente zur Leistungsmessung international valide sind und zugleich dem kulturellen und curricularen Kontext der OECD-Mitgliedsländer Rechnung tragen, dass sie über sehr gute messtechnische Eigenschaften verfügen und sowohl authentisch als auch bildungspolitisch relevant sind.

Zuständig für Design und Implementierung der Erhebungen ist, bei übergreifender Verantwortung des OECD-Sekretariats, ein internationales Konsortium, in dem unter der Leitung des Australian Council for Educational Research (ACER) die folgenden Forschungseinrichtungen zusammenarbeiten: Netherlands National Institute for Educational Measurement (CITO); Educational Testing Service (ETS), USA; National Institute for Educational Research (NIER), Japan; und WESTAT, USA¹.

Die vorliegende Veröffentlichung enthält die Rahmenkonzeption, die den Erhebungen von OECD/PISA zugrunde liegt: Sie definiert die Bereiche, die erfasst werden sollen, und erklärt, was gemessen wird und wie es gemessen wird. Sie beschreibt außerdem die Kontexte und Beschränkungen, die für die Studie bestimmend sind. Verantwortlich für die Veröffentlichung der Rahmenkonzeption ist der Generalsekretär der OECD.

¹ Die hier aufgeführten Institutionen bilden das aktuelle Konsortium (Stand: Januar 2000). Zum Zeitpunkt des Erscheinens der englischen/französischen Originalfassung der Rahmenkonzeption war zusätzlich der Service de Pédagogie Expérimentale (SPE), Universität Liège, Belgien, Teil des Konsortiums. ETS und NIER sind im Juli 1999 neu hinzugekommen.

BETEILIGTE PERSONEN UND INSTITUTIONEN

Die PISA-Rahmenkonzeption wurde von Expertengruppen unter der Leitung von Raymond Adams von ACER entwickelt. Vorsitzender der Expertengruppe im Bereich Lesekompetenz war Dr. Irwin Kirsch vom Educational Testing Service, den Vorsitz der Mathematikgruppe hatte Professor Jan de Lange von der Universität Utrecht, und die Expertengruppe für den Bereich Naturwissenschaften stand unter der Leitung von Professor Wynne Harlen vom Scottish Council for Research. Die Mitglieder der Expertengruppen werden im Anhang 1 aufgeführt. In die Rahmenkonzeption sind darüber hinaus Stellungnahmen von Expertengruppen in den einzelnen Teilnehmerländern eingeflossen. Die Rahmenkonzeption wurde im Dezember 1998 von den im „Board of Participating Countries“ (BPC) vertretenen OECD-Regierungen verabschiedet. Diese Veröffentlichung wurde von der Abteilung für Statistik und Indikatoren des OECD-Direktorats für Bildung, Beschäftigung, Arbeit und Soziales erstellt, hauptsächlich von Andreas Schleicher.

INHALTSVERZEICHNIS

Einleitung	9
Das Design von OECD/PISA 2000	11
Grundlegende Merkmale von OECD/PISA	11
Unterschiede zwischen OECD/PISA und anderen internationalen Schulleistungsstudien	12
Was in den einzelnen Bereichen gemessen wird	14
Durchführung der Tests und Berichterstattung über die Ergebnisse	17
Die Kontextfragebögen und ihre Verwendung	18
OECD/PISA - ein Instrument, das sich weiterentwickelt	19
Die Entwicklung von OECD/PISA und seiner Rahmenkonzeptionen: Ein Kooperationsprojekt	21
<i>Kapitel 1: Lesekompetenz (Reading Literacy)</i>	23
Definition des Gegenstandsbereichs	23
Organisation des Gegenstandsbereichs und Aufgabenmerkmale	25
Erhebungsstruktur	40
Skalenentwicklung	42
Sonstiges	42
<i>Kapitel 2: Mathematische Grundbildung (Mathematical Literacy)</i>	47
Definition des Gegenstandsbereichs	47
Organisation des Gegenstandsbereichs	48
Aufgabenmerkmale	57
Erhebungsstruktur	61
Skalenentwicklung	62
Sonstiges	63
<i>Kapitel 3: Naturwissenschaftliche Grundbildung (Scientific Literacy)</i>	65
Definition des Gegenstandsbereichs	65
Organisation des Gegenstandsbereichs	67
Aufgabenmerkmale	73
Erhebungsstruktur	75
Skalenentwicklung	78
Sonstiges	79
Literaturhinweise	81
<i>Anhang 1: Mitglieder der funktionalen Expertengruppen</i>	85
<i>Anhang 2: Überlegungen für zukünftige Erhebungszyklen von OECD/PISA</i>	89

EINLEITUNG

Wie gut sind Jugendliche auf die Herausforderungen der Zukunft vorbereitet? Sind sie in der Lage, ihre Ideen und Vorstellungen nutzbringend zu analysieren, zu begründen und zu kommunizieren? Verfügen sie über die Voraussetzungen für lebenslanges Lernen? Eltern, Schülerinnen und Schüler, die Öffentlichkeit sowie Lehrkräfte und Schulverwaltungen haben einen Anspruch auf solche Informationen.

Viele Bildungssysteme untersuchen die Leistungen ihrer Schülerinnen und Schüler regelmäßig, um Antworten auf diese Fragen zu erhalten. Internationale Vergleichsuntersuchungen können die nationale Perspektive erweitern und bereichern, indem sie Auskunft über das Leistungsniveau von Schülerinnen und Schülern in anderen Ländern geben und somit den Kontext für die Interpretation der eigenen Ergebnisse erweitern. Sie können Orientierungshilfen für die Unterrichtsgestaltung und für das Lernverhalten von Schülerinnen und Schülern geben sowie Einsichten in die Stärken und Schwächen eines Curriculums vermitteln. In Verbindung mit geeigneten Anreizen können sie Schülerinnen und Schüler zu besserem Lernen, Lehrkräfte zu besserem Unterricht und Schulen zu effizienterem Arbeiten motivieren. Darüber hinaus geben sie den Ländern Instrumente an die Hand, die diese zur Erfassung von Leistungsniveaus einsetzen können, selbst wenn Schulen stärker selbstverwaltet sind und z. B. in Partnerschaft mit den Gemeinden arbeiten.

Regierungen und die allgemeine Öffentlichkeit benötigen verlässliche und international vergleichbare Nachweise über die Leistungsfähigkeit ihrer Bildungssysteme. Um diesem Anspruch zu genügen, hat die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) die internationale Schulleistungsstudie PISA (Programme for International Student Assessment) initiiert. OECD/PISA wird regelmäßig und zügig schulpolitisch relevante und international vergleichbare Indikatoren für Schülerleistungen bereitstellen. Die Leistungsmessungen werden bei 15-jährigen Schülerinnen und Schülern vorgenommen. Mit Hilfe der Indikatoren soll eingeschätzt werden, inwieweit die Bildungssysteme der Teilnehmerstaaten ihre Schülerinnen und Schüler auf lebenslanges Lernen und auf die Übernahme von konstruktiven Rollen als Bürger ihrer Gesellschaft vorbereiten.

OECD/PISA ist Ausdruck einer neuen Selbstverpflichtung der Regierungen der OECD-Mitgliedsstaaten, sich durch Messung von Schülerleistungen auf der Grundlage einer gemeinsamen, international abgestimmten Rahmenkonzeption ein Bild von der Leistungsfähigkeit ihrer Bildungssysteme zu machen. Zwar werden verschiedene Personengruppen die Ergebnisse der Studie für unterschiedliche Zwecke nutzen. Das Hauptziel der Entwicklung und Durchführung dieser groß angelegten internationalen Untersuchung ist jedoch die Gewinnung von empirisch gesicherten Informationen, die als Grundlage von schulpolitischen Entscheidungen dienen können.

Die Ergebnisse der OECD-Schulleistungserhebung, die alle drei Jahre zusammen mit anderen Indikatoren der Bildungssysteme veröffentlicht werden sollen, werden die Schulpolitiker der einzelnen Länder in die Lage versetzen, die Leistungsfähigkeit ihrer Bildungssysteme mit denen anderer Staaten zu vergleichen. Die Befunde werden außerdem neue Anstöße für Bildungsreformen und für die Qualitätssicherung in Schulen geben, insbesondere wenn sich zeigt, dass Schulen oder Bildungssysteme bei gleichen Ausgangsdaten zu deutlich unterschiedlichen Ergebnissen gelangen. Darüber hinaus bieten sie eine Grundlage für die bessere Einschätzung und Kontrolle der Effektivität von Bildungssystemen im nationalen Rahmen.

OECD/PISA ist ein kooperativer Prozess, in dem die wissenschaftliche Kompetenz aus den Teilnehmerstaaten zusammengeführt wird. Übergreifende Entscheidungen werden gemeinschaftlich und auf der Basis gemeinsamer politischer Interessen von den Regierungen dieser Staaten getroffen. Im Folgenden wird zunächst ein kurzer Überblick über das Design von OECD/PISA und eine Zusammenfassung der wichtigsten Merkmale der Leistungsmessungen gegeben. Anschließend wird der kooperative Ansatz und seine Nutzung für die Entwicklung der Rahmenkonzeption als verbindliche Grundlage für die OECD-Leistungsstudie beschrieben. In den weiteren Abschnitten wird diese Rahmenkonzeption dann im Einzelnen ausgeführt: die Bereiche, die Gegenstand der Leistungsmessungen sein sollen, werden definiert und es wird erklärt, was gemessen wird und wie es gemessen wird. Zugleich werden der Kontext und die Rahmenbedingungen beschrieben, die für OECD/PISA bestimmend sind.

Was ist OECD/PISA? Die wichtigsten Merkmale im Überblick

Grundlegendes

- PISA ist eine international standardisierte Leistungsmessung, die von den Teilnehmerstaaten gemeinsam entwickelt wurde und an 15-jährigen Schülerinnen und Schülern in ihren Schulen durchgeführt werden soll.
- Teilnehmer sind 32 Staaten, davon 28 Mitgliedsstaaten der OECD.
- In jedem Land werden zwischen 4.500 und 10.000 Schülerinnen und Schüler getestet.

Inhalt

- PISA erfasst drei Bereiche: Lesekompetenz (reading literacy), mathematische Grundbildung (mathematical literacy) und naturwissenschaftliche Grundbildung (scientific literacy).
- Die Definition der Bereiche soll so erfolgen, dass nicht nur die Beherrschung des im Curriculum vorgesehenen Lehrstoffs abgedeckt wird, sondern auch wichtige Kenntnisse und Fähigkeiten, die man im Erwachsenenleben benötigt, erfasst werden. Die Untersuchung von fächerübergreifenden Kompetenzen ist integraler Bestandteil von PISA.
- Das Hauptaugenmerk liegt auf der Beherrschung von Prozessen, dem Verständnis von Konzepten sowie auf der Fähigkeit, innerhalb eines Bereichs mit unterschiedlichen Situationen umzugehen.

Methoden

- Die Messung erfolgt mit Papier-und-Bleistift-Tests; die Testzeit beträgt insgesamt zwei Stunden.
- Die Tests bestehen aus einer Mischung von Multiple-Choice-Aufgaben und Fragen, für die die Schülerinnen und Schüler eigene Antworten ausarbeiten müssen. Die Items sind in Gruppen zusammengefasst, die sich jeweils auf eine Beschreibung einer realitätsnahen Situation beziehen.
- Insgesamt werden Items für eine Testdauer von 7 Stunden eingesetzt, von denen die Schülerinnen und Schüler jeweils unterschiedliche Kombinationen bearbeiten.
- Die Schülerinnen und Schüler beantworten außerdem einen Schülerfragebogen mit Hintergrundfragen über sie selbst, und die Schulleiter erhalten einen Fragebogen mit Fragen zu ihrer Schule. Die Bearbeitung des Schülerfragebogens wird 20 - 30 Minuten, die des Schulfragebogens etwa 30 Minuten in Anspruch nehmen.

Erhebungszyklus

- Die erste Erhebung wird im Jahr 2000 stattfinden; die ersten Ergebnisse werden im Jahr 2001 vorliegen. Danach erfolgen die Erhebungen in einem Dreijahreszyklus.
- In jedem Zyklus wird ein „Hauptbereich“ gründlicher getestet, dem dann zwei Drittel der Testzeit zugeteilt werden; in den beiden anderen Bereichen werden jeweils nur zusammenfassende Leistungsprofile erfasst. Die Hauptbereiche sind: Lesekompetenz im Jahr 2000, mathematische Grundbildung im Jahr 2003 und naturwissenschaftliche Grundbildung im Jahr 2006.

Ergebnisse

- Ein Profil der Kenntnisse und Fähigkeiten von Schülerinnen und Schülern gegen Ende der Pflichtschulzeit.
- Kontextbezogene Indikatoren, mit denen ein Zusammenhang zwischen den Ergebnissen und den Merkmalen von Jugendlichen und Schulen hergestellt wird.
- Trendindikatoren, die zeigen, wie sich die Ergebnisse im Zeitverlauf ändern.

DAS DESIGN VON OECD/PISA 2000

Grundlegende Merkmale von OECD/PISA

Die OECD-Leistungsmessungen werden die Bereiche Lesekompetenz, mathematische Grundbildung und naturwissenschaftliche Grundbildung abdecken. Die Schülerinnen und Schüler werden außerdem einen Hintergrundfragebogen beantworten. Weitere Hintergrundinformationen werden von den Schulen eingeholt. Die ersten Leistungsmessungen werden im Jahr 2000 durchgeführt, erste Ergebnisse werden 2001 vorliegen. Insgesamt planen 32 Staaten, davon 28 OECD-Mitgliedsstaaten, an OECD/PISA teilzunehmen. Zusammen repräsentieren diese Staaten mehr als ein Viertel der Weltbevölkerung und damit mehr als jede andere internationale Erhebung.

Da es das Ziel von OECD/PISA ist, den kumulativen Ertrag von Bildungssystemen gegen Ende der Pflichtschulzeit zu erfassen, werden die Tests mit 15-jährigen Schülerinnen und Schülern in allgemeinbildenden und berufsbildenden Schulen durchgeführt. In der Regel sollen in jedem Land 4.500 - 10.000 Schülerinnen und Schüler getestet werden. Diese Stichprobenbasis wird es erlauben, die Ergebnisse nach bestimmten Schülermerkmalen getrennt zu analysieren.

Obwohl die Bereiche Lesekompetenz, mathematische Grundbildung und naturwissenschaftliche Grundbildung bestimmten Schulfächern entsprechen, wird die OECD-Studie nicht in erster Linie untersuchen, wie gut die Schülerinnen und Schüler den im Curriculum vorgesehenen Lehrstoff beherrschen. Vielmehr soll bestimmt werden, inwieweit die Jugendlichen in diesen Bereichen die allgemeineren Kenntnisse und Fähigkeiten erworben haben, die sie später als Erwachsene benötigen werden. Die Erfassung von fächerübergreifenden Kompetenzen ist daher integraler Bestandteil von OECD/PISA. Die wichtigsten Gründe für diesen relativ umfassenden Messansatz sind die Folgenden:

- Erstens: Obwohl der Erwerb von spezifischem Wissen im schulischen Lernen wichtig ist, hängt die Anwendung dieses Wissens im Erwachsenenleben entscheidend von den allgemeineren Fähigkeiten und Kenntnissen ab, die der Einzelne erworben hat. Im Bereich Lesen besteht die zentrale Kompetenz in der Fähigkeit, schriftliches Material zu interpretieren und über Inhalt und Eigenschaften von Texten zu reflektieren. Im Bereich Mathematik kommt es bei der Anwendung von mathematischen Kompetenzen im täglichen Leben mehr auf die Fähigkeit an, quantitativ zu argumentieren und Beziehungen oder Abhängigkeiten zu erfassen, als auf die Fähigkeit, die für Schulbücher typischen Fragen zu beantworten. Im Bereich Naturwissenschaften sind für die naturwissenschaftlichen Probleme, die in der Welt der Erwachsenen diskutiert werden, spezifische Kenntnisse, etwa die Namen bestimmter Pflanzen oder Tiere, weniger wichtig als ein Verständnis von umfassenderen Konzepten und Themen wie Energieverbrauch, Artenvielfalt und menschliche Gesundheit.
- Zweitens: Eine Konzentration auf die Curriculuminhalte würde bei einer internationalen Vergleichsstudie dazu führen, dass nur solche Aspekte gemessen werden können, die in den Curricula von allen oder den meisten Ländern enthalten sind. Dies würde zwangsläufig zahlreiche Kompromisse erfordern und zu Messungen führen, die viel zu eingeschränkt wären, um für Regierungen, die sich ein Bild von den Stärken und den Innovationen der Bildungssysteme anderer Länder verschaffen wollen, einen Wert zu haben.
- Drittens: Es gibt bestimmte allgemeine Fähigkeiten, deren Entwicklung für Schülerinnen und Schüler wesentlich ist. Dazu gehören Kommunikationsfähigkeit, Anpassungsfähigkeit, Flexibilität, Problemlösefähigkeit und die Fähigkeit zur Nutzung von Informationstechnologien. Diese Kompetenzen werden fächerübergreifend entwickelt und müssen daher auch fächerübergreifend gemessen werden.

OECD/PISA beruht auf einem dynamischen Modell des lebenslangen Lernens. Nach diesem Modell müssen neue Kenntnisse und Fähigkeiten, die für die erfolgreiche Anpassung an veränderte Gegebenheiten erforderlich sind, kontinuierlich über die gesamte Lebensspanne hinweg erworben werden. Nicht alles, was sie als Erwachsene benötigen werden, können Schülerinnen und Schüler in der Schule lernen. Was sie daher erwerben müssen, sind die Voraussetzungen für erfolgreiches Lernen im späteren Leben. Diese Voraussetzungen sind sowohl kognitiver als auch motivationaler Natur. Die Jugendlichen müssen befähigt werden, ihren eigenen Lernprozess zu organisieren und zu regulieren, selbstständig und in Gruppen zu lernen und Schwierigkeiten im Lernprozess zu überwinden. Dies setzt voraus, dass sie sich ihrer eigenen Denkprozesse sowie ihrer Lernstrategien und -methoden bewusst sind. Darüber hinaus ist zu berücksichtigen, dass Weiterbildung und der Erwerb von zusätzlichem Wissen zunehmend in Situationen erfolgen werden, in denen Menschen zusammenarbeiten und voneinander abhängig sind. Um solche

Aspekte erfassen zu können, werden im ersten Zyklus der OECD/PISA-Studie im Jahr 2000 auch Instrumente entwickelt und eingesetzt, mit denen Informationen über selbstreguliertes Lernen erfasst werden können.

OECD/PISA ist keine einmalige länderübergreifende Messung der Lesekompetenz, der mathematischen Grundbildung und der naturwissenschaftlichen Grundbildung von 15-Jährigen. Es ist ein fortlaufendes Programm, das alle drei Jahre Daten für jeden dieser Bereiche erheben wird. Langfristig wird damit ein Bestand an Informationen aufgebaut, der eine Beobachtung von Entwicklungstrends im Wissens- und Kompetenzbestand von Schülerinnen und Schülern aus den verschiedenen Ländern und aus verschiedenen demographischen Untergruppen erlaubt. In jeder Erhebungswelle wird ein anderer Bereich detailliert untersucht, der dann fast zwei Drittel der Gesamttestzeit in Anspruch nehmen wird. Dieser „Hauptbereich“ wird im Jahr 2000 die Lesekompetenz sein, im Jahr 2003 die mathematische Grundbildung und im Jahr 2006 die naturwissenschaftliche Grundbildung. Mit diesem Zyklus ist dafür gesorgt, dass in jedem dieser Bereiche alle neun Jahre eine gründliche Leistungsanalyse und alle drei Jahre ein „check-up“ stattfindet.

Insgesamt beträgt die Testzeit zwei Stunden. Alle eingesetzten Items zusammen werden jedoch Informationen liefern, die einem fast siebenstündigen Test entsprechen. Die Gesamtmenge der Items wird auf eine Reihe unterschiedlich zusammengesetzter Itemgruppen verteilt. Jede Itemgruppe wird von einer ausreichend großen Zahl von Jugendlichen bearbeitet, so dass das Leistungsniveau der Schülerinnen und Schüler in jedem Land sowie in relevanten Untergruppen innerhalb der Länder (z.B. Jugendliche aus unterschiedlichen sozialen und ökonomischen Kontexten) für jedes Item geschätzt werden kann. Die Schülerinnen und Schüler werden außerdem etwa 20 Minuten mit der Beantwortung des Kontextfragebogens zubringen.

Die Erhebung wird verschiedene Arten von Indikatoren liefern:

- Basisindikatoren, die ein Grundprofil der Kenntnisse und Fähigkeiten der Schülerinnen und Schüler liefern;
- Kontextindikatoren, die zeigen, wie diese Fähigkeiten mit wichtigen demographischen, sozialen, wirtschaftlichen und das Bildungssystem betreffenden Variablen zusammenhängen;
- Trendindikatoren, die sich aus dem fortlaufenden, zyklischen Charakter der Datenerhebung ergeben und Veränderungen des Leistungsniveaus, der Leistungsverteilung und der Zusammenhänge zwischen schüler- bzw. schulbezogenen Hintergrundvariablen und Leistungen im Zeitverlauf zeigen.

Obwohl Indikatoren ein geeignetes Mittel sind, um auf wichtige Probleme aufmerksam zu machen, können sie normalerweise keine Antworten auf politische Fragen liefern. Deshalb wurde im Rahmen von OECD/PISA außerdem ein Plan für schulpolitische Analysen entwickelt, der über die Berichterstattung zum Stand der Indikatoren hinausgeht.

Am ersten OECD/PISA-Zyklus beteiligen sich folgende Länder: Australien, Belgien, Brasilien, China, Dänemark, Deutschland, Finnland, Frankreich, Griechenland, Irland, Island, Italien, Japan, Kanada, Korea, Lettland, Luxemburg, Mexiko, die Niederlande, Neuseeland, Norwegen, Österreich, Polen, Portugal, die Russische Föderation, Spanien, Schweden, die Schweiz, die Tschechische Republik, Ungarn, das Vereinigte Königreich und die USA.

Unterschiede zwischen OECD/PISA und anderen internationalen Schulleistungsstudien

OECD/PISA ist nicht die erste internationale Schulleistungsstudie. Ähnliche Studien wurden in den letzten 40 Jahren mehrfach durchgeführt, vor allem von der International Association for the Evaluation of Educational Achievement (IEA) und im Rahmen vom International Assessment of Educational Progress (IAEP) des Education Testing Service. Qualität und Umfang dieser Studien wurden im Lauf der Jahre immer weiter erhöht. Dennoch geben sie nur partielle und sporadische Auskunft über Schülerleistungen in einer begrenzten Zahl von Fächern. Die drei von der IEA in den Bereichen Naturwissenschaften und Mathematik durchgeführten Studien liefern zwar Hinweise auf Veränderungen in den letzten 30 Jahren. Da jedoch an den frühen Untersuchungen eine relativ geringe Anzahl von Ländern teilnahmen und weil die eingesetzten Tests nur bedingt vergleichbar sind, ergeben sie kein umfassendes Bild.

Wichtiger aber ist, dass in diesen Studien immer nur Leistungen erfasst wurden, die sich direkt auf den in Lehrplänen vorgesehenen Stoff beziehen, und noch dazu lediglich auf solche Bestandteile von Lehrplänen, die allen Teilnehmerstaaten gemeinsam sind. Aspekte des Curriculums, die nur in einem Land oder in einer geringen Anzahl von Ländern vertreten sind, wurden ungeachtet ihrer Bedeutung für die beteiligten Staaten gewöhnlich nicht berücksichtigt.

OECD/PISA zeichnet sich durch einen neuen Ansatz aus, der sich in wichtigen Punkten von anderen Vorgehensweisen unterscheidet:

- *Ursprung:* Die Studie geht zurück auf die Initiative von Regierungen und soll deren schul- und bildungspolitischen Interessen dienen.
- *Regelmäßigkeit:* Aufgrund ihrer Entscheidung, Leistungsmessungen durchzuführen, die sich auf mehrere Bereiche erstrecken und alle drei Jahre aktualisiert werden sollen, haben die beteiligten Länder die Möglichkeit, Fortschritte im Erreichen zentraler Bildungsziele regelmäßig und verlässlich zu überprüfen.
- *Getestete Altersgruppe:* Leistungsmessungen zum Ende der Pflichtschulzeit liefern nützliche Hinweise auf die Leistungsfähigkeit von Bildungssystemen. Zwar setzen die meisten Jugendlichen in den OECD-Staaten ihre Erstausbildung über das Alter von 15 Jahren hinaus fort, in diesem Alter stehen sie jedoch normalerweise kurz vor dem Ende der ersten Phase der Schulbildung, in der für alle Jugendlichen ein mehr oder weniger einheitliches Curriculum gilt. Daher ist es sinnvoll, auf dieser Stufe festzustellen, inwieweit die Schülerinnen und Schüler Kenntnisse und Fähigkeiten erworben haben, die ihnen in der Zukunft nützen werden, unter anderem auch für weitere, individuelle Bildungswege, die sie möglicherweise einschlagen werden.
- *Erfasste Kenntnisse und Fähigkeiten:* Wissen und Fähigkeiten werden im Rahmen von PISA nicht in erster Linie im Sinne eines gemeinsamen Nenners der Curricula aller Teilnehmerländer definiert, sondern im Hinblick auf Fähigkeiten, die als wesentlich für das künftige Leben gelten. Dies ist das eigentlich grundlegende und zugleich anspruchsvollste innovative Merkmal von OECD/PISA. Eine allzu strikte Trennung der „schulischen“ Fähigkeiten von den Fähigkeiten „für das Leben“ wäre willkürlich, denn Schulen hatten schon immer das Ziel, Jugendliche auf das Leben vorzubereiten. Aber dennoch ist die Unterscheidung wichtig. Schulische Curricula bestehen traditionell aus Modulen von Informationen und Techniken, die vermittelt werden sollen. Dabei geht es weniger um die Fähigkeiten, die in jedem Fach und Bereich zur späteren Anwendung im Erwachsenenleben entwickelt werden müssen, und noch weniger um die fächerübergreifend entwickelten, allgemeinen Fähigkeiten des Problemlösens und der Anwendung von eigenen Ideen und Einsichten in konkreten Situationen, denen man im Leben begegnet. OECD/PISA schließt das im Curriculum vorgesehene Wissen und Verständnis nicht aus, testet es aber in erster Linie im Hinblick darauf, inwieweit allgemeinere Konzepte und Fähigkeiten erworben wurden, die für die Anwendung dieses Wissens benötigt werden. Darüber hinaus beschränkt sich PISA auch nicht auf den kleinsten gemeinsamen Nenner des in den Schulen aller Teilnehmerländer vorgesehenen Lehrstoffs.

Diese Ausrichtung der Tests auf die Beherrschung von Prozessen und umfassenden Konzepten ist insbesondere für das Interesse der Länder an einer Entwicklung ihres Humankapitals von Bedeutung. Dieses wird von der OECD wie folgt definiert:

„Wissen, Fähigkeiten, Kompetenzen und sonstige Eigenschaften von Individuen, die für das persönliche, soziale und wirtschaftliche Wohlergehen relevant sind.“

Schätzungen des Bestands an Humankapital bzw. des Kompetenzbestands konnten bisher bestenfalls mittelbar, aus Größen wie beispielsweise dem Niveau von Bildungsabschlüssen, abgeleitet werden. Die Unzulänglichkeit dieser Art von Maßen wird besonders deutlich, wenn sich das Interesse am Humankapital auch auf solche Eigenschaften erstreckt, die eine vollständige soziale und demokratische Teilnahme am gesellschaftlichen Leben ermöglichen und Menschen zu lebenslangem Lernen befähigen.

Mit der direkten Messung von Wissen und Fähigkeiten zum Ende der Pflichtschulzeit untersucht OECD/PISA, inwieweit Jugendliche auf das Leben als Erwachsene vorbereitet werden und in gewissem Maße auch, wie leistungsfähig die Bildungssysteme sind. Leistung soll dabei in Bezug auf grundlegende (und von der Gesellschaft definierte) Ziele und nicht in Bezug auf das Lehren und Lernen eines bestimmten Stoffs gemessen werden. Wenn man will, dass sich Schulen und Bildungssysteme den modernen Herausforderungen widmen, sind solche authentischen Leistungsmaße notwendig.

Was in den einzelnen Bereichen gemessen wird

Tabelle 1 fasst die Struktur der drei OECD/PISA-Bereiche zusammen. Sie enthält die Definitionen der Bereiche und Beschreibungen der Dimensionen, die die Items charakterisieren.

Tabelle 1. Zusammenfassende Darstellung der PISA-Dimensionen

Bereich	Lesekompetenz	Mathematische Grundbildung	Naturwissenschaftliche Grundbildung
Definition	<i>Geschriebene Texte zu verstehen, zu nutzen und über sie zu reflektieren, um eigene Ziele zu erreichen, das eigene Wissen und Potential weiterzuentwickeln und am gesellschaftlichen Leben teilzunehmen.</i>	<i>Die Rolle zu erkennen und zu verstehen, die die Mathematik in der Welt spielt, fundierte mathematische Urteile abzugeben und sich auf eine Weise mit der Mathematik zu befassen, die den Anforderungen des gegenwärtigen und künftigen Lebens einer Person als konstruktivem, engagiertem und reflektierendem Bürger entspricht.</i>	<i>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.</i>
Komponenten/ Dimensionen des Bereichs	Verschiedene Arten von <i>Texten</i> lesen: kontinuierliche Texte, klassifiziert nach Typen (z.B. Beschreibung, Erzählung), und Dokumente, klassifiziert nach Struktur.	Mathematische <i>Inhalte</i> – primär „mathematische Leitideen“. Im ersten Zyklus werden die Leitideen Veränderung und Wachstum sowie Raum und Form verwendet. In künftigen Zyklen werden auch Zufall, quantitatives Denken, Ungewissheit sowie Abhängigkeiten und Beziehungen einbezogen.	<i>Naturwissenschaftliche Konzepte</i> – z.B. Energieerhalt, Anpassung, Zerfall – ausgewählt aus den Hauptbereichen der Physik, Biologie, Chemie usw., wobei sie auf Angelegenheiten angewendet werden, die mit Energieverbrauch, Artenerhalt oder Gebrauch von Materialien zu tun haben.
	Verschiedene Arten von <i>Leseaufgaben</i> ausführen, etwa bestimmte Informationen herausuchen, eine Interpretation entwickeln oder über den Inhalt oder die Form eines Textes reflektieren.	Mathematische <i>Kompetenzen</i> , z.B. Modellierung, Problemlösen, unterteilt in drei Klassen: i) Verfahren ausführen ii) Verbindungen und Zusammenhänge herstellen iii) Mathematisches Denken und Verallgemeinern	<i>Prozedurale Fähigkeiten</i> – z.B. Belege bzw. Nachweise identifizieren, Schlussfolgerungen ziehen, bewerten und kommunizieren. Diese Fähigkeiten hängen nicht von einem bereits vorhandenen Bestand an naturwissenschaftlichen Kenntnissen ab, können jedoch auch nicht ohne einen naturwissenschaftlichen Inhalt angewendet werden.
	Texte lesen, die für verschiedene <i>Situationen</i> geschrieben wurden, z.B. für persönliche Interessen oder um Arbeitsanforderungen zu genügen.	Anwendung von Mathematik in unterschiedlichen <i>Situationen</i> , z.B. Probleme, die Individuen, Gemeinschaften oder die ganze Welt betreffen.	Anwendung von Mathematik in unterschiedlichen <i>Situationen</i> , z.B. auf Probleme, die Individuen, Gemeinschaften oder die ganze Welt betreffen.

Die Definitionen der drei Bereiche setzen den Schwerpunkt auf funktionale Kenntnisse und Fähigkeiten, die dem Einzelnen eine aktive Teilnahme am gesellschaftlichen Leben ermöglichen. Eine solche Teilnahme erfordert mehr als nur die Fähigkeit, Aufgaben auszuführen, die von anderen, beispielsweise einem Arbeitgeber, gestellt werden. Sie setzt auch voraus, dass man in der Lage ist, an Entscheidungsprozessen teilzunehmen. Die komplexeren Aufgaben von OECD/PISA verlangen von den Schülerinnen und Schülern, über das vorliegende Material zu reflektieren und nicht bloß Fragen zu beantworten, auf die es nur eine einzige „richtige“ Antwort gibt.

Um die genannten Definitionen zu operationalisieren, wird jeder Bereich durch drei Dimensionen beschrieben. Diese entsprechen in etwa:

- dem *Inhalt* oder der *Struktur* des Wissens, das die Schülerinnen und Schüler im jeweiligen Bereich erwerben müssen,
- dem Spektrum von *Prozessen*, die ausgeführt werden müssen und unterschiedliche kognitive Fähigkeiten erfordern, und
- der *Situation* oder dem *Kontext*, in dem Kenntnisse und Fähigkeiten angewendet bzw. genutzt werden.

Ziel ist es, ein breites Spektrum jener Fähigkeiten zu erfassen, die zur Bewältigung vielfältiger Aufgaben nötig sind, mit denen jeder einmal konfrontiert werden könnte. Es sei jedoch ausdrücklich darauf hingewiesen, dass die relative Bedeutung der drei Dimensionen noch genauer bestimmt werden muss. Im Feldtest im Jahre 1999 wird eine große Zahl von Fragen mit unterschiedlichen Merkmalen zur Erfassung von allen drei Dimensionen eingesetzt. Anschließend wird eine Entscheidung darüber gefällt, welches Merkmalspektrum am sinnvollsten ist und auf welche Weise die Leistungen auf den verschiedenen Items zu aggregierten Werten zusammengefasst werden können.

Innerhalb dieses durch die drei Dimensionen gegebenen gemeinsamen Rahmens definiert jeder Bereich seine Dimensionen auf eigene Weise. Dabei gibt es einen zentralen Unterschied zwischen der Lesekompetenz einerseits und der naturwissenschaftlichen und mathematischen Grundbildung andererseits. Die Lesekompetenz betrifft nämlich alle Fächer des Curriculums und hat keinen offensichtlichen eigenen „Inhalt“. Zwar ist ein gewisses Verständnis struktureller Merkmale wie Satzstrukturen wichtig, doch ist dieses Wissen beispielsweise nicht mit der Beherrschung eines Spektrums naturwissenschaftlicher Prinzipien oder Konzepte vergleichbar.

Im Folgenden werden die wichtigsten Aspekte der drei Bereiche näher beschrieben (für eine zusammenfassende Darstellung siehe Tabelle 1).

Lesekompetenz wird definiert als die Fähigkeit, schriftliche Texte zweckentsprechend zu nutzen. Dieser Aspekt der Lesekompetenz ist bereits in früheren Studien wie dem International Adult Literacy Survey (IALS) abgesteckt und bestätigt worden. Im Rahmen von OECD/PISA wird er jedoch durch Einführung eines „aktiven“ Elements erweitert, nämlich durch die Fähigkeit, einen Text nicht nur zu verstehen, sondern unter Einbeziehung eigener Gedanken und Erfahrungen über ihn zu reflektieren. Die Messung der Lesekompetenz umfasst die folgenden Dimensionen:

- Erstens die *Art des Lesestoffes* oder Textes. Viele bisherige Studien haben sich auf Prosatexte mit Sätzen und Absätzen bzw. „kontinuierliche“ oder fortlaufende Texttypen beschränkt. PISA bezieht zusätzlich „nicht-kontinuierliche“ Texte in die Untersuchung ein, in denen die Information nicht fortlaufend dargestellt ist, wie z.B. Listen, Formulare, Graphiken oder Diagramme. Die Studie unterscheidet zudem zwischen einer Reihe von Prosaformen wie Erzählung, Darlegung und Argumentation. Diese Unterscheidungen basieren auf der Annahme, dass Erwachsene im Laufe ihres Lebens verschiedensten Formen von Texten begegnen und es nicht ausreicht, eine begrenzte Zahl von üblicherweise in der Schule anzutreffenden Textsorten lesen zu können.
- Zweitens die *Art der Leseaufgabe*. Hier geht es einerseits um die verschiedenen kognitiven Fähigkeiten, die gute Leser benötigen, und andererseits um die Merkmale von Fragen, die im Test gestellt werden. Da davon ausgegangen werden kann, dass die meisten 15-Jährigen lesen können, zielen die PISA-Tests nicht darauf ab, grundlegende Lesefertigkeiten zu erfassen. Es wird vielmehr untersucht, inwieweit die Schülerinnen und Schüler in der Lage sind, Informationen aus Texten herauszusuchen, die Aussage von Texten zu verstehen und zu interpretieren sowie den Inhalt und die Form von Texten kritisch zu bewerten.
- Drittens die *Art des Gebrauchs, für den der Text geschrieben wurde*, d.h. der Kontext oder die Situation, in der ein Text typischerweise gelesen wird. So werden beispielsweise Romane, persönliche Briefe oder Biographien normalerweise für den „Privatgebrauch“ geschrieben, amtliche Dokumente oder Ankündigungen für den „öffentlichen Gebrauch“, Manuale oder Berichte für den Gebrauch im Berufsleben und Lehrbücher oder

Arbeitsblätter für den bildungsbezogenen Gebrauch. Ein wichtiger Grund für diese Unterscheidung ist die Möglichkeit, dass manche Schülergruppen unter Umständen bei einem Typ von Lesesituationen besser abschneiden als bei anderen. Deshalb ist es wünschenswert, in den Lesetest verschiedene Textsorten aufzunehmen.

Mathematische Grundbildung wird definiert als die Fähigkeit, die Rolle der Mathematik zu verstehen und sich auf eine Weise mit der Mathematik zu befassen, die den eigenen Bedürfnissen entspricht. Damit liegt der Schwerpunkt mehr auf der Fähigkeit, mathematische Probleme zu stellen und zu lösen, als auf der Ausführung spezifischer mathematischer Operationen. Die Messung der mathematischen Grundbildung umfasst die folgenden Dimensionen:

- Erstens mathematische *Inhalte*. Diese werden in erster Linie durch „Leitideen“, die dem mathematischen Denken zu Grunde liegen, bestimmt (z.B. Zufall, Veränderung und Wachstum, Raum und Form, quantitatives Denken, Ungewissheit, Abhängigkeit und Beziehungen) und nur in zweiter Linie durch die curricular definierten Bereiche Algebra, Arithmetik und Geometrie. Für OECD/PISA wurde ein Spektrum von zentralen Konzepten ausgewählt, das eher repräsentativ als umfassend sein sollte; dieses Spektrum wurde für den ersten Messzyklus – in dem Mathematik Nebenkomponente ist – noch einmal reduziert, und zwar auf die zwei Leitideen Veränderung und Wachstum sowie Raum und Form. Diese beiden Bereiche erlauben einerseits eine weitgehende Berücksichtigung von Komponenten mathematischer Lehrpläne und beugen andererseits einer unangemessenen Verengung der Tests auf die Erfassung von Rechenfertigkeiten vor.
- Zweitens *Prozesse* der Mathematik. Diese werden durch solche allgemeinen mathematischen Fähigkeiten wie den verständigen Gebrauch mathematischer Sprechweisen, das Modellieren von Situationen und den verständigen Einsatz von Kenntnissen zum Lösen mathematischer Probleme bestimmt. Da man davon ausgeht, dass für die Lösung einer mathematischen Aufgabe immer ein Bündel von Fähigkeiten benötigt wird, ist es nicht das Ziel, diese in verschiedenen Aufgaben getrennt zu erfassen. Statt dessen werden die Fragen anhand der Anforderungen, die sie an das mathematische Denken stellen, drei sogenannten Kompetenzklassen zugeordnet: Zur ersten Kompetenzklasse gehören einfache Berechnungen und Wiedergabe von Definitionen, wie sie in herkömmlichen Mathematiktests häufig vorkommen. In der zweiten Kompetenzklasse wird bereits verlangt, dass Querverbindungen hergestellt werden, um ansonsten einfache Aufgaben zu lösen. In der dritten Kompetenzklasse wird mathematisches Denken, Verallgemeinern und Verstehen von Zusammenhängen erforderlich. Hier müssen Schülerinnen und Schüler die mathematischen Aspekte einer Situation erkennen und die entsprechende Problemstellung eigenständig formulieren.
- Drittens *Situationen*, in denen Mathematik angewendet wird. Die Rahmenkonzeption berücksichtigt fünf Situationen: persönliche, bildungsbezogene, berufliche, öffentliche und wissenschaftliche. Im Bereich Mathematik wird diese Dimension als weniger wichtig angesehen als die Dimensionen Prozess und Inhalt.

Naturwissenschaftliche Grundbildung wird als die Anwendung von naturwissenschaftlichen Kenntnissen und Prozessen definiert, die eine Person befähigt, die natürliche Umwelt nicht nur zu verstehen, sondern auch an Entscheidungen mitzuwirken, die sie betreffen. Die Messung der naturwissenschaftlichen Grundbildung umfasst die folgenden Dimensionen:

- Erstens *naturwissenschaftliche Konzepte*, die Verknüpfungen darstellen, mit deren Hilfe verwandte Phänomene verstanden werden können. Im Rahmen von OECD/PISA werden zwar die gängigsten Konzepte aus Physik, Chemie, Biologie sowie Geowissenschaften angesprochen. In den Tests werden diese Konzepte jedoch nicht einfach abgefragt, sondern sie müssen auf den Inhalt der Aufgaben angewendet werden. Dabei kommen hauptsächlich Inhalte aus drei Anwendungsbereichen zur Sprache: Leben und Gesundheit, Erde und Umwelt sowie Technologie.
- Zweitens *naturwissenschaftliche Prozesse*, wobei es in erster Linie um die Fähigkeit geht, Nachweise zu erbringen und zu interpretieren sowie ihren Implikationen entsprechend zu handeln. Die PISA-Tests sprechen die folgenden fünf Prozesse an: i) naturwissenschaftliche Fragestellungen erkennen; ii) naturwissenschaftliche Belege/Nachweise identifizieren; iii) Schlussfolgerungen ziehen; iv) Schlussfolgerungen kommunizieren; v) Verständnis wissenschaftlicher Konzepte zeigen. Von diesen Prozessen setzt nur der zuletzt genannte spezifisches naturwissenschaftliches Wissen voraus. Da jedoch kein naturwissenschaftlicher Prozess „inhaltsfrei“ sein kann, erfordern die Fragen in den OECD/PISA-Tests immer auch das Verständnis grundlegender naturwissenschaftlicher Konzepte in den oben genannten Bereichen.

- Drittens *naturwissenschaftliche Anwendungssituationen*. Im Rahmen von OECD/PISA wurden diese Situationen hauptsächlich dem Alltag entnommen und weniger dem naturwissenschaftlichen Unterricht oder der Arbeit von Wissenschaftlern. Die Situationen beziehen sich auf Angelegenheiten, die die Person selbst, die Familie, die Gemeinschaft, die ganze Welt oder die historische Entwicklung naturwissenschaftlicher Erkenntnis betreffen.

Durchführung der Tests und Berichterstattung über die Ergebnisse

Aus Gründen der Durchführbarkeit wird OECD/PISA 2000 mit Papier-und-Bleistift-Instrumenten arbeiten. Für spätere Zyklen sollen auch andere Testformen aktiv erprobt werden.

Die Tests setzen sich aus unterschiedlichen Typen von Items zusammen. Es gibt „geschlossene“ Fragen, bei denen die Schülerinnen und Schüler nur einfache Antworten auszuwählen oder zu finden brauchen, die dann direkt mit einer einzigen richtigen Antwort verglichen werden können. Andere Fragen sind offener und erfordern ausführlichere Antworten. Mit diesen Items können umfassendere Konstrukte erfasst werden als in anderen, traditionelleren Studien. Die Messung von Fähigkeiten höherer Ordnung, die meistens mithilfe von offenen Aufgaben erfolgt, stellt eine bedeutende Neuerung von OECD/PISA dar. In welchem Umfang solche Aufgaben eingesetzt werden, hängt davon ab, wie sich diese Methodologie im Feldtest bewährt und ob ein konsistentes Kodierungsverfahren entwickelt werden kann. Die Verwendung offener Items, die im ersten Zyklus eher sparsam eingesetzt werden, dürfte in späteren OECD/PISA-Zyklen an Umfang zunehmen.

In den meisten Fällen werden die Tests aus Gruppen von Items bestehen, die sich jeweils auf einen gemeinsamen Text, Stimulus oder Gegenstand beziehen. Damit können die Fragen mehr in die Tiefe gehen, als wenn mit jedem neuen Item ein völlig neuer Kontext angesprochen würde. Auch haben die Schülerinnen und Schüler so mehr Zeit, das Material zu verarbeiten, das dann zur Erfassung ganz unterschiedlicher Leistungsaspekte verwendet werden kann.

Insgesamt werden die Items der OECD/PISA-Erhebung anders aussehen als die früherer internationaler Leistungsstudien wie etwa der Third International Mathematics and Science Study (IEA/TIMSS), in der vor allem mit kurzen, auf den Unterrichtsstoff bezogenen Multiple-Choice-Fragen gearbeitet wurde. Manche der naturwissenschaftlichen Fragen dieser Studie erfordern beispielsweise nur einfaches Wissen (z.B. die Schülerinnen und Schüler müssen angeben, wie viele Beine und Körperteile Insekten haben) oder eine einfache Anwendung von Wissen (z.B. die Schülerinnen und Schüler müssen bestimmen, welcher Löffel sich nach 15 Minuten im heißen Wasserbad heißer anfühlt: ein Metall-, ein Holz- oder ein Plastiklöffel). Im Gegensatz dazu verlangen die OECD/PISA-Items im Allgemeinen eine Kombination vieler verschiedener Kenntnisse und Fähigkeiten und manchmal auch (wie z.B. das Beispiel-Item 4 der Rahmenkonzeption für naturwissenschaftliche Grundbildung – siehe Abb. 18) eine aktive Beurteilung von Entscheidungen, für die es nicht nur eine einzige richtige Antwort gibt.

Die Ergebnisse von OECD/PISA werden in jedem Bereich als Leistungsniveaus auf einer Leistungsskala dargestellt. Durch die Kalibrierung der Testaufgaben auf Leistungsskalen wird eine Sprache geschaffen, mit der sich die Kompetenzen beschreiben lassen, die Schülerinnen und Schüler auf unterschiedlichen Leistungsniveaus aufweisen. Es werden also Aussagen darüber möglich, was Personen auf einem bestimmten Leistungsniveau einer Skala wissen und können, Personen auf niedrigeren Leistungsniveaus jedoch nicht. Durch die Einbeziehung von Items, die höhere Denkfähigkeiten erfordern, und von Items, die relativ einfache Verständnisebenen ansprechen, wird sichergestellt, dass die Skalen ein breites Spektrum von Kompetenzen abdecken.

Eine wichtige Frage, die nun zu klären sein wird, ist, ob Kompetenzniveaus in jedem Bereich auf mehr als einer Skala dargestellt werden sollten. Lassen sich die Kompetenzen einer Person ohne weiteres aggregieren und einem bestimmten Niveau zuordnen? Oder ist es sinnvoller, verschiedene Aspekte zu unterscheiden und die hinsichtlich dieser Aspekte jeweils erreichten Niveaus getrennt zu beschreiben? Dies hängt von zwei Dingen ab, über die mit den Feldtests mehr Klarheit gewonnen werden soll: erstens davon, inwieweit die Leistungen einer Person bei einer Art von Fragen mit ihren Leistungen bei einer anderen Art von Fragen korrelieren, sowie vom Muster der Leistungsunterschiede zwischen den Dimensionen; zweitens davon, ob es angesichts der Zahl der Items, die in die Tests aufgenommen werden können, überhaupt möglich sein wird, mehr als eine Berichtsskala in jedem Bereich zu konstruieren. Auf jeder Skala würde den Schülerinnen und Schülern jeweils ein Punktwert zugewiesen werden; bei mehr als einer Skala müsste also jede Schülerin und jeder Schüler jeweils einen Wert für die unterschiedlichen Aspekte des jeweiligen Bereichs erhalten. Im jeweiligen Hauptbereich, der die meisten Fragen umfasst (also im

Rahmen von OECD/PISA 2000 die Lesekompetenz), wird aller Wahrscheinlichkeit nach mehr als eine Berichtsskala konstruiert werden können, während es in den jeweiligen Nebenbereichen vermutlich nur eine Skala geben wird.

Die Kontextfragebögen und ihre Verwendung

Zur Erhebung von Kontextinformationen erhalten Schülerinnen und Schüler sowie Schulleitungen Hintergrundfragebögen, deren Beantwortung 20 - 30 Minuten in Anspruch nehmen wird. Diese Fragebögen sind zentrale Instrumente, die es ermöglichen werden, die Ergebnisse der Tests im Zusammenhang mit Schüler- und Schulmerkmalen auszuwerten.

Die Fragebögen enthalten Fragen über:

- die Schülerinnen und Schüler und ihren familialen Hintergrund, u.a. das ökonomische, soziale und kulturelle Kapital der Jugendlichen und ihrer Familien;
- bestimmte Aspekte des Lebens der Schülerinnen und Schüler, etwa ihre Einstellung zum Lernen, ihre Gewohnheiten und ihr Leben sowohl in der Schule als auch im familialen Umfeld;
- die Merkmale der Schulen, etwa Humanressourcen und materielle Ressourcen, öffentliche oder private Aufsicht und Finanzierung, Entscheidungsprozesse und Personalpolitik;
- den Kontext des Unterrichts, u.a. Strukturen und Typen der Institutionen, Klassengröße und Grad der Elternbeteiligung.

Im ersten OECD/PISA-Zyklus wird außerdem ein weiteres Instrument eingesetzt, das die Schülerinnen und Schüler zum selbstregulierten Lernen befragt. Dieses Instrument umfasst die folgenden Komponenten:

- Strategien des selbstregulierten Lernens, von denen abhängt, wie tief und systematisch Informationen verarbeitet werden;
- motivationale Präferenzen und Zielorientierungen, die Einfluss auf die für das Lernen aufgewendete Zeit und Anstrengung sowie auf die Auswahl von Lernstrategien haben;
- selbstbezogene Kognitionsmechanismen, die Handlungsziele und Handlungsabläufe regulieren;
- Strategien der Handlungskontrolle, insbesondere Anstrengung und Beharrlichkeit, die den Lernprozess vor konkurrierenden Intentionen schützen und zur Überwindung von Lernschwierigkeiten beitragen;
- Präferenzen für bestimmte Typen von Lernsituationen, Lernstilen und sozialen Verhaltensweisen, die für kooperatives Lernen erforderlich sind.

Insgesamt werden die OECD/PISA-Kontextfragebögen eine detaillierte Grundlage für eine schulpolitisch orientierte Ergebnisanalyse liefern. Zusammen mit den Informationen, die durch andere OECD-Instrumente und -Programme erhoben werden, wird es damit möglich:

- Unterschiede in Schülerleistungen mit Unterschieden zwischen Bildungssystemen und Unterrichtskontexten zu vergleichen;
- Unterschiede in Schülerleistungen mit Unterschieden zwischen Curriculuminhalten und pädagogischen Prozessen zu vergleichen;
- Beziehungen zwischen Schülerleistungen und bestimmten schulbezogenen Faktoren wie Größe und Ressourcen zu untersuchen sowie länderspezifische Unterschiede hinsichtlich dieser Beziehungen zu betrachten;
- Unterschiede zwischen den einzelnen Ländern hinsichtlich der Frage zu untersuchen, inwieweit Schulen die Auswirkungen, die Hintergrundmerkmale von Schülerinnen und Schülern auf ihre Leistungen haben, verringern oder verstärken;
- Unterschiede zwischen Bildungssystemen und nationalen Kontexten zu betrachten, die mit Unterschieden in Schülerleistungen zusammenhängen.

Die mithilfe der Schüler- und Schulfragebögen erhobenen Kontextinformationen stellen nur einen Teil der Informationen dar, die OECD/PISA insgesamt zur Verfügung stehen. Indikatoren, die die allgemeine Struktur der Bildungssysteme abbilden (demographische und ökonomische Kontexte wie z.B. Kosten; Zahl der angemeldeten Schülerinnen und Schüler; Zahl der Schülerinnen und Schüler, die einen Abschluss machen; Merkmale von Schulen und Lehrkräften sowie bestimmte Unterrichtsverläufe) und den Arbeitsmarkt beschreiben, werden von der OECD bereits jetzt routinemäßig entwickelt und angewendet.

OECD/PISA – ein Instrument, das sich weiterentwickelt

Angesichts der langfristigen Perspektive des Projekts und der unterschiedlichen Gewichtung der Bereiche in den einzelnen Zyklen stellen die Rahmenkonzeptionen für OECD/PISA Instrumente dar, die sich weiterentwickeln werden. Die Rahmenkonzeptionen sollen flexibel sein, damit sie

- weiterentwickelt und veränderten Interessen der Teilnehmerländer angepasst werden können, aber gleichzeitig
- einen Orientierungsrahmen für die konstanten Elemente liefern, die von bleibendem Interesse sein dürften und deshalb in allen Erhebungszyklen enthalten sein sollten.

Bevor ein endgültiges Instrument für OECD/PISA 2000 entwickelt wird, werden die Rahmenkonzeptionen anhand der Ergebnisse der Feldtests im Jahre 1999 modifiziert. Darüber hinaus soll die Testentwicklung auch danach noch fortgesetzt werden, sodass sowohl veränderte Ziele von Bildungssystemen als auch Verbesserungen von Erhebungstechniken berücksichtigt werden können. Hierbei muss jedoch der Nutzen derartiger Weiterentwicklungen und Verbesserungen gegenüber der Notwendigkeit abgewogen werden, verlässliche Vergleiche im Zeitverlauf zu ermöglichen, sodass viele Kernelemente von OECD/PISA über die Jahre hinweg erhalten bleiben werden.

Die Zielsetzungen von OECD/PISA sind anspruchsvoll. Zum ersten Mal soll mit einer internationalen Schulleistungsstudie nicht nur festgestellt werden, ob die Schülerinnen und Schüler den in den Curricula festgelegten Stoff beherrschen, sondern auch, ob sie durch die Kenntnisse und Fähigkeiten, die sie in der Kindheit erworben haben, ausreichend auf das Erwachsenenleben vorbereitet sind. Länder, die die Leistungsfähigkeit ihrer Bildungssysteme im globalen Vergleich bestimmen wollen, brauchen diese Art der Erfolgskontrolle. Die Idealvorstellungen werden sicher nicht alle sofort zu verwirklichen sein, zumal bestimmte Zielsetzungen von OECD/PISA auch praktischen Beschränkungen unterliegen, die sich dadurch ergeben, dass das Erhebungsinstrument zuverlässig und in unterschiedlichen Kulturen vergleichbar sein muss. Die Zielsetzungen sind jedoch klar, und die Erhebungen, die sich in den kommenden Jahren weiterentwickeln werden, sollen ihnen zunehmend gerecht werden.

PISA 2000 erfasst die Fähigkeiten der Schülerinnen und Schüler mithilfe von Papier- und Bleistift-Tests, die innerhalb eines vorgegebenen zeitlichen Rahmens bearbeitet werden. Der Begriff „Test“ hat in verschiedenen Ländern unterschiedliche Konnotationen. In einigen Ländern beinhaltet das Wort „Test“, dass die Ergebnisse Folgen für den einzelnen Schüler bzw. die einzelne Schülerin haben. Im Rahmen von PISA sollen jedoch nicht einzelne Personen geprüft, sondern die Merkmale aller Schülerinnen und Schüler eines Landes kollektiv erfasst werden. Deshalb wird zur Beschreibung von PISA der Begriff „assessment“ verwendet, obwohl aus der Sicht der Schülerinnen und Schüler die Bedingungen ähnlich sein werden wie bei individuellen schulischen Leistungstests.

DIE ENTWICKLUNG VON OECD/PISA UND SEINER RAHMENKONZEPTIONEN: EIN KOOPERATIONSPROJEKT

OECD/PISA ist ein Kooperationsprojekt der Mitgliedsstaaten der OECD, das eine neuartige, regelmäßige Erfassung von Schülerleistungen liefern soll. Die OECD/PISA-Erhebungen werden von den Teilnehmerländern gemeinsam entwickelt und vereinbart und von nationalen Organisationen durchgeführt.

Ein Board of Participating Countries (BPC), in dem alle Teilnehmerländer vertreten sind, legt im Rahmen der Zielsetzungen der OECD die politischen Prioritäten für OECD/PISA fest und überwacht ihre Einhaltung im Verlauf der Implementierung des Programms. Die Prioritätensetzung betrifft die Auswahl von Indikatoren, die Entwicklung von Messinstrumenten und die Berichterstattung über die Ergebnisse. Außerdem werden Arbeitsgruppen mit Experten aus den Teilnehmerländern gebildet, die gewährleisten sollen, dass sich die politischen Zielsetzungen von OECD/PISA mit der höchsten international verfügbaren fachwissenschaftlichen und verfahrenstechnischen Kompetenz in den verschiedenen Erhebungsbereichen verbinden. Durch ihre Beteiligung in diesen Expertengruppen stellen die Länder sicher, dass

- die Instrumente international valide sind und den Kontext der Kulturen und Bildungssysteme der OECD-Mitgliedsstaaten berücksichtigen,
- die Testmaterialien gute messtechnische Eigenschaften aufweisen und
- die Instrumente die Ziele der Authentizität und bildungspolitischen Validität umsetzen.

Die Implementierung in den Teilnehmerländern erfolgt nach vereinbarten administrativen Verfahren durch die Nationalen Projektmanager. Die Nationalen Projektmanager spielen eine entscheidende Rolle für die Sicherstellung der Qualität dieser Implementierung, und sie verifizieren und evaluieren die Erhebungsergebnisse, Analysen, Berichte und Veröffentlichungen.

Verantwortlich für das Design und die Implementierung der Erhebungen innerhalb des vom BPC abgesteckten Rahmens ist ein internationales Konsortium unter Leitung des Australian Council for Educational Research (ACER). Die anderen Partner in diesem Konsortium sind das Netherlands National Institute for Educational Measurement (CITO), Educational Testing Service (ETS), USA; National Institute for Educational Research (NIER), Japan; und WESTAT in den USA¹.

Die übergreifende Management-Verantwortung für das Programm liegt beim OECD-Sekretariat, das seine Implementierung laufend überwacht, dem BPC als Sekretariat dient, Konsensus zwischen den beteiligten Ländern herstellt und als Mittler zwischen dem BPC und dem mit der Implementierung beauftragten internationalen Konsortium fungiert. Das OECD-Sekretariat wird außerdem die Indikatoren entwickeln und in Zusammenarbeit mit dem PISA-Konsortium sowie in enger Abstimmung mit den Mitgliedsstaaten sowohl auf der politischen Ebene (BPC) als auch auf der Ebene der Implementierung (Nationale Projektmanager) die internationalen Berichte und Veröffentlichungen analysieren und vorbereiten.

Die nachfolgend beschriebenen OECD/PISA-Rahmenkonzeptionen wurden von Expertengremien unter der Leitung von ACER entwickelt. Der Vorsitz der Expertengruppe für Lesekompetenz lag bei Dr. Irwin Kirsch vom Educational Testing Service, die Expertengruppe für mathematische Grundbildung wurde von Professor Jan de Lange von der Universität Utrecht geleitet, und Vorsitzende der Expertengruppe für naturwissenschaftliche Grundbildung war Professor Wynne Harlen vom Scottish Council for Research in Education. Außerdem gingen Stellungnahmen von Expertengruppen in jedem der Teilnehmerländer in die Entwicklung der Rahmenkonzeptionen ein. Die Rahmenkonzeptionen wurden abschließend vom BPC stellvertretend für die Regierungen der OECD-Mitgliedsstaaten angenommen.

¹ Die hier aufgeführten Institutionen bilden das aktuelle Konsortium (Stand: Januar 2000). Zum Zeitpunkt des Erscheinens der englischen/französischen Originalfassung der Rahmenkonzeption war zusätzlich der Service de Pédagogie Expérimentale (SPE), Universität Liège, Belgien, Teil des Konsortiums. ETS und NIER sind im Juli 1999 neu hinzugekommen.

Die Entwicklung der OECD/PISA-Rahmenkonzeptionen kann als eine Folge von sechs Schritten beschrieben werden:

- Entwicklung einer Arbeitsdefinition für jeden Bereich und Beschreibung der Annahmen, die dieser Definition zugrunde liegen;
- Evaluierung von Möglichkeiten der Organisation des Aufgabensets, um zu gewährleisten, dass die Ergebnisberichte Politikern und Wissenschaftlern Aufschluss über die Leistungen von 15-jährigen Schülerinnen und Schülern der Teilnehmerländer in den Erhebungsbereichen geben;
- Festlegung von Schlüsselmerkmalen, die bei der Entwicklung von international einsetzbaren Testaufgaben zu berücksichtigen sind;
- Operationalisierung der bei der Testkonstruktion zu berücksichtigenden Schlüsselmerkmale, wobei den Definitionen die einschlägige Literatur sowie Erfahrungen aus anderen großen Leistungsstudien zu Grunde gelegt werden;
- Validierung der Variablen und Bestimmung ihres Beitrags zur Erklärung der Schwierigkeit der Aufgaben;
- Entwicklung eines Interpretationsschemas für die Ergebnisse.

Im vorliegenden Dokument werden die ersten vier Schritte beschrieben. Die letzten beiden können erst abgeschlossen werden, wenn die Ergebnisse aus den Feldtests vorliegen. Für die Entwicklung der Rahmenkonzeptionen selbst galten folgende Grundsätze:

- Eine Rahmenkonzeption sollte mit einer allgemeinen Definition oder Absichtserklärung beginnen, die die Gründe für die Erhebung und für das, was mit ihr gemessen werden soll, abstecken.
- Eine Rahmenkonzeption sollte verschiedene Aufgabenmerkmale bestimmen und angeben, wie diese Merkmale bei der Aufgabenkonstruktion zu verwenden sind.
- Eine Rahmenkonzeption sollte die zu jedem Aufgabenmerkmal gehörenden Variablen spezifizieren und anhand derjenigen Variablen, die den größten Einfluss auf die Schwierigkeit der Aufgaben zu haben scheinen, ein Interpretationsschema für die Leistungsskala entwickeln.

Der größte Nutzen der Konstruktion und Validierung einer Rahmenkonzeption für die einzelnen Bereiche liegt in der Erhöhung der Güte der Messungen. Darüber hinaus bringt die Entwicklung einer Rahmenkonzeption aber weitere potentielle Vorteile mit sich:

- Eine Rahmenkonzeption liefert eine gemeinsame Sprache und Grundlage für die Diskussion darüber, was mit der Erhebung bezweckt wird und was man mit ihr zu messen versucht. Eine solche Diskussion trägt dazu bei, dass ein Konsensus über die Rahmenkonzeption und über die Ziele der Messungen erzielt wird.
- Eine Analyse der Typen von Kenntnissen und Fähigkeiten, die mit guten Leistungen zusammenhängen, schafft eine Grundlage für die Bestimmung von Leistungsstandards oder -niveaus („proficiency levels“). Indem das, was gemessen wird, zunehmend besser verstanden wird und anhand bestimmter Skalen klarer interpretiert werden kann, wird es möglich, eine empirische Grundlage für die Vermittlung umfassender und detaillierter Informationen für verschiedene Zielgruppen zu schaffen.
- Durch die Identifikation und das Verständnis bestimmter Variablen, die die Grundlage guter Leistungen sind, wird die Fähigkeit gefördert, das, was gemessen wird, zu evaluieren, und im Laufe der Zeit Veränderungen an der Erhebung vorzunehmen.
- Die Verbindung von Forschung, empirischer Leistungsmessung und Politik fördert nicht nur die ständige Weiterentwicklung und -verwendung der Erhebung, sondern auch das Verständnis dessen, was mit ihr gemessen wird.

LESEKOMPETENZ (Reading Literacy)

Im Rahmen von OECD/PISA wird der Begriff *Lesekompetenz* (Reading Literacy) sehr breit gefasst. Da in unserer Gesellschaft vergleichsweise wenige junge Erwachsene überhaupt nicht lesen können, ist in der Rahmenkonzeption kein Maß dafür vorgesehen, ob 15-jährige Schüler im technischen Sinn lesen können oder nicht. Sie spiegelt jedoch aktuelle Konzeptionen des Lesens wider. Danach sollten Schüler mit Sekundarschulabschluss in der Lage sein, die Bedeutung eines breiten Spektrums von kontinuierlichen und nicht-kontinuierlichen Texten, die viele verschiedene schulische und außerschulische Situationen betreffen, zu konstruieren, zu erweitern und darüber zu reflektieren.

Definition des Gegenstandsbereichs

Parallel zu Veränderungen in Gesellschaft, Wirtschaft und Kultur haben sich auch die Definitionen von Lesen und Lesekompetenz im Verlauf der Zeit verändert. Die Konzepte des Lernens und insbesondere des lebenslangen Lernens haben die Wahrnehmung der Lesekompetenz sowie die daran geknüpften Ansprüche erweitert. Lesekompetenz wird nicht mehr als eine nur in der Kindheit, also während der ersten Schuljahre erworbene Fähigkeit angesehen, sondern als ein ständig wachsendes Repertoire an Kenntnissen, Fähigkeiten und Strategien, die lebenslang in verschiedenen Kontexten und durch Interaktion mit Altersgenossen erweitert werden.

Kognitionspsychologische Definitionen der Lesekompetenz betonen die interaktive Natur des Lesens und die konstruktive Natur des Verstehens (Bruner, 1990; Dole et al., 1991; Binkley und Linnakylä, 1997). Der Leser erzeugt als Reaktion auf einen Text Bedeutung, indem er vorhandenes Wissen und eine Reihe von text- und situationsbezogenen Anhaltspunkten nutzt, die oft soziales und kulturelles Allgemeingut sind. Bei der Konstruktion von Bedeutung greift der Leser auf verschiedene Prozesse, Fähigkeiten und Strategien zurück, um das Verständnis zu fördern, zu überprüfen und aufrechtzuerhalten. In der Auseinandersetzung mit unterschiedlichen kontinuierlichen und nicht-kontinuierlichen Texten muß der Leser je nach Situation und Zielsetzung andere Prozesse und Strategien anwenden.

Zwei neuere internationale Studien zur Lesekompetenz haben ebenfalls die funktionale Natur des Lesens betont (Reading Literacy Study der International Association for the Evaluation of Education Achievement – IEA/RLS; und das gemeinsam von Statistics Canada und der OECD durchgeführte International Adult Literacy Survey – IALS).

IEA/RLS definiert Lesekompetenz als

„die Fähigkeit, schriftliche Sprachformen zu verstehen und zu nutzen, die von der Gesellschaft gefordert und/oder von der jeweiligen Person als relevant eingeschätzt werden.“

Auch IALS hebt die funktionale Natur der Lesekompetenz und insbesondere ihr Potential für die Entwicklung des Einzelnen und der Gesellschaft hervor. In dieser Definition stehen allerdings die Informationen im Mittelpunkt und nicht die Sprachformen. Lesekompetenz wird definiert als

„Nutzung von gedruckten und geschriebenen Informationen, um sich in der Gesellschaft zurecht zu finden, eigene Ziele zu erreichen und das eigene Wissen und Potential weiterzuentwickeln.“

In diesen Definitionen von Lesekompetenz steht die Fähigkeit des Lesers im Mittelpunkt, geschriebene oder gedruckte Texte für Zwecke einzusetzen, die von der Gesellschaft gefordert oder vom Einzelnen zur Entwicklung der eigenen Kenntnisse und Fähigkeiten angestrebt werden. Beide Definitionen gehen über das einfache Entschlüsseln und wörtliche Verständnis von Texten hinaus und besagen, dass Lesekompetenz sowohl das Verstehen als auch den Gebrauch von schriftlichen Informationen für funktionale Zwecke umfasst. Sie betonen jedoch nicht die aktive und initiative Rolle, die Leser hierbei spielen. Deshalb lautet die im Rahmen von OECD/PISA verwendete Definition der Lesekompetenz folgendermaßen:

„Lesekompetenz (reading literacy) heißt, geschriebene Texte zu verstehen, zu nutzen und über sie zu reflektieren, um eigene Ziele zu erreichen, das eigene Wissen und Potential weiterzuentwickeln und am gesellschaftlichen Leben teilzunehmen.“¹

Mit den folgenden Ausführungen wird diese Definition näher erläutert.

Lesekompetenz ...

Der Ausdruck „Lesekompetenz“ (Reading Literacy) wird anstelle von „Lesen“ benutzt, da er der allgemeinen Öffentlichkeit eine präzisere Vorstellung davon vermitteln dürfte, was im Rahmen von OECD/PISA gemessen wird. Unter „Lesen“ wird oft die einfache Entschlüsselung oder das laute Lesen verstanden, während OECD/PISA das Ziel hat, etwas viel Umfassenderes und Tiefergehendes zu messen. Im Mittelpunkt steht dabei die Anwendung des Lesens in einem breiten Spektrum von Situationen und für unterschiedliche Zwecke. Historisch war mit dem Begriff „literacy“ Lesefähigkeit gemeint, also eine Technik, die Leser einsetzen, um Kenntnisse zu erwerben. Der Begriff „literacy“ allein ist jedoch unzulänglich, da er allzu häufig zu Analphabetismus oder dem für das Leben in einer bestimmten Gesellschaft erforderlichen Minimum an Fähigkeiten in Bezug gesetzt wird. Dennoch kommt diese Auffassung mit ihrem Verweis auf den instrumentellen Charakter der Lesefähigkeit den Konnotationen nahe, die der Begriff „Lesekompetenz“ in der OECD/PISA-Studie hat, in der ein breites Spektrum von Schülern getestet wird. Von diesen Schülern werden einige studieren und vielleicht eine akademische Karriere einschlagen, andere werden verschiedene Bildungswege im Sekundar- oder Tertiärbereich wählen, um sich auf ihre Teilnahme am Arbeitsleben vorzubereiten, und wieder andere werden direkt nach Beendigung der Pflichtschulzeit in den Arbeitsmarkt eintreten. Von all diesen Schülern wird jedoch unabhängig von ihren akademischen oder sonstigen beruflichen Zielen erwartet, dass sie aktive Teilnehmer in ihren jeweiligen Gemeinschaften werden.

... heißt, geschriebene Texte ...

Mit den Wörtern „geschriebene Texte“ sind alle gedruckten, handschriftlichen oder elektronischen Texte gemeint, die Sprache verwenden. Dazu gehören auch bildhafte Darstellungen wie Diagramme, Bilder, Karten, Tabellen oder Grafiken; nicht aber Filme, Fernsehen, visuelle Animationen oder Bilder ohne Worte. Solche bildhaften Texte können selbständig oder eingebunden in kontinuierliche Texte auftreten. Zu den geschriebenen Texten gehören auch elektronische, obwohl diese ihrer Struktur und ihrem Format nach mitunter anders aussehen und andere Lesestrategien erfordern können. In zukünftigen Erhebungszyklen werden vermutlich auch elektronische Texte verwendet. Dies war im ersten Zyklus aus Zeitmangel und aus Gründen der Zugänglichkeit nicht möglich. Der Begriff „Texte“ wurde statt der in der IALS-Definition benutzten Bezeichnung „Information(en)“ gewählt, da man diesen Begriff für nicht adäquat hielt, um auch Literatur zu erfassen.

... zu verstehen, zu nutzen und über sie zu reflektieren ...

Die Worte „reflektieren über“ wurden den Wörtern „verstehen“ (aus IEA/RLS) und „nutzen“ (aus IEA/RLS und OECD/IALS) hinzugefügt, um zu betonen, dass Lesen interaktiv ist: Leser bringen ihre eigenen Überlegungen und Erfahrungen ein, wenn sie sich mit einem Text befassen. Reflektieren kann bedeuten, dass sie über den Inhalt des Textes nachdenken und dabei auf vorhandenes Wissen oder Verständnis zurückgreifen oder dass sie sich Gedanken über die Struktur oder Form des Textes machen.

... um eigene Ziele zu erreichen, das eigene Wissen und Potential weiter zu entwickeln und am gesellschaftlichen Leben teilzunehmen

Mit diesem Teil des Satzes soll das gesamte Spektrum an Situationen beschrieben werden, in denen Lesekompetenz eine Rolle spielt – vom privaten bis zum öffentlichen Bereich; von der Schule bis zum Arbeitsleben, lebenslangen Lernen und aktiven Leben als Bürger. Die Formulierung: „um eigene Ziele zu erreichen, eigenes Wissen und Potential weiter zu entwickeln“ ist Ausdruck der Vorstellung, dass Lesekompetenz Menschen in die Lage versetzt, eigene Bestrebungen zu verwirklichen – sowohl fest umrissene Ziele, wie den Erwerb eines

¹ Zwar werden die der Definition der Lesekompetenz zugrunde liegenden Annahmen in allen Teilnehmerländern verstanden, doch gibt es bestimmte Wörter in einzelnen Sprachen nicht. Um zu zeigen, dass die Bedeutung dieses Abschnitts in andere Sprachen übersetzt werden kann, ohne die Bedeutung des Begriffs „Lesekompetenz“ oder die seiner Definition zugrunde liegenden Annahmen zu verändern, wurden Übersetzungen angefertigt, die auf der Website von OECD/PISA unter <http://www.pisa.oecd.org> abgerufen werden können.

Bildungsabschlusses oder das Finden einer Arbeitsstelle, als auch weniger genau definierbare und weniger unmittelbare Zielvorstellungen, die der Bereicherung des persönlichen Lebens und einer lebenslang fortgesetzten Bildung dienen. Der Begriff „teilhaben“ wurde anstelle des IALS-Begriffs „zurechtzukommen“ gewählt, weil er impliziert, dass Lesekompetenz die Menschen in die Lage versetzt, einen Beitrag zur Gesellschaft zu leisten und gleichzeitig die eigenen Bedürfnisse zu erfüllen. Der Begriff „zurechtzukommen“ beinhaltet eine einschränkende, pragmatische Konnotation, während „teilhaben“ ein soziales, kulturelles und politisches Engagement impliziert. Zur Teilhabe kann auch eine kritische Haltung gehören und damit ein Schritt auf dem Weg zur persönlichen Befreiung, Emanzipation und Eigenverantwortung. Der Ausdruck „gesellschaftlich“ umfaßt das ökonomische und politische ebenso wie das soziale und kulturelle Leben (Linnakylä, 1992; Lundberg, 1991, 1997; MacCarthey und Raphael, 1989).

Organisation des Gegenstandsbereichs und Aufgabenmerkmale

Nachdem die Definition des Bereichs Lesekompetenz sowie die Annahmen, die dieser Definition zugrunde liegen, erläutert worden sind, soll nun eine Rahmenkonzeption für die Organisation des Bereichs entwickelt werden. Diese Organisation muß sich vor allem daran orientieren, wie über die mit den Aufgaben zur Erfassung von Lesekompetenz erzielten Ergebnisse berichtet werden soll. Dies ist insofern wichtig, als die konzeptuelle Organisation das Testdesign beeinflussen kann. Untersuchungen haben gezeigt, dass Lesen keine einzelne, eindimensionale Fähigkeit darstellt und es daher nicht angemessen wäre, Lesekompetenz mit einer einzigen Skala oder einem einzigen Wert auf dieser Skala darzustellen. Die Frage, wie viele und welche Skalen zur Erfassung der Lesekompetenz benötigt werden, muss vorab geklärt werden. Nur so ist sicherzustellen, dass eine ausreichende Anzahl von Aufgaben entwickelt wird, um diese Skalen angemessen definieren und interpretieren zu können.

Die Konstruktion dieser Skalen kann unter verschiedenen Gesichtspunkten erfolgen. Dabei wäre es am einfachsten, auf Arbeiten zurückzugreifen, die anderen nationalen und internationalen Schulleistungsstudien zugrunde lagen. Sowohl im Rahmen der United States National Assessment of Educational Progress (NAEP) als auch bei IEA/RLS gab es drei Ergebnisskalen. Diese Skalen betrafen das Textformat. IEA/RLS berichtete die Leistungen von Schülern der 4. und der 9. Klasse separat für erzählende, darlegende und dokumentierende Texte. Und auch NEAP arbeitete mit drei Skalen: Literatur bzw. lesen, um der literarischen Erfahrung willen; Information bzw. lesen, um sich zu informieren; und Dokumente bzw. lesen, um eine Aufgabe auszuführen. Ähnliche Skalen wurden auch im Rahmen des International Adult Literacy Survey (IALS) benutzt. Neben einer quantitativen Skala für Lesekompetenz verwendete IALS zur Darstellung der Leseleistung eine Skala für Prosa und eine Skala für Dokumente. Bei dieser Erhebung bestand die Prosa vor allem aus darlegenden Texten, während bei den in Schulen durchgeführten Erhebungen des NAEP das Verhältnis von erzählenden und darlegenden Texten ausgewogener war.

Eine zweite Form der Klassifikation von Leseaufgaben geht von den Situationen aus, die der Aufgabenkonstruktion zugrunde liegen. Eines der Ziele von OECD/PISA besteht darin, Lesekompetenz nicht nur in schulischen oder akademischen Kontexten zu messen, sondern in vielen verschiedenen Situationen. Ein Grund dafür ist, dass mit der Erhebung auch der Frage nachgegangen werden soll, ob die Schülerinnen und Schüler der Zielgruppe ausreichend auf das Berufsleben und die Teilnahme am gesellschaftlichen Leben vorbereitet werden.

Eine weitere Möglichkeit zur Bereichsorganisation und internationalen Berichterstattung über die Ergebnisse könnte auf einem Schema beruhen, das die Inhalte der Aufgaben klassifiziert. So könnte etwa zwischen fachlich-technischen und geisteswissenschaftlichen Inhalten unterschieden werden.

Eine andere mögliche Organisationsform könnte von bestimmten Aspekten des Lesens ausgehen, also etwa dem Entwickeln eines umfassenden allgemeinen Verständnisses, Heraussuchen von Informationen, Entwickeln einer textimmanenten Interpretation, Reflektieren über den Inhalt eines Textes und Reflektieren über Struktur und Form eines Textes. Das eingehende Verständnis eines Textes setzt voraus, dass die Leser alle genannten Aspekte des Leseprozesses aktivieren (Langer 1995). Diese Aspekte stellen daher im Rahmen von OECD/PISA eines der wichtigsten Elemente für die Konstruktion von Aufgaben zur Lesekompetenz dar.

Die OECD/PISA-Erhebung zur Erfassung von Lesekompetenz wird einen reichhaltigen Datensatz produzieren, mit dessen Hilfe entschieden wird, wie ein möglichst nützlicher Ergebnisbericht für Politiker, für im Bildungsbereich tätige Personen und für Wissenschaftler aussehen sollte. Grundlage dieser Entscheidung wird eine Kombination von drei Gruppen von Kriterien sein, nämlich konzeptuelle, empirische und politische Kriterien. Die endgültige Entscheidung darüber, welche Skalen für den Bericht verwendet werden sollen, wird erst nach der Erhebung und Analyse der Daten des OECD/PISA-Feldtests fallen.

Zusätzlich zu dieser an der Berichterstattung orientierten Organisation des Bereichs Lesekompetenz müssen Aufgabenmerkmale identifiziert und zunächst zumindest so operationalisiert werden, dass mit der Konstruktion und Auswahl von Aufgaben begonnen werden kann. Bei der Aufgabenkonstruktion kann nur eine begrenzte Zahl von Merkmalen systematisch variiert werden, und nur eine geringe Zahl von Variablen, die mit diesen bestimmten Merkmalen zusammenhängen, dürfte entscheidenden Einfluss auf die Leseleistungen von Schülern haben.

Almond und Mislevy (1998) weisen darauf hin, dass derartige Variablen bei einer Erhebung eine von fünf Rollen übernehmen können. Sie können dazu benutzt werden, um

- den in der Erhebung zu erfassenden Bereich einzugrenzen,
- Eigenschaften zu charakterisieren, die in der Aufgabenkonstruktion verwendet werden sollen,
- die Zusammenstellung von Aufgaben in Testheften oder Testbögen zu steuern,
- die von den Schülern bei der Bearbeitung der Aufgaben erbrachten Leistungen bzw. ihre Antworten zu charakterisieren oder
- zur Charakterisierung von Kompetenzmerkmalen beizutragen.

Einige dieser Variablen können sowohl bei der Aufgabenkonstruktion und Explizierung von Kompetenzen als auch bei der Charakterisierung von Leistungen nützlich sein.

Es ist davon auszugehen, dass eine begrenzte Zahl von Aufgabenmerkmalen den Leseprozess beeinflussen. Diese sollen bei der Konstruktion von Aufgaben und Bewertung der Antworten berücksichtigt werden. Die folgenden Merkmale sollen im Rahmen von OECD/PISA manipuliert werden, um die interaktive Natur des Leseprozesses zu simulieren und zu evaluieren:

- **Situation:** Da Erwachsene geschriebene Texte nicht in einem Vakuum lesen, sondern im Rahmen einer bestimmten Situation, muss ein Spektrum von Situationen identifiziert werden, aus denen dann eine Stichprobe von Materialien für die Messung der Lesekompetenz gezogen werden kann. Es darf nicht vergessen werden, dass ein Ziel von OECD/PISA darin besteht, über Texte, die im Unterricht verwendet werden, hinauszugehen und ein breites Spektrum von Materialien einzusetzen, mit denen die Schülerinnen und Schüler außerhalb des Unterrichts zu tun haben.
- **Texte:** Es dürfte unumstritten sein, dass zur Messung von Lesekompetenz viele unterschiedliche Materialien herangezogen werden sollten. Für das Testdesign und für die Interpretation der Ergebnisse ist jedoch von entscheidender Bedeutung, welche Art von Textmaterial mit welchen Eigenschaften bei der Aufgabenkonstruktion einbezogen wird. Daher wird OECD/PISA mit einem breiten Spektrum von kontinuierlichen und nicht-kontinuierlichen Texten arbeiten. Außerdem wird erwogen, den Schülern diese Materialien sowohl einzeln als auch in Kombination vorzulegen, also ihnen unter Umständen zum gleichen Thema zwei kontinuierliche Texte oder einen kontinuierlichen Text und einen nicht-kontinuierlichen Text (z. B. eine Grafik oder eine Tabelle) vorzulegen.
- **Test-Rubriken:** Dieser Begriff bezieht sich auf die Merkmale der Fragen und Anweisungen der Aufgaben, auf die Formate, die den Schülern für ihre Antworten vorgegeben werden, und auf die Richtlinien für die Bewertung dieser Antworten. Die Fragen und Anweisungen beinhalten ein Ziel oder einen Zweck, den sich die Leser zu Eigen machen sollen, während sie den Text lesen und mit ihm interagieren. Zur Beantwortung der Fragen werden den Schülern nicht nur Multiple-Choice-Formate vorgegeben, sondern auch offene Formate, bei denen sie ein breiteres Spektrum von Prozessen und Strategien aktivieren müssen.

Um diese drei Hauptmerkmale der Aufgaben beim Design der Erhebung sowie später bei der Interpretation der Ergebnisse benutzen zu können, müssen sie operationalisiert werden. Es müssen also die verschiedenen Ausprägungen, die jedes dieser Merkmale annehmen kann, näher spezifiziert werden. Auf diese Weise können die Konstrukteure der Items das Textmaterial und die Aufgaben so kategorisieren, dass diese zur Interpretation der Ergebnisse und zur Organisation des Berichts verwendet werden können. Anhand dieser Variablen kann auch bestimmt werden, welche Anteile des Textmaterials auf die einzelnen Kategorien entfallen sollten. Welche Rolle die Variablen bei der Interpretation dann tatsächlich spielen werden, ist eine empirische Frage. Wenn sie jedoch beim Design der Erhebung nicht berücksichtigt werden, können sie überhaupt keine Rolle spielen.

Situationen

Der Begriff der Situation bezieht sich in diesem Zusammenhang mehr auf den Zweck, für den ein Autor einen Text verfasst, als auf einen Ort oder einen bestimmten Kontext. Zwar soll Lesekompetenz sowohl mit Texten gemessen werden, die mit der Schule assoziiert werden, als auch anhand von Materialien aus dem außerschulischen Umfeld, doch reicht der Ort, an dem ein Text typischerweise gelesen wird, für die Spezifizierung der Situation nicht aus. Schulbücher etwa werden sowohl in der Schule als auch zu Hause gelesen, wobei sich Prozess und Absicht des Lesens an beiden Orten kaum unterscheiden dürften. Wie Hubbard (1989) festgestellt hat, finden außerdem bestimmte Formen des Lesens, die normalerweise mit dem außerschulischen Umfeld der Kinder verknüpft sind, etwa das Lesen von Vereinsregeln oder Sportergebnissen, oft inoffiziell auch in der Schule statt.

Obwohl das Lesen in den meisten Fällen eine Aktivität ist, die allein ausgeübt wird, hat es dennoch auch soziale Aspekte. Andere Menschen sind als Verfasser, als Themen oder als Personen, die Aufgaben stellen (etwa Lehrer), am Lesen beteiligt. Zur Situation gehört also auch der Bezug auf die Menschen und (z. B. im Falle des Lesens am Arbeitsplatz) die Aufgaben, die mit dem Lesen in Zusammenhang stehen.

Für die Zwecke der OECD/PISA-Erhebung mit 15-jährigen Schülern ist die Situation als eine Form der Kategorisierung von Aufgaben nach folgenden Kriterien zu verstehen: die beabsichtigte Verwendung der Aufgabe, die in der Aufgabe implizit oder explizit enthaltenen Bezüge zu anderen Personen und die allgemeinen Inhalte der Aufgabe (siehe Tabelle 2). Das Lesen eines Schulbuchs wäre danach ein Beispiel für eine Bildungssituation, da der primäre Nutzen dieser Tätigkeit im Erwerb von Informationen im Rahmen einer bildungsrelevanten Aufgabe besteht (Gebrauch), die Tätigkeit mit den von Lehrern oder anderen Ausbildern gestellten Aufgaben assoziiert wird (andere Personen) und inhaltlich in der Regel auf Unterrichts- und Lernzwecke zugeschnitten ist (Inhalt).

Tabelle 2. **Lesesituationen**

	Lesen zum privaten Gebrauch	Lesen zum öffentlichen Gebrauch	Lesen für den Beruf	Lesen zu Bildungszwecken
Andere	Selbst Verwandte Freunde	Anonym	Objekte Kollegen Manager	Lehrende
Gebrauch	Neugier Kontakt	Information	Um etwas zu tun	Um zu lernen
Inhalte	Briefe Romane Biografie Bücher und Zeitschriften mit Anleitungen zum Selbermachen Karten	Bekanntmachungen Vorschriften Programme Pamphlete Formulare	Anweisungen Handbücher Zeitpläne Mitteilungen Berichte Tabellen/Grafiken	Texte Karten Schemata Tabellen Grafiken

Obwohl der Inhalt der Texte bei dieser Erhebung nicht systematisch manipuliert wird, soll die Stichprobe der Texte ganz unterschiedlichen Situationen entnommen werden, um bei der Erhebung der Lesekompetenz eine möglichst große inhaltliche Vielfalt zu erreichen. Daher wird besonders auf Ursprung und Inhalt der ausgewählten Texte sowie auf die Fragen und Anweisungen geachtet, über die die Lesekompetenz der Schülerinnen und Schüler erfasst werden soll. Angestrebt wird ein ausgewogenes Verhältnis zwischen dem Ziel, Aufgaben zu konstruieren, die der im Rahmen von OECD/PISA zugrunde gelegten breiten Definition der Lesekompetenz entsprechen, und der Bedingung, dass die Materialien die sprachliche und kulturelle Vielfalt der Teilnehmerländer repräsentieren sollten. Mit dieser Vielfalt soll gewährleistet werden, dass keine Gruppe durch den Inhalt der Erhebung bevorzugt oder benachteiligt wird.

Eine sinnvolle Operationalisierung der Situationsvariablen ist der Arbeit des Europarats (1996) zum Thema Sprache zu entnehmen:

- *Lesen für den privaten Gebrauch (persönlich)*: Bei diesem Typ des Lesens verfolgt die Person ihre eigenen praktischen oder intellektuellen Interessen. Dies schließt auch das Lesen zur Aufrechterhaltung oder Entwicklung von persönlichen Beziehungen zu anderen Menschen ein. Zu den Inhalten gehören in der Regel persönliche Briefe, Romane, Biografien sowie Informationsmaterial, das in der Freizeit bzw. im Rahmen von Freizeitaktivitäten aus Neugier gelesen wird.
- *Lesen für den öffentlichen Gebrauch*: Dieser Typ des Lesens dient der Teilhabe an den Aktivitäten der Gesellschaft. Dazu gehört auch der Umgang mit offiziellen Dokumenten und Informationen über öffentliche Ereignisse. Im Allgemeinen sind die hiermit verknüpften Aufgaben mit mehr oder weniger anonymen Kontakten mit anderen Personen verbunden.
- *Lesen für die Arbeit (beruflich)*: Obwohl berufsbezogenes Lesen nur bei wenigen 15-Jährigen vorkommen dürfte, sollen dennoch Aufgaben einbezogen werden, die insofern für arbeitsbezogenes Lesen typisch sind, als sie eng mit unmittelbar auszuführenden Aufgaben zusammenhängen und inhaltlich für die Ziele dieser Erhebung relevant sind. Außerdem ist es wichtig, zu erfassen, inwieweit die 15-Jährigen auf die Arbeitswelt vorbereitet sind, in die viele von ihnen innerhalb der nächsten ein bis zwei Jahre eintreten werden. Typische Aufgaben dieser Art werden oft als „Lesen um etwas zu tun“ bezeichnet (Sticht, 1975; Stiggins, 1982).
- *Bildungsbezogenes Lesen*: Dieser Typ des Lesens hat normalerweise mit dem Erwerb von Informationen als Teil einer umfassenderen Lernaufgabe zu tun. Das Lesematerial wird häufig nicht vom Leser selbst ausgewählt, sondern von einem Lehrer oder Ausbilder vorgegeben. Der Inhalt ist gewöhnlich speziell auf Ausbildungszwecke zugeschnitten. Typische Aufgaben dieser Art werden häufig als „Lesen, um zu lernen“ bezeichnet (Sticht, 1975; Stiggins, 1982).

Texttypen

Lesen setzt voraus, dass der Leser etwas hat, das er lesen kann. In einer Testsituation muss dieses Etwas – ein Text – in sich kohärent sein. Das heißt, der Text muss für sich stehen können und darf kein zusätzliches gedrucktes Material erfordern². Während es jedoch offensichtlich ist, dass es viele verschiedene Arten von Texten gibt und dass jede Erhebung ein entsprechend breites Spektrum abdecken sollte, ist weniger klar, wie man die Texttypen am besten kategorisiert. Es gibt eine Reihe verschiedener Vorschläge für geeignete Kategorien, aber viele davon sind eher für praktische Zwecke als für die Zwecke der Forschung entwickelt worden. Eine Gemeinsamkeit dieser Kategoriensysteme besteht jedoch darin, dass sich ein bestimmter Text in der Regel nicht eindeutig nur einer Kategorie zuordnen läßt. Ein Kapitel in einem Schulbuch etwa kann Definitionen enthalten (oft als Texttyp identifiziert), Anweisungen zur Lösung spezieller Probleme (ein anderer Texttyp), eine kurze historische Erzählung über die Entdeckung der Lösung (wieder ein anderer Texttyp) und Beschreibungen von Dingen, die in der Regel an der Lösung beteiligt sind (ein vierter Texttyp).

Man könnte zwar beispielsweise eine Definition für die Zwecke einer Erhebung herauslösen und als selbständigen Text behandeln. Damit würde jedoch der Definition der Kontext genommen und ein künstlicher Texttyp geschaffen (Definitionen kommen außer in Lexika fast nie isoliert vor); dies wiederum könnte die Entwickler von Items davon abhalten, Aufgaben zu konstruieren, bei denen Leser Informationen aus einer Definition mit Informationen aus Anweisungen verknüpfen müssen.

Manche Texte präsentieren sich als Schilderungen der Welt, wie sie ist (oder war), und geben vor, Sachtexte oder jedenfalls nicht-fiktionale Texte zu sein. Fiktionale Schilderungen stehen mit der Welt, wie sie ist, in einer eher metaphorischen Beziehung und präsentieren sich als Darstellung der Welt, wie sie sein könnte oder zu sein scheint. Diese Unterscheidung verliert zunehmend an Eindeutigkeit, da Autoren bei der Abfassung von fiktionalen Texten Formate und Strukturen verwenden, die typisch für Sachtexte sind. OECD/PISA wird Lesekompetenz sowohl anhand von Sachtexten als auch anhand von fiktionalen Texten messen und darüber hinaus auch Texte verwenden, die sich keiner der beiden Kategorien eindeutig zuordnen lassen. Unterschiede in der Leseleistung in Bezug auf diese beiden Typen sollen jedoch nicht erfasst werden.

Eine wichtigere Form der Klassifizierung von Texten, die auch im Rahmen von OECD/PISA als zentrale Grundlage der Organisation des Lesetests dient, besteht in der Unterscheidung zwischen kontinuierlichen und nicht-kontinuierlichen Texten. Kontinuierliche Texte bestehen normalerweise aus Sätzen, die in Absätzen organisiert sind. Diese wiederum können Teil von noch größeren Strukturen wie Abschnitten, Kapiteln oder Büchern sein. Nicht-kontinuierliche Texte liegen häufig im Matrixformat vor und beruhen auf Kombinationen von Listen.

Die Organisation von kontinuierlichen Texten wird durch Absätze, Einrückungen und eine hierarchische, durch Überschriften kenntlich gemachte Gliederung des Textes sichtbar, die dem Leser hilft, die Struktur des Textes zu erkennen. Solche Kennzeichnungen liefern außerdem Hinweise auf Textgrenzen (etwa durch Markierung des Endes eines Abschnitts). Das Auffinden von Informationen wird oft durch die Verwendung unterschiedlicher Schriftgrößen, Schriftarten (z. B. kursiv oder fett) sowie Rahmen oder Schattierungen erleichtert. Die Fähigkeit zur Identifikation der mit diesen Formaten gegebenen Hinweise stellt eine wesentliche sekundäre Voraussetzung für effizientes Lesen dar.

Informationen zur Textorganisation werden auch in Form von Diskurs-Markierungen gegeben. Markierungen für Abfolgen etwa (erstens, zweitens, drittens usw.) verweisen auf die Beziehungen zwischen den Einheiten, die ihnen folgen, und zeigen an, in welcher Beziehung diese Einheiten zu dem sie umgebenden Text stehen.

Kontinuierliche Texte werden primär nach ihrem rhetorischen Zweck bzw. Texttyp klassifiziert.

Nicht-kontinuierliche Texte oder Dokumente, wie sie mitunter auch genannt werden, können auf zwei Arten kategorisiert werden. Die eine Art besteht in dem formal-strukturellen Ansatz, der in den Arbeiten von Kirsch und Mosenthal³ verwendet wird. Dabei werden die Texte danach klassifiziert, wie die zur Konstruktion der verschiedenen nicht-kontinuierlichen Texttypen verwendeten Listen zusammengestellt wurden. Die andere Art der Klassifizierung arbeitet mit einfachen Beschreibungen dieser Texte. Im Gegensatz zu diesem Ansatz ist das Verfahren von Kirsch und Mosenthal systematisch und erlaubt eine Kategorisierung aller nicht-kontinuierlichen Texte unabhängig von ihrer Verwendung.

Typen kontinuierlicher Texte

Die Unterscheidung von Texttypen stellt eine Standardmethode zur Klassifizierung kontinuierlicher Texte nach ihrem Inhalt und nach den Zwecken, für die sie verfasst wurden, dar. Jeder Texttyp weist ein typisches Format auf; dieses wird nach der Beschreibung des jeweiligen Typs angegeben⁴.

1. *Beschreibung* ist der Texttyp, in dem sich die Informationen auf physische, *räumliche* Eigenschaften von Objekten oder von Personenmerkmalen beziehen. Beschreibende Texte geben in der Regel Antworten auf „*Was*“-Fragen.
 - *Impressionistische Beschreibungen* stellen Informationen über Beziehungen, Eigenschaften und räumliche Richtungen aus der Perspektive subjektiver Eindrücke dar.
 - *Sachbeschreibungen* stellen Informationen aus der Perspektive objektiver räumlicher Beobachtungen dar. Häufig werden in Sachbeschreibungen auch nicht-kontinuierliche Textformate wie Diagramme und Illustrationen verwendet.
2. *Erzählung* ist der Texttyp, in dem sich die Informationen auf *zeitliche* Eigenschaften von Objekten beziehen. Erzählende Texte geben in der Regel Antworten auf *Wann*-Fragen oder auf Fragen nach der *Abfolge*.
 - *Erzählungen* stellen Veränderungen aus der Perspektive einer subjektiven Auswahl und Schwerpunktsetzung dar und halten Handlungen und Ereignisse in ihrem zeitlichen Ablauf aus der Perspektive subjektiver Eindrücke fest.
 - *Berichte* stellen Veränderungen aus der Perspektive objektiver, situationsbezogener Rahmenbedingungen dar und geben Handlungen und Ereignisse wieder, die von anderen überprüft werden können.
 - *Nachrichten* sollen dem Leser die Möglichkeit geben, sich eine eigene, unabhängige Meinung über Fakten und Ereignisse zu bilden, ohne durch die Ansichten des Berichterstatters beeinflusst zu werden.

³ Das Modell von Kirsch und Mosenthal wurde detailliert in einer Reihe von monatlichen Beiträgen unter dem Titel „*Understanding Documents*“ (Das Verstehen von Dokumenten) beschrieben, die im *Journal of Reading* zwischen 1989 und 1991 veröffentlicht wurden.

⁴ Dieser Abschnitt basiert auf der Arbeit von Werlich (1997). Die Kategorie „Hypertext“ allerdings ist nicht Bestandteil von Werlichs System.

3. *Darlegung* ist der Texttyp, in dem die Informationen in Form von zusammengesetzten Konzepten oder mentalen Konstrukten oder in Form von denjenigen Elementen dargestellt werden, in die diese Konzepte oder mentalen Konstrukte analytisch zerlegt werden können. Der Text erklärt, wie diese Bestandteile miteinander in Beziehung stehen und sich zu einem sinnvollen Ganzen verbinden, und gibt oft Antworten auf *Wie*-Fragen.
 - *Darlegende Essays* geben aus einer subjektiven Perspektive einfache Erklärungen von Begriffen, mentalen Konstrukten oder Konzeptionen.
 - *Definitionen* erklären, wie Ausdrücke oder Namen mit mentalen Konzepten zusammenhängen. Durch Aufzeigen solcher Zusammenhänge erklärt die Definition die Bedeutung von „Wörtern“.
 - *Erläuterungen* stellen eine Form der analytischen Darlegung dar, mit der erklärt wird, wie ein mentales Konzept mit Worten oder Begriffen verbunden werden kann. Das Konzept wird als zusammenhängendes Ganzes behandelt, das verstanden werden kann, wenn es in seine Bestandteile zerlegt wird und wenn die Beziehungen zwischen diesen Bestandteilen benannt werden.
 - *Zusammenfassungen* sind eine Art synthetischer Darlegung und dienen dazu, „Texte“ in einer kürzeren Form als im ursprünglichen Text darzustellen und zu erklären.
 - *Protokolle* sind mehr oder weniger offizielle Aufzeichnungen der Ergebnisse von Sitzungen oder Vorträgen.
 - *Textinterpretationen* sind eine Form sowohl der analytischen als auch der synthetischen Darlegung, mit der abstrakte Konzepte erklärt werden, die in einem (fiktionalen oder nicht-fiktionalen) Text oder einer Gruppe von Texten enthalten sind.
4. *Argumentation* ist der Texttyp, in dem Aussagen über die Beziehungen zwischen Konzepten oder anderen Aussagen gemacht werden. Argumentierende Texte beantworten oft *Warum*-Fragen. Eine andere wichtige Unterkategorie von argumentierenden Texten sind Texte, die überzeugen wollen.
 - *Kommentare* beziehen Konzepte von Ereignissen, Objekten und Ideen auf ein persönliches Denk-, Wert- und Überzeugungssystem.
 - *Wissenschaftliche Argumentation* bezieht Konzepte von Ereignissen, Objekten und Ideen auf Denk- und Wissenssysteme, so dass die daraus resultierenden Aussagen verifiziert werden können.
5. In *Anweisungen* wird gesagt, was zu tun ist.
 - *Anleitungen* geben Anweisungen für bestimmte Verhaltensweisen, die der Erledigung einer Aufgabe dienen.
 - *Regeln, Vorschriften* und *Bestimmungen* legen die Anforderungen für bestimmte Verhaltensweisen fest, die von einer unpersönlichen Autorität, etwa einer Behörde, verfügt werden.
6. Mit dem Begriff *Hypertext* sind Texteinheiten gemeint, die so zusammengestellt sind, dass sie in unterschiedlichen Abfolgen gelesen werden können. Diese Texte sind häufig visuell aufbereitet und können bei den Lesern nicht-lineare Strategien auslösen.

Nicht-kontinuierliche Texte (Struktur und Format)

Nicht-kontinuierliche Texte sind anders organisiert als kontinuierliche Texte und erfordern deshalb andere Arten von Leseverhalten. Zweckmäßigerweise sollten solche Texte unter zwei Gesichtspunkten betrachtet werden. Beim ersten Gesichtspunkt geht es um die Prinzipien, nach denen die Textelemente angeordnet sind. Diese Variable beschreibt also die *Textstruktur* und erfasst diejenigen Eigenschaften von nicht-kontinuierlichen Texten, die funktional den Sätzen und Absätzen von kontinuierlichen Texten entsprechen. Der zweite Gesichtspunkt erfasst häufig anzutreffende *Formate* nicht-kontinuierlicher Texte.

• Strukturen nicht-kontinuierlicher Texte

Alle nicht-kontinuierlichen Texte bestehen aus einer Reihe von einfachen Listen. Während einige davon nichts weiter als einfache Listen darstellen, bestehen die meisten jedoch aus Kombinationen mehrerer Listen. Die Analyse nicht-kontinuierlicher Texte setzt weder bei ihrer Verwendung an, noch bei den üblichen Bezeichnungen, mit denen sie oft versehen werden, sondern sie identifiziert zentrale Struktureigenschaften, die viele unterschiedliche Texte gemein haben. Für die vollständige Beschreibung eines nicht-kontinuierlichen Textes ist eine Kategorie „Struktur“

und eine Kategorie „Format“ erforderlich. Leser, die die Struktur von Texten verstehen, sind besser in der Lage, die Beziehungen zwischen Textelementen zu identifizieren und zu bestimmen, welche Texte ähnlich und welche unterschiedlich sind.

1. *Einfache Listen* bestehen aus einer Zusammenstellung von Elementen einer einzigen Art. Ein Beispiel für eine einfache Liste ist die Liste der Bücher, die in einem Literaturkurs gelesen werden sollen, oder eine Liste von Schülern, die für besondere Leistungen ausgezeichnet wurden. Die Elemente der Liste können geordnet sein, etwa wenn die Liste der Schülerinnen und Schüler einer Klasse alphabetisch nach dem Nachnamen geordnet wird, oder ungeordnet, etwa in einer Liste von Materialien, die für den Kunstunterricht gekauft werden müssen. In der ersten Liste sind die einzelnen Elemente leichter zu finden als in der zweiten. Wenn die ungeordnete Liste lang ist, kann es schwierig sein, festzustellen, ob ein bestimmtes Element enthalten ist oder nicht. Bei der geordneten Liste sollte dies einfacher sein, vorausgesetzt, man kennt das Ordnungsprinzip.
2. *Kombinierte Listen* bestehen aus mindestens zwei einfachen Listen, wobei jedem Element der einen Liste ein Element der anderen Liste zugeordnet ist. Dabei kann eine der Listen als Primärliste verwendet werden; diese Primärliste ist geordnet, damit ihre Elemente und damit auch die dazugehörigen Parallelinformationen in den anderen Listen besser zu finden sind. Eine elementare kombinierte Liste könnte beispielsweise aus einer Liste von Schülern und einer entsprechenden Liste der von diesen Schülern bei einem Test erzielten Noten bestehen. In solchen Listen können einzelne Elemente auch mehr als einmal vorkommen, obwohl dies bei Primärlisten selten der Fall ist. So können z. B. in der Liste von Schülern und Schülernoten die einzelnen Noten mehrmals erscheinen. Eine kombinierte Liste kann auch aus vielen Einzellisten bestehen, wie zum Beispiel eine Liste von Schlagern, in der die Titel, die Sänger (oder Sängerinnen), das Plattenlabel und die Zahl der Wochen aufgeführt werden, die sich ein bestimmter Schlager in der Hitparade gehalten hat. Das Suchen in nicht-indizierten Listen ist schwieriger, und es kann im Fall solcher Listen auch schwierig sein, festzustellen, ob man wirklich alle relevanten Informationen gefunden hat. Anhand der Liste von Schülern und Schülernoten wäre es einfach herauszufinden, welche Note ein bestimmter Schüler oder eine bestimmte Schülerin erhalten hat, vor allem wenn die Namen der Schülerinnen und Schüler in alphabetischer Reihenfolge stehen. Schwieriger hingegen wäre es, alle Schülerinnen und Schüler zu finden, die den Test nicht bestanden haben.
3. *Matrixlisten* bestehen aus drei Listen, die nicht parallel konstruiert sind, sondern sich überschneiden und eine Matrix aus Reihen und Spalten bilden. Das typische Beispiel für eine Matrixliste ist ein Fernsehprogramm, das aus einer Liste der Uhrzeiten, einer Liste der Sender und einer Liste der Sendungen besteht. Die Sendungen stehen in den Zellen, die sich durch Überschneidungen einer Uhrzeit (normalerweise in den Spalten abgetragen) mit einem Sender (normalerweise in den Reihen abgetragen) bilden. Im universitären Bereich kann ein Fachbereich sein Vorlesungsverzeichnis in Form einer Matrix erstellen, in der zum Beispiel die Spalten den Wochentagen, die Reihen den Uhrzeiten und die Zellen den zu einer bestimmten Uhrzeit an einem bestimmten Wochentag angebotenen Lehrveranstaltungen entsprechen. Dadurch können die Studierenden ohne weiteres diejenigen Kurse ermitteln, die sich zeitlich nicht überschneiden. In einer Matrixliste enthält jede Zelle nur einen Eintrag (Titel der Lehrveranstaltung, Fernsehsendung usw.). Viele statistische Tabellen sind Matrixlisten. Beispielsweise stehen in einer Tabelle der Arbeitslosenquoten in Großstädten oft die Städte in den Reihen, Datumsangaben in den Spalten und die zu einem bestimmten Datum aktuelle Arbeitslosenquote für diese Städte in den Zellen. Die Tabelle kann so aufgebaut sein, dass Vergleiche zwischen Zeitpunkten möglich sind, etwa wenn es mehrere Spalten gibt, die für jeweils andere Zeiträume stehen (Monate, Jahre usw.).
4. *Verschachtelte Listen* bestehen aus einem Satz kombinierter Listen. Beispielsweise werden in einigen Matrixlisten die in den Spalten aufgeführten Kategorien, etwa Wochentage, nicht nur mit den in den Reihen abgetragenen Kategorien (Uhrzeiten), sondern auch noch mit einer vierten Liste kombiniert, etwa den Fachbereichen einer Universität. Bei einer echten verschachtelten Liste muss in allen Matrixlisten derselbe Kategorientyp verwendet werden. Die Matrixliste der Arbeitslosenquoten etwa kann die Arbeitslosenzahlen pro Monat jeweils für Männer und Frauen getrennt ausweisen; in diesem Falle ist das Geschlecht Subkategorie der Spalte „Monat“.
5. *Kombinierte Listen* sind Listen, in denen verschiedene Typen von Listen oder verschiedene Listen des gleichen Typs zu einer Liste zusammengefasst werden. Beispielsweise kann die Matrixliste der monatlichen Arbeitslosenquoten in bestimmten Großstädten mit einer anderen Matrixliste kombiniert werden, die die monatlichen Veränderungen der Arbeitslosenquoten in diesen Städten enthält.

- Formate nicht-kontinuierlicher Texte

Die Klassifizierung nicht-kontinuierlicher Texte nach ihrem Format beinhaltet einen anderen Gesichtspunkt. Jeder nicht-kontinuierliche Text kann sowohl nach der Struktur als auch nach dem Format klassifiziert werden. Formulare etwa bilden eine bestimmte Formatkategorie. Jedes Formular hat aber außerdem eine Struktur. Im allgemeinen handelt es sich um eine kombinierte Liste, in der eine Liste von Rubriken mit einer Liste von Leerstellen verbunden ist, in die den Rubriken zugehörige Informationen eingetragen werden. Ein Fahrplan (für Busse, Bahnen oder Flugverbindungen) weist ein Tabellenformat auf, das oft die Struktur einer Matrixliste oder einer verschachtelten Liste hat. Es ist wichtig, das Format zu erkennen, weil Texte mit gleicher Struktur ganz unterschiedlich dargestellt werden können. Inhaltsverzeichnisse von Büchern und Formulare etwa sind normalerweise kombinierte Listen. Ein Formular besteht, wie oben beschrieben, aus einer Liste mit Rubriken und einer Liste mit Leerstellen. In einem Inhaltsverzeichnis werden eine Liste mit den Kapitelüberschriften und eine Liste mit den Seitenzahlen der jeweiligen Kapitelanfänge genauso kombiniert wie in einem Formular die Rubriken und die Leerstellen. Niemand aber würde ein Formular mit einem Inhaltsverzeichnis verwechseln.

1. *Formulare* sind strukturierte und formatierte Texte, in denen Leser spezifische Fragen auf spezifische Weise beantworten sollen. Viele Organisationen verwenden Formulare, um Daten zu erheben. Oft enthalten sie strukturierte oder vorkodierte Antwortformate. Typische Beispiele sind Steuererklärungen, Einwanderungsanträge, Visaanträge, Bewerbungsbögen, statistische Fragebögen usw.
2. *Informationsblätter* dienen im Gegensatz zu Formularen eher der Weitergabe als der Erhebung von Informationen. Sie fassen Informationen strukturiert und formatiert zusammen, so dass der Leser bestimmte Einzelinformationen einfach und schnell finden kann. Informationsblätter können verschiedene Textformen enthalten, auch in Kombination mit Listen, Tabellen, Abbildungen und graphisch besonders aufbereiteten Textteilen (Überschriften, Schrifttypen, Einrückungen, Rahmen usw.), durch die die Informationen für das Auge zusammengefasst und hervorgehoben werden. Fahrpläne, Preislisten, Kataloge und Programme sind Beispiele für diese Form von Dokumenten.
3. *Gutscheine* dienen dem Nachweis, dass ihr Besitzer Anspruch auf bestimmte Dienstleistungen hat. Aus den Informationen, die sie enthalten, muß außerdem hervorgehen, ob der Gutschein noch gültig ist oder nicht. Typische Beispiele sind Fahrscheine, Rechnungen usw.
4. *Bescheinigungen* sind schriftliche Bestätigungen über die Gültigkeit einer Vereinbarung oder eines Vertrags. Bei ihnen ist eher der Inhalt als das Format formalisiert. Im allgemeinen müssen sie von einer oder mehreren autorisierten und kompetenten Personen unterschrieben sein, die den Wahrheitsgehalt der Angaben bezeugen können. Garantien, Schulzeugnisse, Diplome, Verträge usw. sind Dokumente, die diese Eigenschaften aufweisen.
5. *Aufrufe und Anzeigen* sind Dokumente, mit denen der Leser aufgefordert wird, etwas zu tun, beispielsweise Güter oder Dienstleistungen zu kaufen, Versammlungen oder Treffen zu besuchen, eine bestimmte Person für ein öffentliches Amt zu wählen usw. Der Zweck dieser Dokumente ist, die Leser von etwas zu überzeugen. Sie bieten etwas an und fordern sowohl Aufmerksamkeit als auch Handlungen. Anzeigen, Einladungen, Aufrufe, Warnungen und Bekanntmachungen sind Beispiele für dieses Format.
6. *Schaubilder und grafische Darstellungen* sind bildliche Darstellungen von Daten. Sie werden im Rahmen wissenschaftlicher Argumentation sowie in Zeitungen und Zeitschriften verwendet, um Informationen, die in Zahlen- und Tabellenform vorliegen, optisch gut fassbar zu vermitteln.
7. *Diagramme* kommen oft in technischen Beschreibungen vor (beispielsweise zur Beschreibung der Teile eines Haushaltsgeräts) sowie in darlegenden Texten und Anleitungen (beispielsweise als Anleitung für das Zusammensetzen des Haushaltsgeräts). Es ist häufig sinnvoll, verfahrensorientierte Diagramme (wie etwas gemacht wird) von prozessorientierten Diagrammen (wie etwas funktioniert) zu unterscheiden.
8. *Tabellen und Matrizen*. Tabellen sind Matrizen, die aus Reihen und Spalten bestehen. In der Regel haben alle Einträge in einer Spalte und in einer Reihe jeweils gemeinsame Eigenschaften, so dass die Bezeichnungen von Spalten und Reihen Teil der Informationsstruktur des Textes sind. Häufig vorkommende Tabellen sind etwa Fahrpläne, Tabellenkalkulationen, Auftragsformulare und Inhaltsverzeichnisse bzw. Indizes.
9. *Listen* stellen die einfachste Form nicht-kontinuierlicher Texte dar. Sie bestehen aus einer Reihe von Einträgen mit mindestens einer gemeinsamen Eigenschaft, die als Bezeichnung oder Titel für die Liste verwendet werden kann. Die Listeneinträge können geordnet (etwa die alphabetisch geordneten Namen der Schülerinnen und Schüler einer Klasse) oder ungeordnet (etwa eine Einkaufsliste) sein.

10. *Karten* sind nicht-kontinuierliche Texte, die die geographischen Beziehungen von Orten zeigen. Es gibt eine Vielzahl von Kartentypen. Straßenkarten geben Entfernungen und Streckenverläufe zwischen bestimmten Orten an. Thematische Karten stellen die Beziehungen zwischen Orten und sozialen oder physischen Eigenschaften dar.

Test-Rubriken und Aufgabenmerkmale

Die Test-Rubriken bestehen aus drei Merkmalstypen: Fragen oder Anweisungen, die die Aufgabenstellung für die Schülerinnen und Schüler enthalten; Antwortformate, die vorgeben, wie die Schülerinnen und Schüler ihre Fähigkeiten bei dieser Aufgabe nachweisen sollen; und Auswertungsrichtlinien, die festlegen, wie die Antworten der Schülerinnen und Schüler zu bewerten sind. Alle drei Aspekte sollen hier im Einzelnen behandelt werden, wobei dem ersten erheblich mehr Aufmerksamkeit zu widmen ist als den anderen beiden.

Fragen und Anweisungen

Die den Schülern gestellten Aufgaben können aus einer Makroperspektive und aus einer Mikroperspektive betrachtet werden. Auf der Makroebene lassen sich die Aufgaben anhand von fünf allgemeinen Aspekten des Lesens bestimmen. Auf der Mikroebene wurden mehrere Variablen identifiziert, die den Schwierigkeitsgrad der Aufgaben beeinflussen.

- Makroaspekte

Da davon ausgegangen wird, dass die Mehrheit der 15-Jährigen in den Teilnehmerländern die zur Dekodierung von Texten erforderlichen Grundfähigkeiten beherrscht, muß auf diese nicht näher eingegangen werden. Im Rahmen von PISA werden deshalb vor allem komplexere Lesestrategien erfasst (Dole et al., 1991; Paris, Wasik und Turner, 1991).

Um authentische Lesesituationen zu schaffen, werden die folgenden fünf Aspekte gemessen, die zu einem umfassenden Verständnis von kontinuierlichen und diskontinuierlichen Texten gehören. Die Schülerinnen und Schüler sollen ihre Fähigkeiten in all diesen Aspekten nachweisen:

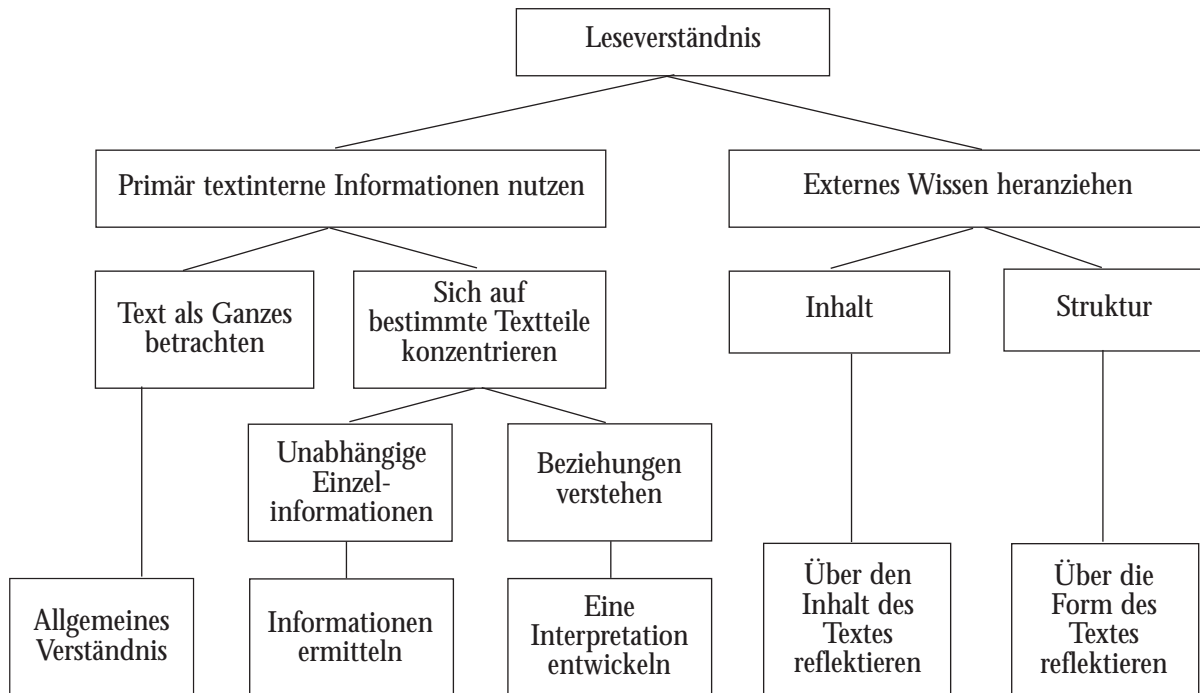
- Allgemeines Textverständnis
- Informationen heraussuchen
- Textinterpretation
- über den Inhalt eines Textes reflektieren und
- über die Form eines Textes reflektieren

Für das vollständige Verständnis eines Textes sind all diese Aspekte wichtig. Es wird erwartet, dass alle Leser unabhängig von ihrer allgemeinen Lesekompetenz in allen genannten Aspekten ein bestimmtes Maß an Kompetenz aufweisen (Langer, 1995). Obwohl natürlich alle fünf Aspekte zusammenwirken – für jeden von ihnen dürften viele der gleichen grundlegenden Fähigkeiten erforderlich sein –, ist die erfolgreiche Bewältigung eines Aspekts nicht unbedingt von der erfolgreichen Bewältigung eines anderen abhängig. Manche Ansätze gehen davon aus, dass die Aspekte Bestandteil des Repertoires jedes Lesers auf jeder Entwicklungsstufe sind und weder eine hierarchische Abfolge noch einen festen Komplex von Fähigkeiten bilden. Diese Annahme wird man jedoch untersuchen müssen, sobald die Items entwickelt wurden und Daten aus den Feldtests vorliegen.

Die fünf Aspekte des Lesens werden über eine Reihe von Fragen und Anweisungen operationalisiert, die den an OECD/PISA beteiligten Schülerinnen und Schülern vorgelegt werden. Mit ihren Antworten auf diese Fragen und Anweisungen sollen die Schülerinnen und Schüler nachweisen, dass sie in der Lage sind, verschiedene kontinuierliche und diskontinuierliche Texte zu verstehen, zu benutzen und über sie zu reflektieren.

Abbildung 1 zeigt die wichtigsten Unterscheidungsmerkmale der fünf Aspekte des Lesens, die in der Regel von OECD/PISA gemessen werden. Auch in dieser notwendigerweise stark vereinfachten Form dürfte die schematische Darstellung nützlich sein, um sich die zwischen den Aspekten bestehenden Beziehungen zu vergegenwärtigen.

Abbildung 1. Unterscheidungsmerkmale der fünf Aspekte des Lesens



Die fünf Aspekte können anhand der folgenden vier Merkmale voneinander unterschieden werden:

- Erstens, inwieweit der Leser Informationen primär dem Text entnehmen oder auf externes Wissen zurückgreifen soll;
- zweitens, inwieweit der Leser den Text als Ganzes betrachten oder sich auf bestimmte im Text enthaltene Informationen konzentrieren soll;
- drittens, ob der Leser spezifische bzw. voneinander unabhängige Informationen finden oder sein Verständnis für die Beziehungen zwischen einzelnen Teilen des Textes nachweisen soll, ob er sich also auf isolierte Textelemente oder auf Verbindungen zwischen Textteilen konzentrieren soll;
- viertens, ob sich der Leser mit dem Inhalt bzw. der Substanz eines Textes oder mit seiner Form bzw. Struktur befassen soll.

Die fünf Aspekte des Lesens werden in Abbildung 1 in der letzten Zeile am Ende der jeweiligen Verzweigungen angegeben. Folgt man den Verzweigungen von oben nach unten, kann man sehen, welche Merkmale zu den einzelnen Aspekten gehören.

Während Abbildung 1 Inhalt und Komplexität der Aspekte von Textverständnis notwendigerweise vereinfacht, soll hier ein erster Versuch unternommen werden, diese Aspekte operational zu definieren und mit bestimmten Typen von Fragen und Anweisungen zu verbinden. Zu beachten ist, dass jeder Aspekt, der hier nur im Hinblick auf das Lesen eines einzigen Textes erläutert wird, auch für mehrere Texte gilt, wenn diese eine Einheit bilden⁵. Die Beschreibung jedes Aspekts hat zwei Teile. Im ersten Teil wird ein allgemeiner Überblick über den Aspekt gegeben, während im zweiten beschrieben wird, auf welche Weise er erfasst werden könnte.

⁵ Die einzelnen Aspekte sollen anhand exemplarischer Items erläutert werden. Ihre Aufnahme in dieses öffentliche Dokument würde jedoch die Testsicherheit gefährden. Beispiele, mit denen die Beschreibung einzelner Aspekte illustriert werden kann, sollen veröffentlicht werden, nachdem die Daten aus den Feldtests erhoben und analysiert wurden und die endgültige Itemauswahl getroffen wurde.

a) Allgemeines Textverständnis

Häufig wollen Leser erst einmal ein allgemeines Verständnis eines Textes entwickeln, ehe sie beschließen, ihn genauer zu lesen. Bei diesem ersten Durchlesen können erfahrene Leser bestimmen, ob der kontinuierliche oder nicht-kontinuierliche Text ihren Zielen entspricht.

Um ein allgemeines Verständnis des Textes zu entwickeln, muss der Leser ihn als Ganzes bzw. unter relativ allgemeinen Gesichtspunkten betrachten. Dieser Vorgang ähnelt der ersten Begegnung mit einer Person oder einem Ort, da der Leser aufgrund seiner ersten Eindrücke Hypothesen oder Vorhersagen über die Inhalte des Textes entwickelt. Diese Eindrücke sind zwar sehr allgemein, aber dennoch sehr wichtig, wenn es darum geht, den relevantesten und interessantesten Lesestoff auszuwählen.

Da die Aufgaben, bei denen es auf ein allgemeines Verständnis ankommt, sich nur auf den Text selbst beziehen, ähneln sie den Aufgaben, bei denen die Leser Informationen heraussuchen oder eine immanente Textinterpretation entwickeln sollen. Im Gegensatz zu diesen Aufgaben müssen Leser bei der Entwicklung eines allgemeinen Verständnisses jedoch das Wesentliche des Textes als Ganzes erfassen – sie müssen erklären können, wovon der Text handelt, sie müssen das behandelte Thema erkennen usw. Dabei sind verschiedene Komponenten wichtig, etwa den Grundgedanken oder das Thema zu bestimmen und – bei einem nicht-kontinuierlichen Text – zu erkennen, wozu der Text im Allgemeinen verwendet wird.

In verschiedenen Aufgaben sollen die Schülerinnen und Schüler ein umfassendes allgemeines Verständnis entwickeln. Dieses erste Verständnis können sie nachweisen, indem sie das Hauptthema oder die zentrale Aussage bzw. den allgemeinen Zweck oder Nutzen des Textes benennen. Beispiele hierfür sind Aufgaben, in denen die Leser aufgefordert werden, einen auf den Text zutreffenden Titel oder eine These auszuwählen bzw. zu formulieren, die Reihenfolge einfacher Anweisungen zu erklären oder die wichtigsten Dimensionen einer Grafik oder Tabelle zu bestimmen. In anderen Aufgaben dieser Art müssen die Schülerinnen und Schüler die Hauptperson, das Umfeld oder das Milieu einer Geschichte beschreiben, ein Thema oder eine Aussage in einem literarischen Text identifizieren, den Zweck oder den Nutzen einer Karte oder einer Abbildung erklären, das Hauptthema oder die Adressaten einer E-Mail-Nachricht bestimmen, herausfinden, welche Art Buch bestimmte Texttypen enthalten könnte, oder den allgemeinen Zweck einer Homepage im Internet erfassen.

In einigen Aufgaben, bei denen es um ein allgemeines Verständnis geht, kann es nötig sein, einen Teil des Textes zu finden, der zu der Frage passt. Dies ist beispielsweise der Fall, wenn ein Thema oder ein zentraler Gedanke im Text explizit formuliert ist. Bei anderen Aufgaben müssen sich die Schülerinnen und Schüler unter Umständen mit mehr als nur einem spezifischen Bezug im Text befassen – beispielsweise wenn sie das Thema des Textes aus dem wiederholten Auftauchen einer bestimmten Kategorie von Informationen ableiten müssen. Die Aufgabe, den Hauptgedanken zu finden, impliziert, dass eine Hierarchie der im Text enthaltenen Gedanken aufgestellt und entschieden werden muss, welche Gedanken die allgemeinen und übergreifenden sind. An einer solchen Aufgabe ist abzulesen, ob ein Schüler oder eine Schülerin zwischen zentralen Ideen und weniger wichtigen Details unterscheiden kann bzw. ob er oder sie in der Lage ist, die Zusammenfassung des Hauptthemas in einem Satz oder Titel zu erkennen.

b) Informationen heraussuchen

Im täglichen Leben benötigen Leser oft eine ganz bestimmte Information. Beispielsweise müssen sie eine Telefonnummer heraussuchen, die Abfahrtszeit eines Busses oder Zugs überprüfen, einen bestimmten Sachverhalt oder Ähnliches ermitteln, der für oder gegen eine Behauptung spricht, die jemand aufgestellt hat. In Situationen wie diesen wollen die Leser eine einzelne Information herausfinden. Dazu müssen sie den Text überfliegen und durchsuchen und die relevante Information lokalisieren und auswählen. Die mit diesem Aspekt des Lesens verbundene Verarbeitung erfolgt häufig auf der Satzebene, obwohl die Information in einigen Fällen auch auf zwei oder mehr Sätze oder Absätze verteilt sein kann.

Erfolgreiches Ermitteln von Informationen erfordert ein unmittelbares Verständnis. Um die benötigte Information zu finden, muss der Leser unter Umständen mehr als eine Information verarbeiten. Wenn der Leser beispielsweise feststellen will, welches der späteste Bus ist, mit dem er sein Ziel noch rechtzeitig erreichen kann, kann er anhand eines Busfahrplans die Ankunfts- und Abfahrtszeiten der verschiedenen Busse vergleichen, die die Strecke befahren. Dazu wird er natürlich mehr als nur eine Information suchen.

In Aufgaben, bei denen es um das Herausuchen von Informationen geht, müssen die Schülerinnen und Schüler die mit der Frage gegebenen Informationen mit wörtlichen oder synonymen Informationen im Text zusammenführen, um auf diese Weise die verlangte neue Information zu ermitteln. Hierbei beschränkt sich die

Informationssuche auf den Text selber und auf die in ihm enthaltenen expliziten Informationen. Die Schülerinnen und Schüler müssen die Information entsprechend der in den Fragen oder Anweisungen genannten Spezifikationen finden. Sie müssen die wesentlichen Elemente einer Aussage herausfinden oder identifizieren (z.B. Personen, Ort / Zeit, Kontext) und dann nach einer Entsprechung im Text suchen, die wörtlich oder synonym sein kann.

Geht es um das Finden von Informationen, müssen die Schülerinnen und Schüler unter Umständen auch mit Mehrdeutigkeiten fertig werden. In einigen Aufgaben sollen sie explizite Informationen suchen, etwa eine Zeit- oder Ortsangabe in einem Text oder einer Tabelle. Eine schwierige Version dieses Aufgabentyps wäre beispielsweise die Suche nach synonymen Informationen. Hierzu muss unter Umständen eine Kategorisierung vorgenommen oder zwischen zwei ähnlichen Informationen unterschieden werden. Durch systematische Variation der Elemente, von denen der Schwierigkeitsgrad dieser Prozesse abhängt, können unterschiedliche Leistungsniveaus innerhalb dieses Aspekts des Textverständnisses gemessen werden.

c) Textinterpretation

Um eine Interpretation zu entwickeln, müssen Leser ihre ersten Eindrücke erweitern und ein genaueres oder umfassenderes Verständnis des Gelesenen erreichen. Dazu müssen sie den Text durchgehen und Informationen aus verschiedenen Textteilen miteinander verbinden sowie sich mit bestimmten Details als Bestandteile des Ganzen befassen.

Für Aufgaben dieser Kategorie ist logisches Verständnis erforderlich: Die Leser müssen die Anordnung der Informationen im Text verarbeiten, und hierzu ist es nötig, die Interaktion zwischen lokalen und globalen Zusammenhängen zu verstehen. In manchen Fällen braucht der Leser, um eine Interpretation zu entwickeln, nur zwei aufeinander folgende Sätze im Rahmen ihres lokalen Zusammenhangs zu verarbeiten, was ihm möglicherweise sogar durch explizite Kennzeichnungen dieses Zusammenhangs erleichtert wird. In schwierigeren Fällen (z.B. Bestimmung von Ursache und Wirkung) sind unter Umständen keine solchen Kennzeichnungen vorhanden.

Ein Text enthält außer den explizit formulierten auch weitere Informationen. Daher ist das Ziehen von Schlussfolgerungen ein wichtiger mentaler Prozess, der im Textverständnis viele verschiedene Funktionen hat. Schlussfolgerungen nutzen Informationen und Gedanken, die während des Lesens aktiviert werden, aber im Text nicht explizit formuliert sind. Sie werden (mehr oder weniger) von der Weltkenntnis des Lesers bestimmt. Manche dieser Schlussfolgerungen werden als notwendig für das Verständnis betrachtet und im Zusammenhang mit der Verarbeitung der sprachlichen Mittel (beispielsweise Verweisketten) gesehen. Sie spielen eine wichtige Rolle für die Kohärenz der beim Lesen entwickelten Interpretation. Mit anderen Schlussfolgerungen werden auf der Grundlage der in dem Text enthaltenen Daten und der Kenntnisse des Lesers neue Informationen erzeugt.

Zur Erfassung dieses Aspekts können z. B. Aufgaben verwendet werden, die Vergleiche und Gegenüberstellungen von Informationen, das Ziehen von Schlussfolgerungen sowie das Identifizieren und Auflisten von Belegen erfordern. In Aufgaben, bei denen es um „Vergleich und Gegenüberstellung“ geht, müssen die Schülerinnen und Schüler zwei oder mehr Informationen aus dem Text in Zusammenhang setzen. Um in solchen Aufgaben explizite oder implizite Informationen aus einer oder mehreren Quellen angemessen zu bearbeiten, muss der Leser oft die beabsichtigte Verbindung oder Kategorie schlussfolgern.

Ein anderes Beispiel für Messungen dieses Aspekts des Textverständnisses sind Aufgaben, in denen die Schülerinnen und Schüler Schlussfolgerungen über die Absichten des Verfassers ziehen und Belege für diese Schlussfolgerungen angeben sollen.

Weitere Beispiele für charakteristische Fragen zu diesem Aspekt sind das Erschließen der Bedeutung aus dem Kontext, die Ermittlung der Motive oder Absichten einer bestimmten Person und die Bestimmung von Ursache und Wirkung.

d) Über den Inhalt eines Textes reflektieren

Um über den Inhalt eines Textes zu reflektieren, müssen die Leser die in einem Text enthaltenen Informationen mit Kenntnissen aus anderen Quellen verknüpfen. Sie müssen außerdem Behauptungen, die im Text aufgestellt werden, mit dem eigenen Weltwissen vergleichen. In vielen Situationen müssen Leser ihren eigenen Standpunkt begründen und vertreten können. Dazu müssen sie in der Lage sein, Aussagen und Intentionen eines Textes zu verstehen und diese mentale Repräsentation mit dem zu vergleichen, was sie aufgrund ihres Vorwissens oder aufgrund von Informationen aus anderen Texten wissen bzw. annehmen. Sie müssen sich auf Belege im Text beziehen und diese mit Informationen aus anderen Quellen vergleichen. Dabei müssen sie allgemeines und spezifisches Wissen anwenden und in der Lage sein, abstrakt zu denken.

Dieser Aspekt des Leseverständnisses erfordert ein hohes Maß an metakognitiven Fähigkeiten. Die Leser müssen ihr eigenes Denken und ihre Reaktionen auf einen Text überwachen, während sie potentielle mentale Modelle testen. Um den Anforderungen dieses Aufgabentyps gerecht zu werden, müssen relevante Informationen abgerufen und in einen kohärenten Zusammenhang gebracht werden können.

In Aufgaben, die Kategorien des Verarbeitens von Texten erfassen, sollten zum Beispiel textexterne Belege oder Argumente angeführt, einzelne Informationen oder Belege auf ihre Relevanz geprüft oder Vergleiche mit moralischen oder ästhetischen Regeln (Normen) angestellt werden. Unter Umständen müssen die Schülerinnen und Schüler auch weitere Informationen einbringen oder identifizieren, die das Argument des Verfassers stützen, oder einschätzen, ob die im Text enthaltenen Belege oder Informationen ausreichend sind.

Das externe Wissen, mit dem Textinformationen verknüpft werden müssen, kann aus dem Wissensbestand der Schülerinnen und Schüler stammen, aus anderen Texten, die ebenfalls im Test enthalten sind, oder aus Überlegungen bzw. Ideen, die in den Fragen explizit formuliert werden.

e) Über die Form eines Textes reflektieren

In Aufgaben dieser Kategorie müssen die Leser sich vom Text lösen, ihn objektiv betrachten und seine Beschaffenheit und Angemessenheit beurteilen. Dazu müssen sie in der Lage sein, Textmerkmale wie Ironie, Humor und logischen Aufbau kritisch zu bewerten und in ihren Auswirkungen zu verstehen. Dieser Aspekt beinhaltet auch die Fähigkeit, Voreingenommenheiten aufzuspüren und Versuche zu erkennen, den Leser auf subtile Weise von etwas zu überzeugen.

Hierbei ist entscheidend, dass die Bedeutung von Textstruktur, Gattung und Stilebene/Sprachebene bekannt ist. Die Kenntnis dieser zum Handwerkszeug eines Autors gehörenden Textmerkmale bestimmt in hohem Maße das Verständnis der Standards, die in Aufgaben dieser Art angesprochen werden. Um zu bewerten, wie gut ein Verfasser eine bestimmte Eigenschaft darzustellen oder jemanden zu überzeugen vermag, müssen Leser nicht nur über inhaltliches Wissen, sondern auch über die Fähigkeit verfügen, sprachliche Nuancen zu erkennen – beispielsweise zu verstehen, wie die Wahl eines bestimmten Adjektivs die Interpretation beeinflusst. Für einen solchen, in die Tiefe gehenden Leseprozess ist es erforderlich, dass der Leser imstande ist, logisch zu denken und kritisch zu analysieren, zu erklären, ob der Verfasser einen bestimmten Sinn angemessen darstellt, Fakten von Meinungen zu unterscheiden usw. Darüber hinaus muss der Leser die wichtigen Texteinheiten auswählen, sekundäre Elemente integrieren und eine Position begründen.

Typische Testaufgaben, in denen über die Form eines Textes reflektiert werden muss, verlangen Beurteilungen der Brauchbarkeit eines Textes für einen angegebenen Zweck sowie der Art und Weise, wie der Autor bestimmte Textmerkmale einsetzt, um diesen Zweck zu erreichen. In anderen Fragen sollen Schülerinnen und Schüler beschreiben oder kommentieren, wie der Autor einen bestimmten Stil einsetzt und welche Absichten und Einstellungen er hat.

• Mikroaspekte

Bei der praktischen Umsetzung der fünf Aspekte, in denen die Schülerinnen und Schüler ihre Kompetenz nachweisen sollen, können drei Prozessvariablen berücksichtigt werden, die den Untersuchungen zum Lesen und zur Lesefähigkeit im Rahmen anderer internationaler Vergleichsstudien (IEA/RLS und IALS) zu entnehmen sind. Diese Variablen sind: die Art der zu ermittelnden Informationen, die Art der Entsprechung zwischen gegebenen und zu ermittelnden Informationen und die Plausibilität von ablenkenden Informationen (Distraktoren). In den folgenden Abschnitten werden die allgemeinen Merkmale dieser drei Variablen erläutert und Überlegungen dazu angestellt, in welchem Format die entsprechenden Fragen beantwortet und wie die Antworten kodiert werden sollen.

a) Art der zu ermittelnden Information

Damit ist die Art der Information gemeint, die die Schülerinnen und Schüler zur erfolgreichen Beantwortung einer Testfrage ermitteln müssen. Je konkreter die zu ermittelnde Information ist, desto einfacher ist die Aufgabe. In früheren Untersuchungen, denen umfangreiche Erhebungen der Lesekompetenz von Erwachsenen und Kindern zugrunde lagen (Kirsch, 1995, Kirsch und Mosenthal, 1994; Kirsch, Jungeblut und Mosenthal, 1998), wurde die Variable Informationstyp auf einer Fünf-Punkte-Skala abgebildet. Dabei entsprach der Wert 1 den konkretesten und daher am einfachsten zu verarbeitenden Informationen und der Wert 5 den abstraktesten Informationen, die auch am schwierigsten zu verarbeiten waren. Fragen, bei denen die Schülerinnen und Schüler eine Person, ein Tier oder

eine Sache (also vorstellbare Substantive) identifizieren sollten, galten beispielsweise als einfach, weil zu ihrer Bearbeitung nur ganz konkrete Informationen ermittelt werden mussten. Daher erhielten sie den Skalenwert 1. Fragen, bei denen die Testpersonen Zielsetzungen, Bedingungen oder Absichten ermitteln mußten, galten als schwieriger, weil diese Art von Informationen abstrakter sind. Diese Fragen erhielten den Skalenwert 3. Fragen, bei denen die Schülerinnen und Schüler ein „Äquivalent“ finden mußten, galten als die abstraktesten; sie erhielten den Skalenwert 5. Das Äquivalent war in diesen Fällen meist ein Ausdruck oder eine Wendung, die wenig geläufig war und deren Definitionen oder Interpretationen durch Schlußfolgerung aus dem Text abgeleitet werden mussten.

b) Art der Entsprechung

Damit ist die Art und Weise gemeint, wie Schülerinnen und Schüler mit dem jeweiligen Text arbeiten müssen, um eine Frage richtig zu beantworten. Dazu gehören auch die Prozesse, mit denen die in der Frage enthaltenen Informationen (die gegebenen Informationen) mit den nötigen Informationen aus dem Text (den neuen Informationen) verknüpft werden, sowie die Prozesse, die erforderlich sind, um die richtige Antwort aus den vorhandenen Informationen herauszusuchen oder zu konstruieren.

Vier Arten von Strategien zur Ermittlung von Entsprechungen wurden identifiziert: einfaches Finden (locating), zyklisches Finden (cycling), Integrieren (integrating) und Generieren (generating). In Aufgaben mit einfachem Finden müssen die Schülerinnen und Schüler zu in der Frage enthaltenen Informationen die gleichlautende oder synonyme Information im Text finden. Beim zyklischen Erschließen müssen ebenfalls eine oder mehrere Entsprechungen zu bestimmten Informationen gefunden werden, aber anders als beim einfachen Finden müssen die Schülerinnen und Schüler hier eine Reihe von Merkmalsabgleichungen durchführen, um die in der Frage formulierten Bedingungen zu erfüllen. Beim Integrieren müssen zwei oder mehr Informationen aus dem Text in eine bestimmte Art von Beziehung zueinander gebracht werden. Zur Herstellung dieser Beziehung kann es zum Beispiel nötig sein, Ähnlichkeiten (durch Vergleiche) oder Unterschiede (durch Gegensatzbildung) zu identifizieren, ein Verhältnis (kleiner oder größer) zu ermitteln oder eine Ursache-Wirkung-Beziehung zu bestimmen. Die entsprechenden Informationen können in einem einzigen Absatz des Textes stehen oder auf mehrere Absätze oder Abschnitte verteilt sein. Beim Integrieren müssen zunächst diejenigen Informationen im Text gefunden werden, die den in der Frage angegebenen Kategorien entsprechen. Anschließend muss dann zwischen diesen den Kategorien zugeordneten Textinformationen eine Beziehung hergestellt werden, die dem in der Frage spezifizierten Typ von Beziehung entspricht. In bestimmten Fällen müssen die Schülerinnen und Schüler die Kategorien und/oder Beziehungen jedoch erst selber bilden, ehe sie die im Text enthaltenen Informationen integrieren können. Dies entspricht der vierten Strategie zur Ermittlung von Entsprechungen, dem Generieren.

Die Art der Entsprechung von Frage und Text bestimmt also die Strategie, die die Schülerinnen und Schüler anwenden müssen. Darüber hinaus ist der Bearbeitungsprozess aber auch noch von mehreren anderen Bedingungen abhängig, die zur Gesamtschwierigkeit der Aufgabe beitragen. Die erste ist die Zahl der Sätze, die bei der Suche bearbeitet werden müssen. Die Schwierigkeit der Aufgabe nimmt mit der Menge der Informationen zu, die in der Frage enthalten sind, von der die Suche ausgeht. Fragen zum Beispiel, die aus nur einem Hauptsatz bestehen, sind in der Regel einfacher als Fragen, die aus mehreren Haupt- oder Nebensätzen bestehen. Die Schwierigkeit nimmt außerdem mit der Anzahl der Antworten zu, die die Schülerinnen und Schüler geben müssen. Fragen, für die eine einzige Antwort genügt, sind leichter als Fragen, die drei oder mehr Antworten erfordern. Darüber hinaus sind Fragen, die die Zahl der erforderlichen Antworten genau angeben, meist einfacher als Fragen, die diese Spezifizierung nicht enthalten. Die Frage: „Welche drei Gründe gibt es ...“, ist zum Beispiel einfacher zu beantworten als die Frage: „Welche Gründe gibt es ...“. Die Schwierigkeit der Aufgaben wird außerdem dadurch beeinflusst, inwieweit die Schülerinnen und Schüler Schlussfolgerungen ziehen müssen, um die in der Frage enthaltenen Informationen den entsprechenden Informationen im Text zuzuordnen und um die zu ermittelnde Information zu bestimmen.

c) Plausibilität der ablenkenden Informationen (Distraktoren)

Hier geht es darum, in welchem Ausmaß die Informationen im Text Gemeinsamkeiten mit der Information aufweisen, die zur Bearbeitung einer Frage ermittelt werden muss, ihr aber nicht voll entsprechen. Aufgaben ohne ablenkende Informationen im Text gelten als die einfachsten. Ihre Schwierigkeit steigt mit zunehmender Anzahl der Distraktoren, mit zunehmender Anzahl der Merkmale, die die ablenkenden Informationen mit der richtigen Antwort gemeinsam haben, und mit zunehmender Nähe zwischen den Distraktoren und der richtigen Antwort im Text. Beispielsweise werden Aufgaben in der Regel als relativ schwierig bewertet, wenn es einen oder mehrere Distraktoren gibt, die einige, aber nicht alle der in der Frage festgelegten Bedingungen erfüllen und in einem

anderen Absatz oder Abschnitt des Textes stehen als in dem, der die richtige Antwort enthält. Als die schwierigsten Aufgaben gelten solche, bei denen es zwei oder mehr Distraktoren gibt, die in den meisten ihrer Merkmale mit der richtigen Antwort übereinstimmen und im selben Absatz oder Informationsknoten auftauchen wie die richtige Antwort.

d) Antwortformate

Zur Messung des Lesekompetenz sind Fragen in der Vergangenheit sowohl mit Multiple-Choice-Antworten wie auch Fragen mit frei formulierten Antworten (constructed-response questions) benutzt worden. Die Testliteratur liefert jedoch nur wenig Anhaltspunkte dafür, welche Formate zur Messung bestimmter Strategien oder Prozesse am besten geeignet sind. So stellte Bennett (1993) fest: „Obwohl kognitive Psychologen diese Annahme stark vertreten, hat die empirische Forschung nicht eindeutig nachweisen können, dass Aufgaben mit frei zu formulierenden Antworten grundsätzlich andere Fähigkeiten messen als Multiple-Choice-Fragen“ (S. 8). So kommt Traub in seinem Überblick der Forschung zu den Unterschieden zwischen beiden Antwortformaten bei Leseverständnistests zu dem Schluss, dass es keine Anzeichen für einen deutlichen Formateffekt gibt (Traub, 1993).

Allerdings gibt es zu Formateffekten auch nur wenig empirische Literatur. Traubs Überblick nennt nur zwei Untersuchungen – eine mit College-Studenten und eine mit Schülern der 3. Klasse. Dabei wurden in der Untersuchung mit den College-Studenten (Ward, Dupree und Carlson, 1987) tatsächlich komplexere Aspekte der Lesekompetenz gemessen. Frederickson stellt in seiner Ansprache als Präsident der American Psychological Association (1984) jedoch fest, dass die eigentliche Verzerrung bei Tests auf die Einschränkungen zurückzuführen ist, die die alleinige Verwendung von Multiple-Choice-Items mit sich bringt. Darüber hinaus sind möglicherweise die Schülerinnen und Schüler in einigen OECD-Ländern mit dem Format von standardisierten Multiple-Choice-Items nicht vertraut. Die Aufnahme verschiedenartiger offener Items dürfte daher eine ausgewogenere Auswahl von Aufgabentypen gewährleisten, die Schülern auf der ganzen Welt aus dem Unterricht bekannt sind. Diese Ausgewogenheit könnte auch dazu führen, dass die Konstrukte breiter gefasst werden.

Es gibt ein breites Spektrum von Aufgaben mit frei formulierten Antworten. Manche erfordern kaum subjektive Beurteilung durch die Kodierer. Dies gilt beispielsweise für Fragen, bei denen die Leser für ihre Antwort nur Teile des Textes oder bestimmte Wörter kennzeichnen müssen. Bei anderen ist ein erhebliches Maß an subjektiver Beurteilung durch die Kodierer notwendig, etwa wenn die Leser einen Text mit eigenen Worten zusammenfassen sollen.

Da zur Frage des Methodeneffekts weder eindeutige Befunde noch überzeugende Ratschläge von Itementwicklern vorliegen, dürfte die beste Lösung für die Erfassung von Lesekompetenz aus einer Mischung von Multiple-Choice-Items und Items mit frei formulierten Antworten bestehen.

e) Kodierung

Bei dichotomen (nur zwei Werte zulassenden) Multiple-Choice-Items ist die Kodierung relativ einfach; entweder hat der Schüler oder die Schülerin die richtige Antwort ausgewählt, oder er oder sie hat sie nicht ausgewählt. Sogenannte „partial credit models“, also Modelle, die abgestufte Punktwerte vorsehen, erlauben eine komplexere Kodierung von Multiple-Choice-Items. Da manche falschen Antworten „richtiger“ sind als andere, werden bei diesem Modell für die Auswahl „fast richtiger“ Antworten Teilwerte der Höchstpunktzahl vergeben. Psychometrische Modelle für solche polytomen Kodierungen haben sich bewährt und sind in mancher Hinsicht den dichotomen Bewertungen sogar vorzuziehen, da sie die in den Antworten enthaltenen Informationen besser nutzen. Die Interpretation polytomer Bewertungen ist jedoch komplexer, da hier jede Frage mehreren Positionen auf der Schwierigkeitsskala zugeordnet werden muss: einer Position für die vollständig richtige Antwort und jeweils einer anderen Position für teilweise richtige Antworten.

Bei dichotomen Items mit frei formulierten Antworten ist die Kodierung ebenfalls relativ einfach, die Festlegung der richtigen Antworten ist jedoch schwieriger. Je mehr die Schülerinnen und Schüler sich eigene Gedanken machen müssen, um eine Frage zu beantworten, desto größer werden die Unterschiede zwischen den richtigen Antworten sein. Um die Vergleichbarkeit der Kodierungen innerhalb der Teilnehmerländer und über die Länder hinweg zu gewährleisten, sind Kodiertraining und Kodierkontrolle in erheblichem Umfang erforderlich. Hier ist ein Gleichgewicht zwischen Festlegung und Offenheit zu finden. Sind die Kodieranweisungen zu spezifisch, werden ungewöhnlich formulierte, aber richtige Antworten unter Umständen als falsch kodiert; sind sie zu offen, werden Reaktionen und Antworten, mit denen die Aufgabe nicht vollständig erfüllt wird, unter Umständen als richtig kodiert.

Items mit frei formulierten Antworten sind für eine Kodierung mit abgestuften Punktwerten besonders gut geeignet, wobei dadurch der Kodierprozess (und damit auch die Entwicklung von Kodieranweisungen) komplexer wird. Abgestufte Punktwerte ermöglichen auch eine differenzierte Bewertung von Aufgaben, bei denen ein bestimmter Antworttyp ein komplexeres Verständnis des Textes anzeigt als ein anderer Antworttyp, obwohl beide Antworten „richtig“ sind. Zumindest für die komplexeren Items mit frei formulierten Antworten empfiehlt sich daher die Verwendung von abgestuften Punktwerten.

Erhebungsstruktur

In diesem Abschnitt wird beschrieben, wie sich die Testaufgaben im Bereich Lesekompetenz auf die verschiedenen Situationen, Textformate, Aspekte und Itemtypen verteilen.

Es wäre nahe liegend, die Aufgaben gleichmäßig auf die vier Situationen zu verteilen (Tabelle 3); den berufsbezogenen Situationen wurde jedoch aus zwei Gründen ein geringeres Gewicht gegeben: Erstens ist es wichtig, die Abhängigkeit der Testleistungen, die mit berufsbezogenen Texten verknüpft sein kann, von berufsbezogenem Spezialwissen zu verringern. Zweitens wird davon ausgegangen, dass sich alle Typen von Fragen und Anweisungen auch in Verbindung mit anderen Situationen konstruieren lassen, zu denen 15-Jährige einen besseren inhaltlichen Zugang haben.

Tabelle 3. **Empfohlene Verteilung der Leseaufgaben auf die Situationen**

Situation	% aller Aufgaben
Privat	28
Bildungsbezogen	28
Berufsbezogen	16
Öffentlich	28
Insgesamt	100

Die Vielfalt der Texte, die die Schülerinnen und Schüler in der Regel von OECD/PISA lesen werden, ist ein wichtiges Merkmal der Erhebung. Die Tabellen 4 und 5 zeigen die empfohlene Verteilung von kontinuierlichen und nicht-kontinuierlichen Texten. Wie man sieht, werden kontinuierliche Texte rund zwei Drittel der Testaufgaben ausmachen. Innerhalb dieser Kategorie werden darlegende Texte den größten Anteil beanspruchen (33 %) und anweisende Texte den kleinsten Anteil (7 %). Die übrigen Typen kontinuierlicher Texte werden mit je 20 % gleichmäßig verteilt sein. Ein Drittel der Texte für den Lesetest wird aus nicht-kontinuierlichen Texten bestehen. Diese werden ganz überwiegend (zu 66 %) aus Tabellen, Diagrammen und Grafiken bestehen. Die übrigen nicht-kontinuierlichen Texte sind Karten, Anzeigen und Formulare, die 15-jährige Schülerinnen und Schüler lesen und benutzen können sollten. Es sei noch einmal daran erinnert, dass es sich bei diesen Prozentsätzen um Zielvorgaben für den Haupttest handelt. Die Textauswahl für den Feldtest und später auch für den Haupttest wird nicht nur nach strukturellen Merkmalen wie Format und Texttyp erfolgen. Berücksichtigt werden auch die kulturelle Vielfalt und das Spektrum der Schwierigkeit der Texte, das potentielle Interesse, das die Schülerinnen und Schüler an den Texten haben könnten, und die Authentizität der Texte.

Tabelle 4. **Empfohlene Verteilung der Leseaufgaben auf Typen von kontinuierlichem Text**

Texttyp	% der kontinuierlichen Texte	% des Tests
Erzählung	20	13
Darlegung	33	22
Beschreibung	20	13
Argumentation	20	13
Anweisung	7	5
Insgesamt	100	66

Tabelle 6 zeigt die empfohlene Verteilung der Leseaufgaben auf die fünf oben definierten Aspekte des Lesens. Den größten Prozentsatz machen Aufgaben aus, in denen eine immanente Textinterpretation entwickelt werden soll. Insgesamt erfassen etwas mehr als zwei Drittel der Aufgaben (70 %) die ersten drei Aspekte. Bei jedem dieser drei Aspekte – allgemeines Verständnis entwickeln, Informationen herausuchen und Interpretation entwickeln – geht es darum, inwieweit der Leser die in erster Linie im Text enthaltenen Informationen verstehen und benutzen kann. Bei den restlichen Aufgaben (30 %) müssen die Schülerinnen und Schüler entweder über die im Text enthaltenen Inhalte bzw. Informationen oder über Struktur und Form des Textes reflektieren. Tabelle 7 zeigt die Verteilung der Aufgaben nach Textformat und Aspekt.

Tabelle 5. **Empfohlene Verteilung der Leseaufgaben auf Typen von nicht-kontinuierlichem Text**

Texttyp	% der nicht-kontinuierlichen Texte	% des Tests
Diagramme/Graphen	33	11
Tabellen	33	11
Schematische Zeichnungen	10	3
Karten	10	3
Formulare	8	3
Anzeigen	6	2
Insgesamt	100	33

Tabelle 6. **Empfohlene Verteilung der Leseaufgaben auf die Aspekte des Lesens**

Aspekt	% des Tests
Informationen herausuchen	20
Allgemeines Textverständnis	20
Immanente Textinterpretation	30
Über den Inhalt reflektieren	15
Über die Form reflektieren	15
Insgesamt	100

Tabelle 7. **Empfohlene Verteilung der Aufgaben nach Format und Aspekt**

Aspekt	% des Tests	% der kontinuierlichen Texte	% der nicht-kontinuierlichen Texte
Informationen herausuchen	20	13	7
Allgemeines Textverständnis	20	13	7
Immanente Textinterpretation	30	20	10
Über den Inhalt reflektieren	15	10	5
Über die Form reflektieren	15	10	5
Insgesamt	100	66	34

Die Festlegung des genauen Anteils der Items mit frei formulierten Antworten wird sowohl auf praktischen als auch auf konzeptuellen Überlegungen basieren. Tabelle 8 zeigt die vorgeschlagene Verteilung der Aufgaben mit frei formulierten Antworten sowie der Multiple-Choice-Aufgaben auf die fünf Aspekte des Lesens.

Tabelle 8. Empfohlene Verteilung der Aufgaben mit frei formulierten Antworten und mit Multiple-Choice-Antworten auf die fünf Aspekte des Lesens

Aspekt	% des Tests	% der Aufgaben mit frei formulierten Antworten	% der Items mit frei formulierten Antworten	% der Items mit Multiple-Choice-Antworten
Informationen herausuchen	20	35	7	13
Allgemeines Textverständnis	20	35	7	13
Immanente Textinterpretation	30	35	11	19
Über den Inhalt reflektieren	15	65	10	5
Über die Form reflektieren	15	65	10	5
Insgesamt	100		45	55

Tabelle 8 zeigt, dass ungefähr 45 % des Lesetests aus Items mit frei formulierten Antworten bestehen werden, die den Kodierern ein subjektives Urteil abverlangen. Die übrigen 55 % sind Multiple-Choice-Items sowie solche Items mit frei formulierten Antworten, bei denen die Kodierer nur in geringem Maße subjektiv urteilen müssen. Aus der Tabelle geht außerdem hervor, dass Multiple-Choice-Items und Items mit frei formulierten Antworten zwar für alle fünf Aspekte eingesetzt werden, aber nicht gleichmäßig verteilt sind: Bei den ersten drei Aspekten der Lesekompetenz ist der Anteil der Multiple-Choice-Items höher.

Skalenentwicklung

Eine wesentliche Voraussetzung dafür, dass OECD/PISA seinen Zweck erfüllen kann, ist die Entwicklung von Skalen zur Beschreibung von Schülerleistungen. Die Entwicklung solcher Skalen ist ein iterativer Prozess, da die ersten Fassungen, die auf bereits vorliegenden Erfahrungen und Forschungen im Bereich des Lernens und der kognitiven Entwicklung in der Mathematik beruhen, auf der Grundlage der beim OECD-PISA-Feldtest gewonnenen empirischen Daten weiterentwickelt werden sollen.

Zwei der in der Rahmenkonzeption enthaltenen Organisationsaspekte werden als Grundlage für die Skalenentwicklung erwogen: die Texttypen (kontinuierlich und nicht-kontinuierlich) und die Makroaspekte (ein allgemeines Verständnis entwickeln, Informationen herausuchen, eine immanente Textinterpretation entwickeln, über den Inhalt eines Textes reflektieren und über die Form eines Textes reflektieren). Dies bedeutet, dass es über eine Gesamtskala für Lesekompetenz hinaus entweder zwei oder fünf weitere Skalen geben wird.

Sonstiges

Außer der Frage, wie Lesekompetenz im Rahmen dieser Erhebung definiert und gemessen werden soll, müssen auch noch andere Aspekte der Erhebung angesprochen werden. Drei betreffen die Kontextfragebögen, die die Schülerinnen und Schüler beantworten sollen, zwei weitere das Verhältnis zu anderen Leistungsstudien. Diese Fragen werden in den folgenden Abschnitten behandelt.

Fragebögen

Manche Bereiche lassen sich besser mit Fragebögen erfassen als mit Leistungstests. Ein solcher Bereich ist die Erhebung von Daten zu den Lesepraktiken und -interessen, ein anderer betrifft bestimmte Aspekte der Metakognition und ein dritter hat mit der Rolle zu tun, die die Technologie im Leben der an OECD/PISA teilnehmenden Schülerinnen und Schüler spielt.

Lesepraktiken und -interessen

Der OECD/PISA-Kontextfragebogen für Schülerinnen und Schüler wird – genau wie die Fragebögen in anderen großen Studien zu Lesen und Lesekompetenz (Council of Ministers of Education, Canada, 1994; Elley, 1992; Jones, 1995; Smith, 1996; Taube und Mejdning, 1997) – Fragen zu ihren Lesepraktiken und zum allgemeinen Kontext ihrer Leseaktivitäten innerhalb wie außerhalb der Schule enthalten. Solche Fragen sind sowohl in deskriptiver als auch in erklärender Hinsicht von Interesse. Mit ihrer Hilfe kann nämlich die Population der 15-jährigen Schülerinnen und Schüler hinsichtlich folgender Merkmale beschrieben werden: Zugang zu gedruckten Materialien, Interesse an bestimmten Leseaktivitäten und Einstellung zu diesen Aktivitäten sowie aktuelle Lesepraktiken. Außerdem können diese Variablen dazu beitragen, einen Teil der Varianz der Leseleistung zu erklären.

Für die Auswahl der in die Befragung einzubeziehenden Typen von Information waren vor allem zwei Prinzipien maßgeblich:

- die erfassten Informationen müssen bildungspolitisch relevant sein;
- die Messung der Lesekompetenz sollte durch Fragebögen ergänzt werden, mit denen Informationen über die Einstellungen der Schülerinnen und Schüler zum Lesen und zu bestimmten Lesepraktiken erhoben werden.

Dieser Teil des Kontextfragebogens erfasst:

- *Inwieweit Schülerinnen und Schüler zu Hause, in der Schule oder in öffentlichen Zusammenhängen mit verschiedenen Arten schriftlichen Materials konfrontiert werden.* Hierzu gehören folgende Fragen: Wie viele Bücher gibt es zu Hause? Besitzt die Schülerin oder der Schüler eigene Bücher? Werden in der Familie regelmäßig Tageszeitungen oder Wochenzeitschriften gekauft? Wird eine öffentliche Bibliothek oder eine Schulbibliothek benutzt? usw.
- *Lesepraktiken/Lesegewohnheiten.* Es ist wichtig, sicherzustellen, dass die Vielfalt und Häufigkeit der verschiedenen Arten von Druckerzeugnissen sowohl in Bezug auf die Texttypen und Textformate, die in der Rahmenkonzeption unterschieden werden, als auch in Bezug auf den Kontext der Erhebung angemessen repräsentiert sind. Aufgrund zeitlicher und methodischer Beschränkungen kann eine vollständige Erfassung aller Leseaktivitäten nicht geleistet werden. Statt dessen muss man ein sinnvolles Gleichgewicht finden, das einerseits der Notwendigkeit Rechnung trägt, die unterschiedlichen Druckerzeugnisse in vielfältigen Situationen aufzulisten (um die Vielfalt zu messen), und andererseits die verschiedenen Beschränkungen berücksichtigt.
- *Einstellungen zum Lesen und Leseinteressen.* Einstellungen zum Lesen und Motivation dürften einen Einfluss auf Lesepraktiken und Leseleistungen haben; diese wiederum können durch die Schaffung eines dem Leseverständnis förderlichen Klimas innerhalb und außerhalb der Schule beeinflusst werden. Im Rahmen von OECD/PISA wird dieser Aspekt mit einigen gezielten, rasch zu beantwortenden Fragen gemessen (etwa: Bekommst du gern Bücher geschenkt? Gehst du gern in die Bücherei? usw.). Darüber hinaus wird die Präferenz für das Lesen mit der Präferenz für andere Freizeitaktivitäten verglichen (Fernsehen, Musik, Ausgehen, Videospiele usw.). Dabei könnte sich herausstellen, dass die Beantwortung von Fragen dieses Typs weniger stark durch eine Orientierung am sozial Erwünschten (compliance effects) beeinflusst wird, als es sonst bei Studien zum Leseverhalten oft der Fall ist.

Metakognition

Verschiedene Untersuchungen haben gezeigt, dass es bei jugendlichen Lesern einen Zusammenhang zwischen metakognitivem Wissen und Leseleistung gibt (Ehrlich et al., 1993; Ehrlich 1996). So wurden interindividuelle Unterschiede beim Verständnis des Begriffs Lesen gefunden, bei den Zielen und Zwecken des Lesens, bei den Strategien, die angewendet werden, um das Gelesene zu verstehen, und beim Aufdecken von Widersprüchen im Text. Diese Unterschiede hängen mit zwei Grundelementen der Metakognition zusammen: dem Wissen der Schülerinnen und Schüler über Kognition und der Regulation von Kognition. Das erste Element betrifft die Fähigkeit, über die eigenen kognitiven Prozesse zu reflektieren, wozu auch das Wissen gehört, wann, wie und warum bestimmte kognitive Aktivitäten auszuführen sind. Das zweite, die Regulation, bezieht sich auf die Anwendung von Strategien zur Kontrolle der eigenen kognitiven Bemühungen (Baker, 1991).

Dem Interesse an einer Messung von Metakognition im Rahmen von OECD/PISA liegt die Überzeugung zugrunde, dass die Ergebnisse bildungspolitische Relevanz haben und die Lese- und Lernpraxis beeinflussen könn-

ten, insbesondere da man davon ausgeht, dass metakognitive Fähigkeiten nicht nur bezogen auf Leseaufgaben, sondern allgemeiner gelehrt und angewendet werden können. Die Herausforderung besteht darin, eine Methode zu finden, mit der man Metakognition in groß angelegten Erhebungen wie OECD/PISA messen kann. Ein Großteil der Literatur zu diesem Thema beruht auf experimentellen Studien mit Schülern, die jünger sind als die Zielgruppe von OECD/PISA (Myers und Paris, 1978). Damit scheint es kein Instrument zu geben, von dem angenommen werden kann, dass es im Rahmen dieser Studie verlässliche und valide Ergebnisse liefern würde.

Da es kein verlässliches Instrument zur Messung der Metakognition von 15-jährigen Schülern gibt, und keine Ressourcen für die Entwicklung eines derartigen Instruments zur Verfügung stehen, wurde entschieden, Metakognition im ersten Zyklus weder im Feldtest noch in der Haupterhebung zu messen. Angesichts des Interesses an diesem Thema wird jedoch die Entwicklung eines Instruments zur Messung von Metakognition in künftigen Zyklen in Betracht gezogen.

Technologie

In unserer sich rasch verändernden Welt muss die Untersuchung von Lesegewohnheiten und Metakognition auch auf Fragen ausgeweitet werden, bei denen Technologie und insbesondere Computer eine Rolle spielen. Die Verfügbarkeit von elektronischen Texten und ihre Verwendung bei der Gewinnung und dem Austausch von Informationen werden im Leben von Schülern in den kommenden Jahren eine immer größere Bedeutung bekommen.

Um die zunehmende Berücksichtigung von Technologie in künftigen OECD/PISA-Zyklen vorzubereiten, wird in diesem Zyklus ein kurzer Fragebogen eingesetzt, der folgende Informationen erfasst: Zugang, den die Schülerinnen und Schüler zu Hause, in der Schule, bei der Arbeit oder in öffentlichen Einrichtungen an ihrem Wohnort zu Computern haben; Einstellungen zur Nutzung des Computers; Häufigkeit der Computernutzung in verschiedenen Situationen; Typen der mit Computern ausgeführten Aktivitäten.

Verknüpfungen mit anderen Studien

Für eine detaillierte Interpretation der bei OECD/PISA gewonnenen Leistungsdaten ist es wichtig, möglichst viele der bereits vorliegenden Informationen einzubeziehen. Solche Zusatzinformationen werden zum Teil aus den Fragebögen kommen. Darüber hinaus können aber auch durch Verknüpfungen mit anderen Studien Daten gewonnen werden, die die Interpretation der Ergebnisse bereichern dürften.

OECD/PISA ist nicht die einzige groß angelegte Studie, in der Lesekompetenz erfasst wird. Auch im Rahmen des bereits genannten International Adult Literacy Survey (IALS) wurde die Lesefähigkeit von Erwachsenen gemessen, und es wäre durchaus nützlich, die Ergebnisse von OECD/PISA auch im Kontext dieser Studie zu interpretieren, obwohl sich die Instrumente von OECD/PISA und IALS hinsichtlich ihrer Art und ihres Designs erheblich unterscheiden. Eine Verknüpfung mit IALS würde auch eine Verknüpfung der Messungen für Schülerinnen und Schüler mit den Messungen für Erwachsene ermöglichen. Dies läßt sich zum Teil dadurch erreichen, dass Items aus IALS in den OECD/PISA-Test aufgenommen werden, womit eine direkte Verknüpfung hergestellt wird. Dabei war zu überlegen, wie viele Items aus IALS übernommen werden müssten, um sowohl eine konzeptionelle als auch eine statistische Verknüpfung gewährleisten zu können. Außerdem galt es, den Wunsch, eine Verknüpfung zwischen IALS und OECD/PISA herzustellen mit dem Ziel zu vereinbaren, etwas zu messen, das über frühe Studien hinausgeht.

Eine Verknüpfung mit mehr als einer der drei IALS-Skalen etwa wäre nicht möglich, da die Testzeit für die Einbeziehung eines breiteren Spektrums von IALS-Items nicht ausreicht. Da es im Rahmen von OECD/PISA wahrscheinlich weniger Aufgaben mit nicht-kontinuierlichen als mit kontinuierlichen Texten geben wird, wäre eine Verwendung der IALS-Dokumentenskala wenig sinnvoll. Um nämlich eine solche Verknüpfung zu erlauben, müssten vermutlich so viele IALS-Items in den OECD/PISA-Test aufgenommen werden, dass sie diese Kategorie dominieren würden. Gleiches gilt für die quantitative Skala von IALS, da auch dieser überwiegend nicht-kontinuierliche Aufgaben zugrunde liegen. Die Prosa-Aufgaben aus IALS hingegen scheinen für eine Verwendung im Rahmen von OECD/PISA geeignet, da sie unter den kontinuierlichen Texten kein Übergewicht bilden würden.

Um diese Verknüpfung untersuchen zu können, wurden zwei Itemblöcke der Prosa-Skala von IALS in den Feldtest von OECD/PISA übernommen. Damit soll geprüft werden, wie gut die Item-Parameter aus IALS auf die OECD/PISA-Population passen. Da anzunehmen ist, dass die meisten Parameter passen werden, soll eine angemessene Anzahl von IALS-Prosaaufgaben in die Haupterhebung von OECD/PISA aufgenommen werden.

Diese Items dienen einem doppelten Zweck. Erstens werden die Item-Parameter aus IALS dazu verwendet, die Leistungen im Bereich der Prosaliteratur in den OECD/PISA-Ländern zu messen. Zweitens sollen diese Items als Teil der OECD/PISA-Rahmenkonzeption für Lesekompetenz kodiert werden und dazu beitragen, die Leseleistung von 15-jährigen Schülern in den Teilnehmerländern zu messen.

Da im Rahmen von OECD/PISA auch andere Leistungen gemessen werden (Mathematik und Naturwissenschaften), bietet die Studie die Möglichkeit, Beziehungen zwischen diesen Bereichen zu schätzen. Simple Ansätze würden einfach davon ausgehen, dass ähnliche Populationen in den gleichen Ländern zur gleichen Zeit erfasst werden, ohne den Versuch zu machen, diese Beziehungen zu modellieren. Die interessanteren, aber auch komplexeren Ansätze würden mit Items arbeiten, die alle Bereiche erfassen. Mit Hilfe der Ergebnisse solcher gemeinsamen Items können dann die Beziehungen zwischen Lesen und Naturwissenschaften oder Mathematik direkt modelliert werden. Obwohl psychometrische Modelle für solche bereichsübergreifenden Messungen existieren, gestaltet sich die Entwicklung entsprechender Testaufgaben doch als aufwendig, und die Interpretation solcher bedingten Ergebnisse ist sehr viel komplizierter als bei traditionellen Modellen, bei denen jede Aufgabe nur eine einzige Position auf einer einzigen Skala einnimmt. Da die Informationen aus gemeinsamen Items jedoch sehr wichtig sind, wird der OECD/PISA-Feldtest sowohl einen Block von vollständig integrierten Items enthalten (Items, die für mehr als einen Erhebungsbereich kodiert werden), als auch mehrere Blöcke, die separate, aber jeweils vom gleichen Text oder von der gleichen Situation ausgehende Items für Lesen, Naturwissenschaften und Mathematik umfassen.

MATHEMATISCHE GRUNDBILDUNG (Mathematical Literacy)

Definition des Gegenstandsbereichs

Bei mathematischer Grundbildung geht es um die Fähigkeit von Schülerinnen und Schülern, ihre mathematischen Kompetenzen zu nutzen, um die Herausforderungen der Zukunft zu bestehen. Der Bereich erfasst somit die Fähigkeit, mathematische Ideen in effektiver Weise zu analysieren, zu begründen und mitzuteilen, indem mathematische Probleme in einer Vielzahl von Bereichen und Situationen gestellt, formuliert und gelöst werden.

Die Definition mathematischer Grundbildung im Rahmen von OECD/PISA lautet:

„Mathematische Grundbildung ist die Fähigkeit einer Person, die Rolle zu erkennen und zu verstehen, die Mathematik in der Welt spielt, fundierte mathematische Urteile abzugeben und sich auf eine Weise mit der Mathematik zu befassen, die den Anforderungen des gegenwärtigen und künftigen Lebens dieser Person als konstruktivem, engagiertem und reflektierendem Bürger entspricht.“

Diese Definition der Reichweite mathematischer Grundbildung bedarf einiger Erläuterungen.

Mathematische Grundbildung..

Der Begriff *Grundbildung (literacy)* wurde gewählt, um zu betonen, dass mathematische Kenntnisse und Fähigkeiten, wie sie im traditionellen Curriculum der Schulmathematik definiert werden, im Rahmen von OECD/PISA nicht im Vordergrund stehen. Stattdessen liegt der Schwerpunkt auf der funktionalen Anwendung von mathematischen Kenntnissen in ganz unterschiedlichen Kontexten und auf ganz unterschiedliche, Reflexion und Einsicht erfordernde Weise. Eine solche Verwendung von Mathematik ist natürlich nur auf der Basis von umfangreichen mathematischen Grundkenntnissen und -fähigkeiten (wie sie in der Schule oft gelehrt werden) möglich. Auch im sprachlichen Bereich wird der Begriff *literacy* nicht auf einen breiten Wortschatz und umfassende Kenntnisse grammatischer, phonetischer, orthografischer usw. Regeln beschränkt, er setzt diese Kenntnisse jedoch voraus. Ebenso ist mathematische Grundbildung nicht auf die Kenntnis mathematischer Terminologien, Fakten und Verfahren sowie auf die Fähigkeit zur Durchführung bestimmter Operationen und Anwendung bestimmter Methoden einzuschränken, setzt diese aber voraus.

... die Welt ...

Mit dem Begriff *Welt* ist das natürliche, soziale und kulturelle Umfeld gemeint, in dem eine Person lebt. So stellt Freudenthal (1983) fest: „Unsere mathematischen Konzepte, Strukturen und Ideen wurden als Instrumente zur Ordnung von Phänomenen der physischen, sozialen und mentalen Welt erfunden.“

... sich befassen ...

Mit dem Ausdruck *sich befassen* sind nicht nur physische oder soziale Handlungen im engeren Sinne gemeint. Er beinhaltet auch die folgenden Aktivitäten: sich über Mathematik zu äußern, Stellung zu ihr zu nehmen, sich auf sie zu beziehen, ihre Leistungsfähigkeit zu beurteilen und sie wertzuschätzen. Die Definition mathematischer Grundbildung beschränkt sich also nicht auf die funktionale Anwendung der Mathematik, sondern umfasst auch ihre ästhetischen und spielerischen Elemente.

... des gegenwärtigen und künftigen Lebens ...

Die Wendung *des gegenwärtigen und künftigen Lebens* dieser Person bezieht sich auf deren privates, berufliches und soziales Leben mit Gleichaltrigen und Verwandten, aber auch auf ihr Leben als Bürger einer Gemeinschaft.

Eine besonders wichtige Fähigkeit, die mit dieser Auffassung von mathematischer Grundbildung angesprochen ist, ist die Fähigkeit, mathematische Probleme in ganz unterschiedlichen Bereichen und Situationen zu stellen, zu formulieren und zu lösen. Das Spektrum dieser Situationen reicht von rein mathematischen Problemen bis zu Situationen, in denen zunächst eine mathematische Struktur gar nicht erkennbar ist – in denen also diese Struktur von der Person, die das Problem stellt oder löst, erst erkannt werden muss.

Es sei noch einmal ausdrücklich darauf hingewiesen, dass es bei dieser Definition nicht nur um ein bestimmtes Minimum an mathematischem Grundwissen geht, sondern auch um die Verwendung von Mathematik in einem breiten Spektrum von unterschiedlichen Situationen.

Einstellungen und Gefühle wie zum Beispiel Selbstvertrauen, Neugier, Interesse und Wertschätzung für die Mathematik sowie der Wunsch, bestimmte Dinge zu tun oder zu verstehen, sind zwar nicht Bestandteil der OECD/PISA-Definition mathematischer Grundbildung, aber doch wichtige Voraussetzungen. Grundsätzlich ist es natürlich möglich, auch ohne derartige Einstellungen und Gefühle über mathematische Grundbildung zu verfügen. Es ist jedoch wenig wahrscheinlich, dass eine Person diese Grundbildung im zuvor definierten Sinne zur Geltung bringt, wenn ihr das Selbstvertrauen, die Neugier, das Interesse oder der Wunsch fehlt, etwas zu tun oder zu verstehen, das mathematische Komponenten enthält.

Organisation des Gegenstandsbereichs

Für die Zwecke von OECD/PISA ist es sinnvoll, eine Reihe von Aspekten mathematischer Grundbildung zu benennen. Dabei werden zwei Hauptaspekte und zwei Nebenaspekte unterschieden. Die Hauptaspekte sind:

- mathematische Kompetenzen und
- mathematische Leitideen¹.

Die Nebenaspekte sind:

- curriculare Teilbereiche der Mathematik und
- Situationen und Kontexte.

Die Hauptaspekte dienen dazu, den Bereich, der mit dieser Erhebung erfasst wird, abzustecken und Leistungen zu beschreiben. Die Nebenaspekte werden verwendet, um zu gewährleisten, dass der Bereich angemessen breit abgedeckt ist und ein ausgewogenes Spektrum von Aufgaben ausgewählt wird.

Es ist wichtig, darauf hinzuweisen, dass diese vier Aspekte nicht zu einem einzelnen Klassifizierungsschema kombiniert werden können. Zwei Aspekte, nämlich „mathematische Leitideen“ und „curriculare Teilbereiche der Mathematik“, stellen alternative Möglichkeiten der Beschreibung mathematischer Inhalte dar.

Mathematische Kompetenzen sind allgemeine Fähigkeiten und Kompetenzen wie Problemlösen, Gebrauch der mathematischen Sprache und mathematische Modellbildung.

Mathematische Leitideen sind bedeutungshaltige, stark miteinander vernetzte mathematische Konzepte, wie sie in realen Situationen und Kontexten auftreten. Einige dieser Ideen sind in der Mathematik und im allgemeinen Sprachgebrauch fest eingeführt, wie etwa *Zufall*, *Veränderung und Wachstum*, *Abhängigkeit und Beziehungen* sowie *Form*. „Leitideen“ wurden gewählt, weil sie nicht zu einer künstlichen Aufspaltung der Mathematik in einzelne, voneinander getrennte Themenbereiche führen.

Mit dem Aspekt der *curricularen Teilbereiche der Mathematik* ist der Inhalt der Schulmathematik angesprochen, wie er in vielen Lehrplänen enthalten ist. Im Rahmen von OECD/PISA werden die Bereiche *Arithmetik*, *Messen und Schätzen von Größen*, *Algebra*, *Funktionen*, *Geometrie*, *Wahrscheinlichkeit*, *Statistik* und *diskrete Mathematik* einbezogen. Die curricularen Teilbereiche werden als Nebenaspekte berücksichtigt, um sicherzustellen, dass sie in der Erhebung angemessen abgedeckt werden. Für die Auswahl der in OECD/PISA aufzunehmenden Inhalte ist jedoch der wichtigere und allgemeinere Aspekt der *mathematischen Leitideen* maßgeblich.

Der zweite Nebenaspekt betrifft die *Situationen* bzw. den allgemeinen Rahmen, in dem die mathematischen Probleme präsentiert werden. Die Situationen können sich auf schulische, berufliche, öffentliche und persönliche Bereiche beziehen.

In den folgenden Abschnitten werden die vier Aspekte im Einzelnen beschrieben.

¹ Im deutschen Sprachraum ist auch der Begriff „fundamentale Ideen“ geläufig. Es ist jedoch darauf hinzuweisen, dass diese beiden Konzepte nicht übereinstimmen (Anmerkung der Übersetzer).

Mathematische Kompetenzen

Den ersten Hauptaspekt der OECD/PISA-Rahmenkonzeption für mathematische Grundbildung bildet der Bereich „mathematische Kompetenzen“. Hierbei handelt es sich um eine nicht hierarchisch aufgebaute Liste von allgemeinen, für alle Ebenen des Lehrens und Lernens von Mathematik relevanten mathematischen Fähigkeiten. Diese Liste enthält die folgenden Elemente:

1. *Die Fähigkeit, mathematisch zu denken.* Dazu gehört: Fragen zu stellen, die für die Mathematik charakteristisch sind („gibt es...?“, „wenn ja, wie viele?“, „wie finden wir...?“); zu wissen, welche Art von Antworten die Mathematik für solche Fragen bereithält; zwischen unterschiedlichen Arten von Aussagen zu unterscheiden (Definitionen, Sätze, Vermutungen, Hypothesen, Beispiele, Bedingungen); Reichweite und Grenzen mathematischer Konzepte zu verstehen und zu berücksichtigen.
2. *Die Fähigkeit, mathematisch zu argumentieren.* Dazu gehört: zu wissen, was mathematische Beweise sind und wie sie sich von anderen Arten der mathematischen Argumentation unterscheiden; verschiedene Arten von mathematischen Argumentationsketten nachzuvollziehen und zu bewerten; heuristisches Gespür („was kann [nicht] passieren und warum?“); Entwicklung von mathematischen Argumenten.
3. *Die Fähigkeit zur mathematischen Modellierung.* Dazu gehört: den Bereich oder die Situation, die modelliert werden soll, zu strukturieren; „Mathematisierung“ (Übersetzung der „Realität“ in mathematische Strukturen); „De-Mathematisierung“ (mathematische Modelle im Rahmen der modellierten „Realität“ zu interpretieren); mit einem mathematischen Modell zu arbeiten; das Modell zu validieren; das Modell und seine Ergebnisse zu reflektieren, zu analysieren und kritisch zu beurteilen; über das Modell und seine Ergebnisse (einschließlich der Grenzen dieser Ergebnisse) zu kommunizieren; den Prozess der Modellbildung zu beobachten und zu steuern.
4. *Die Fähigkeit, Probleme zu stellen und zu lösen.* Dazu gehört: verschiedene Arten von mathematischen Problemen zu stellen, mathematische Probleme zu formulieren und zu definieren („reine“, „angewandte“, „offene“ und „geschlossene“); und verschiedene Lösungswege für unterschiedliche Arten von mathematischen Problemen zu finden.
5. *Die Fähigkeit, mathematische Darstellungen zu nutzen.* Dazu gehört: verschiedene Formen der Darstellung von mathematischen Objekten und Situationen sowie die Wechselbeziehungen zwischen diesen Darstellungsformen zu erkennen, zu interpretieren und zu unterscheiden; verschiedene Darstellungsformen je nach Situation und Zweck auszuwählen und zwischen ihnen zu wechseln.
6. *Die Fähigkeit, mit den symbolischen, formalen und technischen Elementen der Mathematik umzugehen.* Dazu gehört: die symbolische und formale Sprache zu dekodieren und zu interpretieren und ihre Beziehung zur natürlichen Sprache zu verstehen; natürliche Sprache in die symbolische/formale Sprache zu übersetzen; mit Aussagen und Ausdrücken umzugehen, die Symbole und Formeln enthalten; Variablen zu benutzen; Gleichungen zu lösen und Berechnungen vorzunehmen.
7. *Die Fähigkeit zu kommunizieren.* Dazu gehört: sich mündlich und schriftlich in verschiedenen Formen zu Sachverhalten mit mathematischem Inhalt zu äußern und entsprechende schriftliche oder mündliche Aussagen von anderen Personen zu verstehen.
8. *Die Fähigkeit, Hilfsmittel einzusetzen und zu gebrauchen.* Dazu gehört: die verschiedenen Hilfsmittel (einschließlich solche aus dem Bereich der Informationstechnologie), die bei mathematischen Aktivitäten hilfreich sein können, zu kennen und anzuwenden sowie die Grenzen dieser Hilfsmittel einzuschätzen.

Kompetenzklassen

Im Rahmen von OECD/PISA ist nicht vorgesehen, Items zu entwickeln, mit denen diese Fähigkeiten jeweils getrennt erfasst werden. Wenn man echte Mathematik betreibt, kommen gewöhnlich viele (wenn nicht alle) dieser Fähigkeiten gleichzeitig zum Einsatz, so dass jeder Versuch der Messung einzelner Fähigkeiten zu künstlichen Aufgaben und damit zu einer unnötigen Aufspaltung des Bereichs der mathematischen Grundbildung führen würde.

Um den Aspekt der *mathematischen Kompetenzen* durch die Konstruktion von Items und Tests operationalisieren zu können, ist es hilfreich, die Fähigkeiten zu drei größeren Kompetenzklassen zusammenzufassen. Diese drei Kompetenzklassen sind:

- Kompetenzklasse 1: Wiedergabe, Definitionen und Berechnungen
- Kompetenzklasse 2: Querverbindungen und Zusammenhänge herstellen, um Probleme zu lösen
- Kompetenzklasse 3: Einsichtsvolles mathematisches Denken und Verallgemeinern

Alle oben genannten Fähigkeiten dürften in jeder der drei Kompetenzklassen eine Rolle spielen; die Fähigkeiten gehören also nicht jeweils nur in eine einzige Kompetenzklasse. Die Klassen bilden ein konzeptuelles Kontinuum, das von der reinen Wiedergabe von Fakten und einfachen Rechenfähigkeiten über die Fähigkeit, Verknüpfungen zwischen verschiedenen Bereichen herzustellen, um einfache realitätsnahe Probleme zu lösen, bis zur dritten Klasse reicht, nämlich der „Mathematisierung“ von realitätsnahen Problemen und der Reflexion über Lösungen im Kontext dieser Probleme unter Anwendung mathematischen Denkens, Argumentierens und Verallgemeinerns (der Begriff der „Mathematisierung“ wird weiter unten ausführlich erläutert).

Diese Beschreibung legt nahe, dass die Kompetenzklassen insofern eine Hierarchie bilden, als Aufgaben, für die Kompetenzen der Klasse 3 erforderlich sind, im Allgemeinen schwieriger sein dürften als Aufgaben, die Kompetenzen der Klasse 2 erfordern. Dies bedeutet jedoch nicht, dass die Fähigkeiten der Klasse 2 Voraussetzung für jede Fähigkeit der Klasse 3 sind. Tatsächlich haben frühere Untersuchungen (de Lange, 1987; Shafer und Romberg, im Druck) gezeigt, dass man in Kompetenzen der Klasse 1 nicht unbedingt herausragend sein muss, um innerhalb der Kompetenzklassen 2 oder 3 gute Leistungen zu erbringen, und dass Schülerinnen und Schüler mit guten Leistungen in Kompetenzklasse 3 nicht unbedingt auch in Kompetenzklasse 1 besonders gut sein müssen.

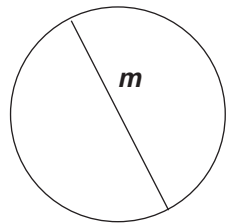
Ausgehend von der Definition mathematischer Grundbildung wird innerhalb von OECD/PISA besonderer Wert darauf gelegt, dass die Schülerinnen und Schüler Aufgaben bearbeiten, die ihre Fähigkeiten in allen drei Kompetenzklassen aufzeigen. Deshalb werden in den Tests Aufgaben aus allen drei Kompetenzklassen enthalten sein, sodass die Verantwortlichen für das Schul- und Bildungswesen sich einen Überblick darüber verschaffen können, inwieweit ihre Schulen und Curricula gewährleisten, dass die in den einzelnen Kompetenzklassen erforderlichen Fähigkeiten entwickelt werden.

Kompetenzen der Klasse 1: Wiedergabe, Definitionen, Berechnungen

In dieser Kompetenzklasse geht es um Aspekte, die in vielen standardisierten Tests und internationalen Vergleichsstudien erfasst werden. Dazu gehören Faktenwissen, Wissen um mathematische Darstellungen, Erkennen von Äquivalenzen, Abrufen der Definitionen mathematischer Objekte und Eigenschaften, Verwendung von Routineverfahren, Anwendung von Standardalgorithmen und Entwicklung von technischen Fertigkeiten. Außerdem gehören in diese Kompetenzklasse der Umgang mit Formeln und Ausdrücken, die die üblichen mathematischen Symbole enthalten, sowie Berechnungen. Zur Messung der Kompetenzen dieser Klasse können in der Regel Items mit Multiple-Choice-Format oder mit kurzer freier Antwort (wobei es nur eine richtige Antwort gibt) eingesetzt werden.

Bei der Kompetenzklasse 1 geht es insbesondere um die Fähigkeit zum Umgang mit symbolischen, formalen und technischen Elementen, die oben beschrieben wurde. Abbildung 2 zeigt Beispielaufgaben, die dieser Klasse zugeordnet sind.

Abbildung 2. **Beispiele für Aufgaben der Kompetenzklasse 1**

<p>Löse die Gleichung $7x - 3 = 13x + 15$ Was ist der Mittelwert von 7, 12, 8, 14, 15, 9? Schreibe 69% als Bruchzahl</p> <p>Die Strecke m ist der _____ des Kreises</p>	
---	---

Kompetenzen der Klasse 2: Querverbindungen und Zusammenhänge herstellen, um Probleme zu lösen

In dieser Kompetenzklasse kommt es darauf an, verschiedene Stoffgebiete und Teilbereiche der Mathematik miteinander in Beziehung zu setzen und Informationen zu verknüpfen, um einfache Probleme zu lösen. Hierzu müssen die Schülerinnen und Schüler entscheiden, welche Strategien und mathematischen Instrumente einzusetzen sind. Bei den Problemen handelt es sich zwar nicht um Routineaufgaben, eine „Mathematisierung“ ist jedoch nur auf relativ niedrigem Niveau erforderlich.

Innerhalb der Kompetenzklasse 2 wird von Schülerinnen und Schülern erwartet, dass sie je nach Situation und Zielsetzung unterschiedliche Methoden verwenden. Für das Herstellen von Zusammenhängen ist außerdem die Fähigkeit erforderlich, verschiedene Aussagen wie Definitionen, Behauptungen, Beispiele, Bedingungen und Beweise zu unterscheiden und miteinander in Beziehung zu setzen.

Diese Kompetenzklasse umfasst mehrere der oben genannten mathematischen Fähigkeiten. Zur Lösung der Beispielaufgaben ist offensichtlich ein gewisses Maß an mathematischem Denken oder Argumentieren notwendig, also die Anwendung der Fähigkeit, mathematisch zu argumentieren (siehe Abbildung 3). Darüber hinaus müssen die Schülerinnen und Schüler ein „Modell“ des Problems bilden, um es lösen zu können – die Fähigkeit zur mathematischen Modellierung ist also ebenfalls erforderlich. Beim Prozess der Problemlösung wird die Fähigkeit, Probleme zu stellen und zu lösen, benötigt. Werden bei diesem Prozess verschiedene Darstellungsformen benutzt (z. B. eine Tabelle, ein Diagramm oder eine Zeichnung), erfordert dies die Fähigkeit zur mathematischen Darstellung.

Aus der Sicht der mathematischen Sprache sind das Dekodieren und Interpretieren von symbolischer und formaler Sprache sowie das Verständnis ihrer Beziehung zur natürlichen Sprache eine weitere wichtige Fähigkeit in dieser Kompetenzklasse. Items in dieser Kompetenzklasse sind oft in einen Kontext eingebettet, der den Schülerinnen und Schülern mathematische Entscheidungen abverlangt.

Abbildung 3 enthält zwei Beispiele für Aufgaben aus dieser Kompetenzklasse. Anders als bei den Beispielen aus Kompetenzklasse 1 ist weder ihre Zugehörigkeit zu einem bestimmten curricularen Teilbereich unmittelbar offensichtlich noch welche Methode, welche Strategie oder welcher Algorithmus zur Lösung des Problems am besten geeignet wäre. In einigen Fällen hängt der curriculare Teilbereich von der Strategie ab, die die Schülerinnen und Schüler wählen, und es gibt viele gleich gute Lösungsstrategien.

Abbildung 3. Beispiele für Aufgaben der Kompetenzklasse 2

Du hast mit deinem Auto zwei Drittel der Gesamtstrecke zurückgelegt. Du bist mit einem vollen Tank losgefahren, und jetzt ist dein Tank noch zu einem Viertel gefüllt. Hast du ein Problem?

Maria lebt zwei Kilometer von der Schule entfernt, Martin fünf Kilometer. Wie weit leben Maria und Martin voneinander entfernt?

Kompetenzen der Klasse 3: Einsichtsvolles mathematisches Denken und Verallgemeinern

Um Items dieser Kompetenzklasse zu lösen, müssen die Schülerinnen und Schüler Situationen „mathematisieren“, d.h. die in der Situation enthaltene Mathematik erkennen und herausarbeiten sowie mathematische Methoden anwenden, um das Problem zu lösen. Die Schülerinnen und Schüler müssen weiterhin analysieren und interpretieren, eigene Modelle und Strategien entwickeln und mathematische Argumente, einschließlich Beweise und Verallgemeinerungen, darlegen.

Zu diesen Fähigkeiten zählt auch die Gesamtanalyse des Modells und das Reflektieren über den Modellierungsprozess. In dieser Kompetenzklasse sollten die Schülerinnen und Schüler in der Lage sein, Probleme nicht nur zu lösen, sondern auch zu stellen.

All diese Kompetenzen können nur dann richtig eingesetzt werden, wenn die Schülerinnen und Schüler in der Lage sind, in verschiedenen Formen angemessen zu kommunizieren: mündlich, schriftlich, visuell usw. Diese Kommunikation wird als Prozess verstanden, der zwei Richtungen umfasst: Schülerinnen und Schüler sollten sowohl ihre mathematischen Ideen kommunizieren können als auch die mathematischen Mitteilungen von anderen verstehen.

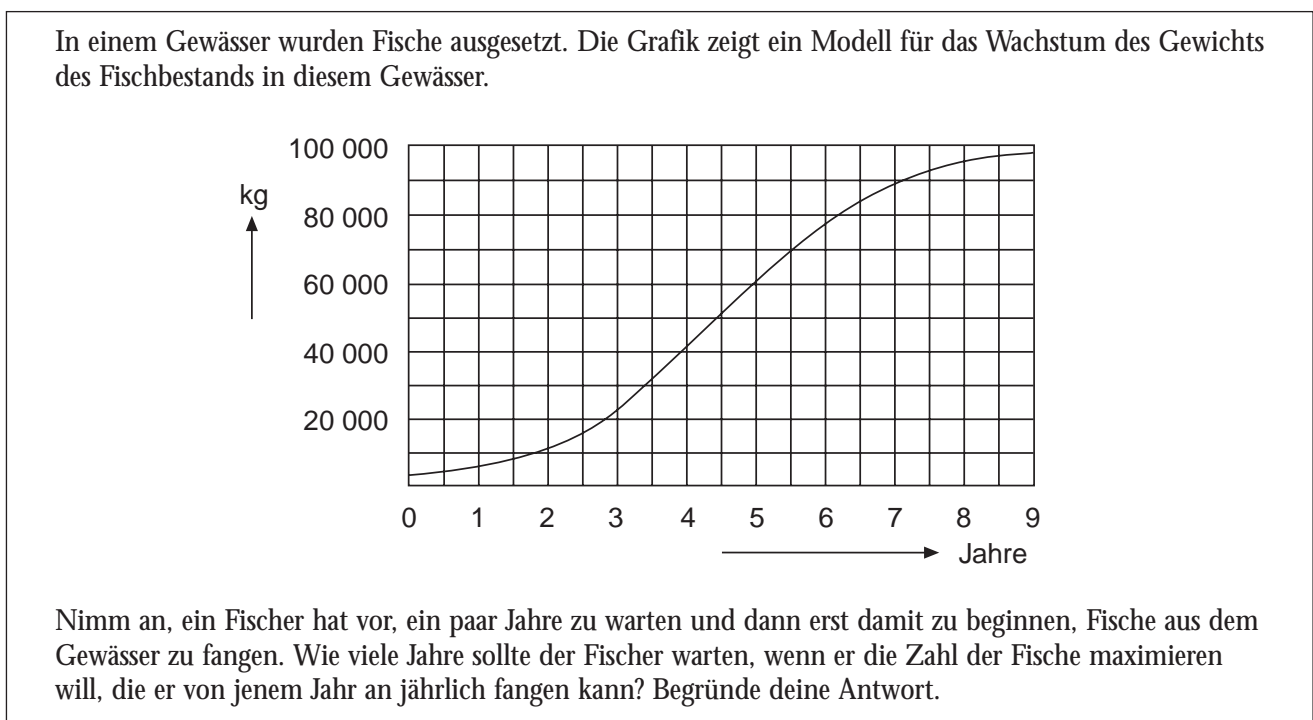
Schließlich ist wichtig, dass die Schülerinnen und Schüler auch über Einsicht in das Wesen der Mathematik verfügen müssen, einschließlich ihrer kulturellen und historischen Elemente, sowie in die Anwendung der Mathematik in anderen Kontexten und curricularen Bereichen, in denen mathematische Modellbildung eingesetzt wird.

Die Kompetenzen in Klasse 3 beinhalten oft Fähigkeiten und Kompetenzen aus anderen Klassen.

Diese Kompetenzklasse ist ein zentrales Element mathematischer Grundbildung. Leider ist sie jedoch auch am schwierigsten zu messen, vor allem innerhalb so umfangreicher Erhebungen wie OECD/PISA. Multiple-Choice-Items zum Beispiel sind zur Messung dieser Kompetenzen häufig nicht geeignet. Besser geeignet sind Fragen, die ausführlichere Antworten und mehrere Lösungsmöglichkeiten zulassen. Es ist jedoch sehr schwierig, solche Items zu entwickeln und die Schülerantworten zu bewerten. Da diese Kompetenzklasse aber ein wesentlicher Bestandteil mathematischer Grundbildung nach der im Rahmen von OECD/PISA zugrunde gelegten Definition ist, wurde versucht, sie in die Erhebung aufzunehmen und sie zumindest in begrenztem Maße zu erfassen.

Abbildung 4 zeigt ein Beispiel für eine Aufgabe der Kompetenzklasse 3.

Abbildung 4. **Beispiel für Aufgaben der Kompetenzklasse 3**



„Mathematisierung“

Der Begriff der „Mathematisierung“, wie er im Rahmen von OECD/PISA verwendet wird, bezieht sich auf das Strukturieren der Realität mithilfe mathematischer Ideen und Konzepte. Er bezeichnet die organisierende Tätigkeit, bei der erworbene Kenntnisse und Fähigkeiten angewandt werden, um unbekannte Regelmäßigkeiten, Beziehungen und Strukturen aufzuspüren (Treffers und Goffree, 1985). Dieser Prozess wird manchmal auch *horizontale Mathematisierung* genannt (Treffers, 1986) und umfasst unter anderem die folgenden Aktivitäten:

- Erkennen der Mathematik, die in einem allgemeinen Kontext enthalten ist
- Schematisieren
- Formulieren und Visualisieren eines Problems
- Entdecken von Beziehungen und Regelmäßigkeiten
- Erkennen von Ähnlichkeiten zwischen unterschiedlichen Problemen (de Lange, 1987)

Nachdem ein Problem in ein mathematisches übersetzt wurde, kann es mit mathematischen Werkzeugen gelöst werden. Das heißt, mathematische Werkzeuge können angewendet werden, um realitätsnahe Probleme, die mathematisch modelliert worden sind, zu bearbeiten und zu verfeinern. Dieser Prozess wird als *vertikale Mathematisierung* bezeichnet und ist durch folgende Aktivitäten gekennzeichnet:

- Darstellen einer Beziehung mithilfe einer Formel
- Nachweisen von Regelmäßigkeiten
- Verfeinern und Anpassen von Modellen
- Kombinieren und Integrieren von Modellen
- Verallgemeinern

Der Prozess der Mathematisierung beinhaltet also zwei Phasen: horizontale Mathematisierung, d.h. die Übersetzung der realen Welt in die mathematische Welt, und vertikale Mathematisierung, d.h. die Bearbeitung eines Problems innerhalb der mathematischen Welt und die Anwendung mathematischer Werkzeuge zur Lösung des Problems. Das Reflektieren über die Lösung im Hinblick auf das Ausgangsproblem ist ebenfalls ein wichtiger Schritt im Prozess der Mathematisierung, der jedoch selten angemessen berücksichtigt wird.

An dieser Stelle ließe sich einwenden, dass die Mathematisierung in allen Kompetenzklassen eine Rolle spielt, weil bei allen Problemen, die in einen Kontext eingebettet sind, zunächst erkannt werden muss, was mathematisch relevant ist. Die Art der Mathematisierung, die in Kompetenzklasse 3 erforderlich ist, nimmt innerhalb von OECD/PISA jedoch einen besonderen Stellenwert ein. Diese Form der Mathematisierung geht nämlich über das einfache Erkennen von bereits bekannten Problemen hinaus.

Die unterschiedlichen Komplexitätsgrade der Mathematisierung werden in den beiden folgenden Beispielen deutlich. Beide sind für Schülerinnen und Schüler zwischen 13 und 15 Jahren bestimmt und beide beziehen sich auf ähnliche mathematische Konzepte. Beim ersten ist jedoch nur eine einfache, beim zweiten eine komplexere Mathematisierung erforderlich.

Bei dem in Abbildung 5 gezeigten Beispiel handelt es sich um ein Item der Kompetenzklasse 2; in diesem Fall ist eine Mathematisierung nur auf relativ niedrigem Niveau erforderlich.

Abbildung 5. **Beispielitem mit einfacher „Mathematisierung“**

In einer Klasse sind 28 Schüler. Das Verhältnis von Mädchen zu Jungen ist 4:3. Wie viele Mädchen sind in der Klasse?

Quelle: TIMSS Mathematics Achievement in the Middle Years, S. 98.

Die für das Beispiel in Abbildung 6 erforderliche Mathematisierung hingegen entspricht der Kompetenzklasse 3. Hier müssen Schülerinnen und Schüler nicht nur die relevante Mathematik erkennen, sondern auch ein mathematisches Argument entwickeln und argumentativ darstellen können.

Abbildung 6. **Beispielitem mit komplexerer „Mathematisierung“**

Der Verteidigungshaushalt eines Landes beträgt im Jahr 1980 30 Millionen Dollar. Der Gesamthaushalt für dasselbe Jahr beträgt 500 Millionen Dollar. Im folgenden Jahr beträgt der Verteidigungshaushalt 35 Millionen Dollar, der Gesamthaushalt 605 Millionen Dollar. Die Inflationsrate lag in dem Zeitraum, für den diese beiden Haushalte galten, bei 10 Prozent.

- a) Du wirst eingeladen, für eine Organisation der Friedensbewegung einen Vortrag zu halten. Dort willst du erklären, dass der Verteidigungshaushalt in diesem Zeitraum abgenommen hat. Beschreibe, wie du das machen würdest.
- b) Du wirst eingeladen, in einer Militäarakademie einen Vortrag zu halten. Dort willst du erklären, dass der Verteidigungshaushalt in diesem Zeitraum zugenommen hat. Beschreibe, wie du das machen würdest.

Quelle: de Lange (1987), mit freundlicher Genehmigung; s. auch MSEB (1991)

„Mathematische Leitideen“

Wie bereits ausgeführt, verfolgt OECD/PISA mit der Erfassung mathematischer Grundbildung andere Zwecke als frühere vergleichende Schulleistungsstudien, bei denen es ebenfalls um Mathematik ging, insbesondere IEA/TIMSS. Einige der wichtigsten Unterschiede werden im Folgenden noch einmal beschrieben.

Bei der Entwicklung der Tests für IEA/TIMSS wurde sehr viel Wert auf eine angemessene Erfassung der Curricula der teilnehmenden Länder gelegt, und zur Beschreibung nationaler Curricula wurde ein detailliertes, auf den traditionellen curricularen Inhalten beruhendes Schema entwickelt. Im Unterricht wird die Mathematik jedoch häufig in streng voneinander getrennte Bereiche unterteilt, und Berechnungen und Formeln stehen im Vordergrund. Schülerinnen und Schülern am Ende der Sekundarstufe I ist in der Regel nicht bewusst, dass die Mathematik kontinuierlich wächst und sich auf neue Felder und *Situationen* ausweitet. Daher bezogen sich die Instrumente von IEA/TIMSS überwiegend auf mathematisches Faktenwissen, das isoliert und mit sehr kurzen Items getestet wurde.

Im Gegensatz dazu steht im Rahmen von OECD/PISA mathematische Grundbildung, wie sie oben definiert wurde, im Mittelpunkt. Ziel von OECD/PISA ist es, um es noch einmal ausdrücklich zu betonen, Schülerleistungen in ihrer ganzen Breite und im Zusammenhang zu messen und nicht nur fragmentiertes Faktenwissen abzufragen, das zur Kompetenzklasse 1 gehört. Für PISA sind daher Wechselbeziehungen und allgemeine Ideen zentral. Mathematik ist eine Sprache, die Muster beschreibt – sowohl natürliche als auch von Menschen erdachte Muster. Mathematische Grundbildung setzt voraus, dass Schülerinnen und Schüler diese Muster erkennen und ihre Vielfalt, Regelmäßigkeit und Wechselbeziehungen verstehen.

Aus diesen Gründen werden die traditionellen Inhaltsbereiche der Mathematik im Rahmen von OECD/PISA nicht als zentrale Dimension mathematischer Grundbildung behandelt. Vielmehr werden die zu testenden Inhalte anhand *mathematischer Leitideen* organisiert.

Die Konzeption der „*mathematischen Leitidee*“ ist nicht neu. Im Jahre 1990 veröffentlichte das Gremium Mathematical Sciences Education Board die Publikation *On the Shoulders of the Giant: New Approaches to Numeracy* (Senechal, 1990). Dabei handelt es sich um ein nachdrückliches Plädoyer dafür, Schülerinnen und Schülern zu helfen, tiefer in mathematisches Denken einzudringen, die der Mathematik zugrunde liegenden Konzepte zu erkennen und so zu einem besseren Verständnis der Bedeutung dieser Konzepte in der realen Welt zu gelangen. Hierzu müssen Ideen, die tief im mathematischen Denken verankert sind, erkundet werden, und zwar ohne sich durch die Grenzen, die mit den gegenwärtigen curricularen Teilbereichen gesetzt werden, beschränken zu lassen. Dieser Ansatz wird auch von anderen Mathematikern unterstützt; eine der bekannteren Publikationen hierzu ist *Mathematics: The Science of Patterns* (Devlin, 1994, 1997).

Es gibt sehr viele „mathematische Leitideen“; der Bereich der Mathematik ist so reich und so vielfältig, dass es gar nicht möglich wäre, sie alle aufzuzählen. Um den Bereich der mathematischen Grundbildung im Rahmen von OECD/PISA abzustecken, war es wichtig, solche „Leitideen“ auszuwählen, die so vielfältig und so tiefgehend sind, dass sie wesentliche Charakteristika der Mathematik repräsentieren. Um dieser Anforderung Genüge zu tun, werden im Rahmen von OECD/PISA die folgenden mathematischen Leitideen verwendet:

- Zufall
- Veränderung und Wachstum
- Raum und Form
- quantitatives Denken
- Ungewissheit
- Abhängigkeit und Beziehungen

Im ersten Zyklus von OECD/PISA kann mathematische Grundbildung aufgrund der begrenzten Testzeit selbstverständlich nicht in ihrer ganzen Breite erfasst werden. Der erste Zyklus konzentriert sich deshalb auf die folgenden beiden mathematischen Leitideen:

- Veränderung und Wachstum
- Raum und Form

Diese Beschränkung im ersten Erhebungszyklus hat vor allem zwei Gründe:

- erstens kann mit diesen beiden Ideen ein breites Spektrum von Themen aus den oben genannten Inhaltsbereichen erfasst werden und
- zweitens werden mit ihnen bestehende Curricula angemessen abgedeckt.

Quantitatives Denken wurde in diesem ersten Erhebungszyklus ausgeklammert, da sonst eine Überbetonung der typischen Rechenfertigkeiten zu befürchten wäre.

Die für den ersten Zyklus ausgewählten Ideen sollen nun ausführlicher beschrieben werden.

Veränderung und Wachstum

Jedes Phänomen unserer Umwelt beinhaltet Veränderung. Beispiele hierfür sind: Organismen, die sich beim Wachsen verändern, der Zyklus der Jahreszeiten, Ebbe und Flut, zyklische Schwankungen von Arbeitslosigkeit, Wetterveränderungen und der Dow-Jones-Index. Einige Wachstumsprozesse können direkt mit mathematischen Funktionen oder Modellen beschrieben werden; es kann sich um lineare, exponentielle, periodische, logistische, um diskrete oder kontinuierliche Prozesse handeln. Viele Prozesse aber fallen unter mehrere Kategorien, so dass oft Datenanalysen nötig sind. Mit Hilfe moderner Computertechnologien sind bessere Näherungsverfahren und ausgefeiltere Methoden zur Visualisierung von Daten entwickelt worden. Die Muster von Veränderungen in der Natur und in der Mathematik halten sich nicht an die traditionellen Inhaltsbereiche.

Sensibilität für Muster von Veränderungen umfasst laut Stuart (1990) folgende Aspekte:

- das Darstellen von Veränderungen in verständlicher Form
- das Verstehen von grundlegenden Typen der Veränderung
- das Erkennen besonderer Typen der Veränderung
- das Anwenden dieser Techniken auf die Außenwelt
- die Kontrolle eines sich verändernden Universums zu unserem eigenen Vorteil

Diese Fähigkeiten stehen in enger Beziehung sowohl zu unserer Definition mathematischer Grundbildung als auch zu den weiter oben in dieser Rahmenkonzeption definierten Fähigkeiten.

In der Leitidee von Veränderung und Wachstum sind viele verschiedene Teilbereiche der traditionellen mathematischen Gebiete enthalten. Dies ist besonders offensichtlich für die Bereiche Relationen und Funktionen und ihre

graphischen Darstellungen. Die Begriffe Reihe, Steigung, Gradient sind eng mit Funktionen verknüpft. Eine Untersuchung von Wachstumsraten für verschiedene Wachstumsphänomene führt zu linearen, exponentiellen, logarithmischen, periodischen und logistischen Wachstumskurven mit bestimmten Eigenschaften und Beziehungen untereinander. Diese wiederum führen zu Aspekten der Zahlentheorie, wie zum Beispiel den Fibonacci-Zahlen oder dem Goldenen Schnitt. Außerdem können Verbindungen zwischen diesen Ideen und geometrischen Darstellungen eine Rolle spielen.

Die Geometrie kann auch benutzt werden, um Muster in der Natur, Kunst und Architektur zu untersuchen. Ähnlichkeit und Kongruenz können dabei ebenso eine Rolle spielen wie das Wachstum einer Fläche in Relation zum Wachstum ihres Durchmessers oder Umfangs.

Wachstumsmuster können in algebraischen Formen und diese wiederum als Graphen dargestellt werden.

Wachstum lässt sich auch empirisch messen. Dabei stellt sich unter anderem die Frage, welche Schlüsse aus den Wachstumsdaten gezogen werden können, wie die Wachstumsdaten dargestellt werden können usw. Aspekte der Datenanalyse und der Statistik sind also weitere Inhaltsbereiche, die in diesem Zusammenhang angesprochen werden können.

Raum und Form

Muster sind nicht nur in Wachstums- und Veränderungsprozessen zu finden, sondern überall um uns herum: in Sprache, Musik, Videos, Verkehr, Bauwerken und Kunst. Formen wie Häuser, Kirchen, Brücken, Seesterne, Schneeflocken, Stadtpläne, Kleeblätter, Kristalle und Schatten sind Muster. Geometrische Muster können als relativ einfache Modelle für vielerlei Arten von Phänomenen dienen, und es ist möglich und erstrebenswert, sie auf allen Ebenen zu untersuchen (Grünbaum, 1985). Form ist ein zentrales, in seiner Bedeutung zunehmendes und faszinierendes Thema in der Mathematik, das eng mit der traditionellen Geometrie (allerdings weniger mit der Schulgeometrie) zusammenhängt, hinsichtlich Inhalt, Bedeutung und Methode jedoch weit über diese hinaus geht (Senechal, 1990).

Bei der Untersuchung von Form und Konstruktionen suchen wir, wenn wir die Bestandteile einer Form analysieren und Formen in unterschiedlichen Darstellungen und Dimensionen erkennen, nach Ähnlichkeiten und Unterschieden. Die Untersuchung von Formen hängt eng mit dem Konzept der „Raumerfassung“ („grasping space“, Freudenthal, 1973) zusammen. Damit ist gemeint, dass wir lernen, den Raum, in dem wir leben, zu erkennen, zu erforschen und zu beherrschen, um besser in ihm leben, atmen und uns bewegen zu können.

Um dahin zu gelangen, müssen wir in der Lage sein, die relativen Positionen von Objekten zu verstehen. Wir müssen uns bewusst sein, wie wir die Dinge sehen und warum wir sie so sehen, wie wir sie sehen. Wir müssen lernen, uns im Raum, innerhalb von Konstruktionen und Formen zu bewegen. Dies bedeutet, dass Schülerinnen und Schüler in der Lage sein sollten, die Beziehung zwischen den Formen und ihren Bildern oder visuellen Darstellungen zu verstehen, etwa die Beziehung zwischen einer Stadt und Fotografien oder Karten dieser Stadt. Sie müssen außerdem verstehen, wie dreidimensionale Objekte zweidimensional dargestellt werden können, wie Schatten sich bilden und zu interpretieren sind, was Perspektive ist und wie sie funktioniert.

Wie hier beschrieben, ist die Untersuchung von Raum und Form offen und dynamisch und fügt sich gut in das für OECD/PISA definierte Konzept mathematischer Grundbildung und mathematischer Fähigkeiten ein.

Curriculare Teilbereiche der Mathematik

Natürlich können und sollen auch die traditionellen curricularen Teilbereiche der Mathematik nicht übergangen werden. Deshalb wurden sie explizit als sekundäre Ordnungsaspekte des Bereichs mathematischer Grundbildung in OECD/PISA aufgenommen. Dieser Aspekt kann zur Ausgewogenheit der Items und zu einer angemessenen Streuung der Inhalte in Bezug auf schulische Curricula beitragen. Die curricularen Teilbereiche der Mathematik, die im Rahmen von OECD/PISA berücksichtigt werden, sind:

- Arithmetik
- Größen und Größenordnungen
- Algebra
- Funktionen
- Geometrie

- Wahrscheinlichkeit
- Statistik
- diskrete Mathematik

Diese Liste der curricularen Teilbereiche der Mathematik wurden in Zusammenarbeit mit den an PISA beteiligten Ländern entwickelt. Die OECD/PISA-Tests werden Items zu jedem dieser Teilbereiche enthalten.

Situationen

Ein wichtiger Bestandteil der Definition mathematischer Grundbildung ist die Anwendung von Mathematik in vielen verschiedenen Situationen. Es wird berücksichtigt, dass die Wahl mathematischer Methoden und Formen der Ergebnisdarstellung oft von der Einbettung abhängt, in der die Probleme präsentiert werden. Jede Situation sollte es den Schülerinnen und Schülern erlauben, sich am Prozess der Mathematisierung zu beteiligen, indem sie erkennen, wie sich die Methoden, die sie in einer bestimmten Situation gelernt haben, auch in anderen, ähnlichen Situationen erfolgreich anwenden lassen.

Situationen können im Hinblick auf ihre „Distanz“ zu den Schülerinnen und Schülern bewertet werden. Im Rahmen von OECD/PISA werden Situationen verwendet, die sich hinsichtlich dieser Distanz bzw. Nähe unterscheiden. Die größte Nähe weisen Kontexte auf, die dem persönlichen Leben entnommen sind, gefolgt von Situationen aus den Bereichen Schule, Arbeit und Sport (bzw. Freizeit im Allgemeinen) sowie aus der lokalen Gemeinschaft und Gesellschaft, wie sie Schülerinnen und Schülern im täglichen Leben begegnen. Am weitesten entfernt von Schülerinnen und Schülern sind schließlich die wissenschaftlichen Kontexte. Zu diesen gehören unter anderem Beweise für abstrakte Vermutungen, Verallgemeinerungen von numerischen oder räumlichen Mustern usw.

Hiermit wird also eine mehr oder weniger kontinuierliche Skala entwickelt, die als ein weiterer Aspekt der OECD/PISA-Rahmenkonzeption angesehen werden kann und fünf Situationen umfasst: persönlich, bildungsbezogen, beruflich, öffentlich und wissenschaftlich.

Ein mit den *Situationen* zusammenhängender Aspekt ist die Authentizität des für die Problemstellung benutzten Kontextes. Dieser Aspekt soll nun im Einzelnen behandelt werden.

Aufgabenmerkmale

In den vorangegangenen Abschnitten wurde der Bereich der mathematischen Grundbildung definiert und seine Struktur beschrieben. In diesem Abschnitt nun geht es um die Testaufgaben, die zur Erfassung der Schülerleistungen verwendet werden. Es werden die Art der Aufgaben, ihr mathematischer Kontext, ihr Format und der Prozess der Aufgabenentwicklung beschrieben.

Itemkontexte

Bei der Entwicklung der Testaufgaben muss das Problem der mathematischen Kontexte berücksichtigt werden, in die die Items eingebettet sind. Der Begriff *Kontext* wird hier entsprechend dem in der Mathematikdidaktik üblichen Gebrauch verwendet. Ein Kontext ist ein *außer-mathematischer* oder *inner-mathematischer* Rahmen, in dem die Elemente eines *komplexen mathematischen Zusammenhangs* (z.B. eines Problems, einer Aufgabe oder einer Ansammlung mathematischer Objekte, Beziehungen, Phänomene usw.) interpretiert werden sollen. Es handelt sich entweder um einen Rahmen, in dem der jeweilige mathematische Komplex bereits enthalten ist (inner-mathematischer Kontext), oder um einen Rahmen, der sich zur Aktivierung dieses mathematischen Komplexes eignet und in den dieser Komplex dann eingebettet wird (außer-mathematischer Kontext). Die Einbettung eines mathematischen Komplexes in eine außer-mathematische Situation setzt immer das explizite oder implizite (unausgesprochene) Vorhandensein eines mathematischen Modells voraus. Dies stellt die Übersetzung dieser Situation (bzw. bestimmter Aspekte dieser Situation) in dem betreffenden mathematischen Komplex dar.

Diese Definition lässt Raum für ein breites Spektrum von Kontexten. Kontexte können anderen Fächern entnommen werden, dem Berufsleben, dem Alltag, dem Leben in der Gemeinschaft und Gesellschaft usw. Auch Freizeitkontexte wie Sport und Spiel fallen darunter. Weiterhin ist die in früheren Abschnitten als Neben aspekt

definierte „Situation“ eine Form des Kontextes. In den Testaufgaben werden ganz unterschiedliche Kontexte verwendet. Dies ist nötig, um kulturelle Vielfalt zu gewährleisten und um das ganze Spektrum der Rollen zu erfassen, die die Mathematik spielen kann.

Die in den Tests von OECD/PISA verwendeten Kontexte werden schwerpunktmäßig authentisch sein. Ein Kontext wird dann als authentisch angesehen, wenn er innerhalb der tatsächlichen Erfahrungen und Praktiken der Teilnehmer in realen Zusammenhängen angesiedelt ist. Diese Definition setzt jedoch nicht voraus, dass die Schülerinnen und Schüler, die getestet werden, aktiv an diesen Zusammenhängen teilnehmen. Fragen nach den Erträgen von Sparguthaben, die zu realistischen Zinssätzen bei einer Bank angelegt wurden, können durchaus authentisch sein, auch wenn sie außerhalb des aktuellen Erfahrungsbereichs der zu testenden Schülerinnen und Schüler liegen.

Zu berücksichtigen ist auch, dass ein Kontext nicht schon deshalb authentisch ist, weil er realitätsnahe Elemente enthält. So enthalten die Aufgaben in Abbildungen 7 und 8 zwar realitätsnahe Elemente. Sie sind jedoch nicht authentisch, weil sich diese Probleme in einem außerschulischen Rahmen so niemandem stellen würden. Die Kontexte dieser Aufgaben sind gewählt worden, um auf der Oberfläche eine realitätsnahe Einbettung vorzutäuschen. Im Rahmen von OECD/PISA werden Kontexte dieses Typs so weit wie möglich vermieden.

Abbildung 7. **Eine realitätsnahe, aber nicht authentische Aufgabe**

Was kostet ein T-Shirt?
Was kostet ein Getränk?
Gib an, welche Überlegungen zu deinen Antworten geführt haben.

Abbildung 8. **Beispielitem mit einem konstruierten Kontext**

Welche der folgenden Gleichungen könnte benutzt werden, um folgendes Problem zu lösen?
Bill wog letzten Sommer 107 Pfund. Er nahm 4 Pfund ab, nahm dann aber wieder 11 Pfund zu. Wieviel wiegt er jetzt?

- a) $107 - (4+11) = A$
- b) $(107 - 4) + 11 = A$
- c) $(107 + 11) + 4 = A$
- d) $- 4 + 11 = 107 + A$
- e) $(107 - 11) + 4 = A$

Die Betonung der Authentizität von Kontexten im Rahmen von OECD/PISA schließt jedoch nicht aus, dass auch wichtige und/oder interessante inner-mathematische Kontexte (die mitunter auch „virtuell“ sein können) aufgenommen werden. Ein Beispiel hierfür ist die Aufgabe in Abbildung 9. Hier handelt es sich um einen künstlichen bzw. verallgemeinerten Kontext, der authentisch sein mag oder auch nicht. Solche Kontexte werden, wenn sie mathematisch interessant und relevant sind, auch in den OECD/PISA-Tests vorkommen. Eine der größten Stärken der Mathematik liegt ja gerade darin, dass sie benutzt werden kann, um hypothetische Szenarien zu erklären und potentielle Systeme oder Situationen zu untersuchen, auch wenn diese in der Realität nicht möglich wären.

Abbildung 9. **Beispielitem mit einem „virtuellen“ Kontext**

Wäre es möglich, ein Münzsystem (oder ein Briefmarkensystem) aufzubauen, dem nur die Einheiten 3 und 5 zugrunde liegen?
Welche Geldbeträge könnte man mit diesem System bezahlen? Falls ein solches System möglich ist, wäre es auch wünschenswert?

Aufgabenformate

Bei der Entwicklung von Testinstrumenten muss genau bedacht werden, welche Auswirkungen die Aufgabenformate auf die Leistungen der Schülerinnen und Schüler und damit auf die Definition des zu testenden Konstrukts haben. Diese Frage ist in einem Projekt wie OECD/PISA besonders wichtig, da hier das Spektrum der in Frage kommenden Formate durch den groß angelegten, länderübergreifenden Rahmen erheblich eingeschränkt ist.

Wie die Lesekompetenz soll auch die mathematische Grundbildung im Rahmen von OECD/PISA mit einer Kombination aus Items mit Multiple-Choice-Formaten und Items mit offenem Antwortformat gemessen werden. Dabei wird es unter den Items mit offenem Antwortformat sowohl Aufgaben geben, für die es nur eine richtige Lösung gibt (offene Frage/eine richtige Antwort), als auch Fragen, die mehrere richtige Antwortmöglichkeiten haben (offene Frage/mehrere richtige Antworten). Anhang 2 enthält darüber hinaus Überlegungen zu einem breiteren Spektrum von Antwortformaten, die im zweiten Zyklus der Erhebung eingesetzt werden könnten, wenn Mathematik der Hauptbereich ist.

Im Zusammenhang mit der zweiten Mathematik-Studie der IEA stellen Travers und Westbury (1989) fest: „Die Konstruktion und Auswahl von Multiple-Choice-Items für die unteren Ebenen kognitiven Verhaltens – Rechnen und Verstehen – war nicht schwierig.“ Aber, so fahren sie fort, „auf den höheren Ebenen ergaben sich Schwierigkeiten“. Multiple-Choice-Formate sind zwar sinnvoll (siehe das Beispiel in Abbildung 10), aber nur in begrenztem Umfang und nur zur Erfassung von Zielen (oder Verhaltensweisen) und Lernergebnissen auf der untersten Ebene. Für Ziele auf höheren Ebenen und für komplexere Prozesse sollten eher andere Testformate verwendet werden, wobei die einfachste Strategie der Einsatz offener Fragen ist.

Offene Items, die nur eine richtige Antwort haben, stellen ähnliche Fragen wie Multiple-Choice-Items. Sie sind dadurch gekennzeichnet, dass die Schülerinnen und Schüler eine Antwort frei formulieren müssen, die dann leicht als falsch oder richtig beurteilt werden kann. Wenn die Antworten nicht maschinell ausgewertet werden, ist dies das bevorzugte Format für die Erfassung von Kompetenzklasse 1. Anders als bei Multiple-Choice-Items spielt hier das Erraten richtiger Antworten in der Regel keine große Rolle, und es müssen keine Distraktoren (die das zu testende Konstrukt beeinflussen) verwendet werden. Ein Beispiel für ein offenes Item, bei dem es nur eine richtige Antwort und viele mögliche falsche Antworten gibt, findet sich in Abbildung 11.

Items, die mehrere richtige Antworten haben, müssen die Schülerinnen und Schüler ausführlicher beantworten, und der Antwortprozess erfordert in der Regel Aktivitäten höherer Ordnung. Bei solchen Items müssen die Schülerinnen und Schüler oft nicht nur eine Antwort formulieren, sondern auch die Schritte angeben, die zu ihr geführt haben, bzw. erklären, wie sie zu der Antwort gekommen sind. Die wichtigste Eigenschaft dieser Art von Items ist, dass sie es Schülerinnen und Schülern erlauben, Lösungen auf verschiedenen Ebenen mathematischer Komplexität zu entwickeln. Abbildung 12 zeigt ein Beispiel.

Abbildung 10. **Beispielitem mit einer begrenzten Zahl von Antwortmöglichkeiten**

Eine Robbe muss atmen, auch wenn sie schläft. Martin hat eine Robbe eine Stunde lang beobachtet. Zu Beginn seiner Beobachtung tauchte die Robbe zum Meeresboden und begann zu schlafen. Nach 8 Minuten ließ sie sich langsam zurück an die Oberfläche treiben und holte Atem.

Drei Minuten später war sie wieder auf dem Meeresboden, und der ganze Prozess fing von vorne an.

Nach einer Stunde war die Robbe:

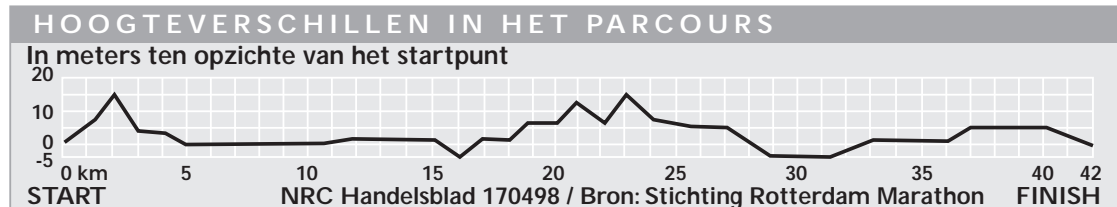
- a) auf dem Meeresboden
- b) auf dem Weg nach oben
- c) beim Atemholen
- d) auf dem Weg nach unten

Abbildung 11. **Beispielitem mit einer richtigen Antwort und vielen falschen Antworten**

Tepla Louroupe gewann 1998 den Rotterdamer Marathonlauf. „Es war einfach“, sagte sie, „die Strecke war ganz eben.“

Hier siehst du eine Grafik mit den Höhenunterschieden der Rotterdamer Marathonstrecke:

[Höhenunterschiede der Strecke relativ zum Startpunkt – in Metern]



Wie groß war der Unterschied zwischen der höchsten und der niedrigsten Stelle der Strecke?

Abbildung 12. **Beispielitem mit einer frei formulierten Antwort**

Indonesien liegt zwischen Malaysia und Australien. Die folgende Tabelle zeigt einige Daten zur Bevölkerung Indonesiens und ihrer Verteilung auf die Inseln.

Region	Fläche (km ²)	Prozent der Gesamtfläche	Bevölkerung im Jahre 1980 (in Mio.)	Prozent der Gesamtbevölkerung
Java/Madura	132 187	7,50	91,281	65,21
Sumatra	473 606	26,88	27,981	19,99
Kalimantan (Borneo)	539 460	30,62	6,721	4,80
Sulawesi (Celebes)	189 216	10,74	10,377	7,41
Bali	5 561	0,31	2,470	1,77
Irian Jaya	421 981	23,95	1,145	0,82
INSGESAMT	1 762 011	100,00	139,975	100,00

Eines der Hauptprobleme Indonesiens ist die ungleichmäßige Verteilung der Bevölkerung auf die Inseln. Der Tabelle ist zu entnehmen, dass auf der Insel Java, die nur 7,5 % der Gesamtfläche einnimmt, etwa 65 % der Bevölkerung leben.

Frage: Entwerfe eine grafische Darstellung (oder mehrere grafische Darstellungen), die die ungleichmäßige Verteilung der indonesischen Bevölkerung zeigt.

Quelle: de Lange und Verhage (1992), mit freundlicher Genehmigung.

In der OECD/PISA-Studie werden 25–35% der Mathematik-Testzeit auf offene Items mit mehreren richtigen Antworten entfallen. Die Bewertung dieser Items muss von geschulten Kodierern vorgenommen werden. Diese müssen in der Lage sein, Kodieranweisungen zu verwenden, die ein gewisses Maß an professionellem Urteilsvermögen erfordern. Da bei diesen Items die Gefahr besteht, dass verschiedene Kodierer zu unterschiedlichen Ergebnissen kommen, wird OECD/PISA Untersuchungen der Reliabilität der Kodierungen durchführen, um das Ausmaß der Nichtübereinstimmung zu kontrollieren. Erfahrungen mit Untersuchungen dieses Typs haben gezeigt, dass es durchaus möglich ist, klare Kodieranweisungen zu entwickeln und zuverlässige Bewertungen zu erzielen.

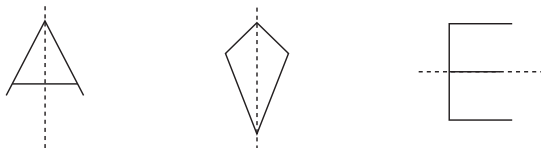
Im ersten Erhebungszyklus von OECD/PISA wird ein Aufgabenformat verwendet, bei dem mehrere Items an ein gemeinsames Stimulusmaterial anknüpfen. Indem zunehmend komplexere Fragen gestellt werden, wird den Schülerinnen und Schülern Gelegenheit gegeben, sich eingehender mit einem Kontext oder Problem zu befassen. Die ersten Fragen sind meistens Multiple-Choice-Items oder offene Items mit einer einzigen richtigen Antwort, auf die dann in der Regel offene Items mit mehreren richtigen Antworten folgen. Dieses Aufgabenformat ist für alle Kompetenzklassen geeignet.

Ein Grund für die Verwendung solcher Aufgabenformate ist, dass mit ihnen realistische Aufgaben entwickelt werden können, die die Komplexität realitätsnaher Situationen widerspiegeln. Ein weiterer Grund ist die effiziente Nutzung der Testzeit, da diese Aufgabenformate die Zeit reduzieren, die für den „Einstieg“ in den mit der Situation gegebenen Stoff benötigt wird. Bei der Entwicklung der OECD/PISA-Aufgaben wird darauf geachtet, dass jeder einzelne Punkt, der innerhalb einer solchen Aufgabe erreicht werden kann, von allen anderen innerhalb der Aufgabe zu erzielenden Punkten unabhängig ist. Wichtig ist außerdem, Verzerrungen zu minimieren, die aufgrund der Verwendung einer zu geringen Anzahl verschiedener Situationen entstehen könnten. Dies wird bei der Konstruktion des Tests ebenfalls beachtet.

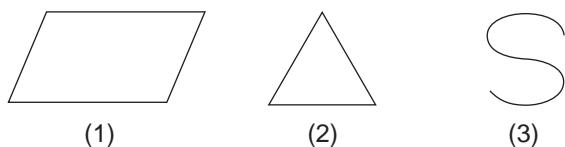
Abbildung 13 zeigt eine Aufgabe mit dem hier beschriebenen Format.

Abbildung 13. **Eine Aufgabe mit mehreren Items**

Wenn eine Figur so zusammengefaltet werden kann, dass die beiden Hälften genau aufeinander liegen, ist die Faltlinie eine Symmetrieachse.




Frage A
Welche der folgenden Figuren besitzen Symmetrieachsen?



(1) (2) (3)

Frage B
Zeichne im nebenstehenden Quadrat alle Symmetrieachsen ein.



Frage C
Welche der ersten acht Buchstaben des Alphabets haben genau zwei Symmetrieachsen?

Frage D
Hans sagte: „Ich kenne eine Regel, nach der man beurteilen kann, wann eine Figur mit vier Seiten eine Symmetrieachse hat: Wenn nämlich die Dreiecke auf beiden Seiten der Linie die gleiche Größe und die gleiche Form haben, gibt es eine Symmetrieachse, an der entlang die Figur zusammengefaltet werden kann.“ Erkläre, warum du Hans zustimmst oder nicht zustimmst.

Erhebungsstruktur

In diesem Abschnitt wird die Struktur des mathematischen Teils in den OECD/PISA-Testheften für den ersten Erhebungszyklus beschrieben, in dem insgesamt 60 Minuten Testzeit für Mathematik vorgesehen sind.

Im ersten Erhebungszyklus wird die Testzeit gleichmäßig auf die beiden mathematischen Leitideen „Veränderung und Wachstum“ bzw. „Raum und Form“ verteilt, während die Verteilung auf die drei Kompetenzklassen ungefähr 1:2:1 sein wird. Diese Information wird in Tabelle 9 zusammengefasst, die die Anzahl der Items für jede Leitidee und jede Kompetenzklasse angibt. Die Items selbst sind danach aufgeschlüsselt, ob sie durch mehrere Kodierer bewertet werden müssen.

Auf die Nebenaspekte „curriculare Teilbereiche“ und „Situationen“, die in Tabelle 9 nicht aufgeführt sind, verteilen sich die Items ungefähr gleichmäßig. Im Rahmen von OECD/PISA haben also alle oben genannten neun curricularen Teilbereiche und fünf Situationen gleiches Gewicht.

Tabelle 9. Empfohlene Verteilung der Items und der zu erzielenden Punkte auf die fundamentalen Ideen und Kompetenzklassen

Itemtyp	Veränderung und Wachstum			Raum und Form		
	Kompetenzklassen			Kompetenzklassen		
	Klasse 1	Klasse 2	Klasse 3	Klasse 1	Klasse 2	Klasse 3
Ein Kodierer	6(6)	5(5)		6(6)	5(5)	
Mehrere Kodierer		2(5)	2(5)		2(5)	2(5)

Anm.: Die Zahl in Klammern ist die zu erreichende Punktzahl.

Wie aus Tabelle 9 hervorgeht, wird der Test voraussichtlich folgendes umfassen:

- 15 Items für jede der beiden Leitideen
- 21 Punktwerte für jede der beiden Leitideen
- 8 Items mit Mehrfachkodierung und 22 Items mit Einfachkodierung
- 12 Punktwerte in Kompetenzklasse 1, 20 Punktwerte in Kompetenzklasse 2 und 10 Punktwerte in Kompetenzklasse 3

Langfristig sollen die Kompetenzklassen 2 und 3 im Rahmen von OECD/PISA größeres Gewicht erhalten als im ersten Erhebungszyklus, in dem die Testzeit für Mathematik sehr begrenzt ist.

Nachstehend wird die Zusammensetzung eines typischen halbstündigen Moduls für die Erfassung mathematischer Grundbildung im ersten Zyklus beschrieben:

- eine kleine Anzahl (2-4) Multiple-Choice-Items oder offene Items mit einer einzigen richtigen Antwort zur Erfassung von Kompetenzklasse 1 oder 2
- eine kleine Anzahl (1-2) von Aufgaben mit jeweils 2 oder 3 Items innerhalb ein und desselben Kontexts zur Erfassung von Kompetenzklasse 1 oder 2
- ein Block von mehreren Items innerhalb eines Kontexts; die Items werden mit relativ einfachen Aufgaben zur Erfassung von Kompetenzklasse 1 beginnen und dann zu komplexeren Aufgaben zur Erfassung von Kompetenzklasse 3 übergehen

Skalentwicklung

Eine wesentliche Voraussetzung dafür, dass die mit OECD/PISA angestrebten Ziele erreicht werden können, ist die Entwicklung von Skalen zur Beschreibung von Schülerleistungen. Der Prozess der Skalentwicklung ist insofern iterativ, als erste Entwürfe, die auf bereits vorliegenden Erfahrungen und Untersuchungen im Bereich Lernen und kognitive Entwicklung in der Mathematik basieren, auf Grundlage der im Feldtest von OECD/PISA gewonnenen Daten weiterentwickelt werden sollen.

Die Entscheidung, welche Skalen zur Berichterstattung über die Ergebnisse für mathematische Grundbildung zu verwenden sind, muss noch getroffen werden. Die naheliegendsten Alternativen sind die folgenden: i) eine einzige Skala für mathematische Grundbildung insgesamt, ii) eine separate Skala für jede Leitidee oder iii) eine separate Skala für jede Kompetenzklasse. Die Entscheidung für eine dieser drei Alternativen wird nach Auswertung der Daten aus dem Feldtest getroffen.

Um den Informationsgehalt der OECD/PISA-Ergebnisse weiter zu erhöhen, werden zumindest bei einigen Items nicht einfach nur Punktwerte für richtige Antworten vergeben, sondern auch die von den Schülerinnen und Schülern jeweils gewählten Strategien kodiert. Beispielsweise wäre es durchaus nicht überraschend, wenn Schülerinnen und Schüler aus zwei Ländern zwar auf der Skala für mathematische Grundbildung gleiche Punktwerte erzielen, sich aber erheblich darin unterscheiden, wie sie dieses Niveau erreicht haben: In dem einen Land könnten sich die Schülerinnen und Schüler mehr an formale Strategien und Routineverfahren halten, während in dem anderen Land möglicherweise eher informelle, auf gesundem Menschenverstand beruhende, Strategien gewählt werden.

Sonstiges

Verknüpfungen mit anderen Erhebungsbereichen

OECD/PISA setzt andere Schwerpunkte als frühere Leistungsvergleiche im Bereich Mathematik, etwa die Third International Mathematics and Science Study (IEA/TIMSS). Während die in IEA/TIMSS eingesetzten Instrumente auf dem gemeinsamen Nenner der nationalen Curricula aller Teilnehmerländer basieren, soll im Rahmen von OECD/PISA mathematische Grundbildung im oben definierten Sinne gemessen werden. Wer IEA/TIMSS kennt, wird feststellen, dass zwischen den „mathematischen Kompetenzen“ von OECD/PISA und den „Leistungserwartungen“ („performance expectations“) von IEA/TIMSS gewisse Beziehungen bestehen und dass die im Rahmen von OECD/PISA unterschiedenen „curricularen Teilbereiche der Mathematik“ Ähnlichkeiten mit den in IEA/TIMSS verwendeten Curriculuminhalten aufweisen. Im Rahmen von OECD/PISA werden die curricularen Teilbereiche jedoch den Leitideen untergeordnet, die den Bereich der mathematischen Grundbildung abgrenzen. Außerdem gehören die meisten Items im IEA/TIMSS-Leistungstest zu Kompetenzklasse 1, während die OECD/PISA-Tests auch die Kompetenzklassen 2 und 3 erfassen sollen. Aus diesen Gründen ist nicht zu erwarten, dass eine psychometrische Verknüpfung der Skalen von IEA/TIMSS und OECD/PISA möglich sein wird.

Hilfsmittel

Es gibt drei Möglichkeiten, mit der Frage der Benutzung von Taschenrechnern und anderen Hilfsmitteln im Rahmen von OECD/PISA umzugehen:

- die Benutzung von Taschenrechnern wird grundsätzlich untersagt
- die Benutzung von Taschenrechnern wird auf einen im Rahmen des OECD/PISA-Tests zur Verfügung gestellten Rechner beschränkt
- die Benutzung von eigenen Taschenrechnern und sonstigen Hilfsmitteln wird freigestellt

OECD/PISA hat sich für die dritte Möglichkeit entschieden. Sie wurde gewählt, weil sie die authentischste Messung dessen ermöglicht, was Schülerinnen und Schüler leisten können, und für den Vergleich der Leistungsfähigkeit von Bildungssystemen am informativsten ist. Die in einem Bildungssystem getroffene Entscheidung, Schülerinnen und Schülern den Gebrauch von Taschenrechnern zu erlauben, unterscheidet sich nicht prinzipiell von Entscheidungen über andere Aspekte des Unterrichts, die in Bildungssystemen getroffen und von OECD/PISA nicht kontrolliert werden.

Im Übrigen ist das Argument, bei den ersten beiden Optionen wären die Vergleiche gerechter, weil die Testbedingungen scheinbar identisch sind, nur bei oberflächlicher Betrachtung richtig. Schülerinnen und Schüler, die es gewohnt sind, Aufgaben mit Hilfe eines Taschenrechners zu lösen, sind benachteiligt, wenn ihnen dieses Hilfsmittel genommen wird. Schülerinnen und Schüler, die einen Rechner erhalten, mit dem sie nicht vertraut sind, können diesen andererseits unter Umständen nicht effektiv nutzen oder lassen sich möglicherweise dazu verleiten, ihn unnötig oder unangemessen einzusetzen. Das einfache Problem $6+4\cdot 3=?$ etwa ist für viele Schülerinnen und Schüler mit einem einfachen Taschenrechner schwerer zu lösen als ohne, insbesondere wenn sie mit Taschenrechnern nicht vertraut sind.

Deshalb wurde die Entscheidung getroffen, Schülerinnen und Schülern im Rahmen von OECD/PISA die Nutzung von Taschenrechnern und anderen, in ihren Schulen normalerweise benutzten Hilfsmitteln zu erlauben. Die Testitems werden jedoch so gewählt, dass der Gebrauch von Taschenrechnern die Testleistungen nicht verbessern dürfte.

NATURWISSENSCHAFTLICHE GRUNDBILDUNG (Scientific Literacy)

Eine für die eigenverantwortliche Lebensgestaltung junger Menschen wichtige Kompetenz ist die Fähigkeit, aus vorliegenden Informationen und Befunden angemessene und vorsichtige Schlussfolgerungen zu ziehen, Behauptungen anderer Personen anhand der angeführten Belege zu kritisieren und durch Belege gestützte Aussagen von bloßen Meinungen zu unterscheiden. Den Naturwissenschaften kommt hier eine besondere Rolle zu, da sie sich mit der rationalen Überprüfung von Ideen und Theorien anhand von Befunden aus unserer Umwelt befassen. Dies heißt nicht, dass die Naturwissenschaften Kreativität und Phantasie ausschließen, die für die Weiterentwicklung unseres Verständnisses der Welt schon immer eine zentrale Rolle gespielt haben. Auf manche Ideen, die „plötzlich aufgetaucht“ zu sein scheinen, ist man nur aufgrund eines Vorgangs gekommen, den Einstein als den „Weg der Intuition“ beschrieben hat, „dem ein Gefühl für die hinter den Dingen liegende Ordnung zu Hilfe kommt“ (Einstein, 1933). Welche Ideen man dann aber aufgreifen konnte, hing von ihrer sozialen Annehmbarkeit zu der betreffenden Zeit ab, so dass Entwicklungen naturwissenschaftlichen Wissens nicht nur von der Kreativität einzelner Menschen abhängen, sondern immer auch von der Kultur, in der sie erdacht werden. Wird der kreative Sprung aber vollzogen und ein neuer theoretischer Bezugsrahmen formuliert, muss dieser anschließend äußerst sorgfältig an der Realität überprüft werden. Wie Hawking (1988) schrieb:

„Gut ist eine Theorie, wenn sie zwei Voraussetzungen erfüllt: Sie muss eine grosse Klasse von Beobachtungen auf der Grundlage eines Modells beschreiben, das nur einige wenige beliebige Elemente enthält, und sie muss bestimmte Voraussagen über die Ergebnisse künftiger Beobachtungen ermöglichen.“ (Hawking, 1991, S.23-24)

Theorien, die diese Voraussetzungen nicht erfüllen – oder nicht überprüft werden können –, sind keine naturwissenschaftlichen Theorien. Ein gebildeter Bürger muss in der Lage sein, zwischen Fragen, die von den Naturwissenschaften beantwortet werden können, und Fragen, bei denen dies nicht der Fall ist, ebenso zu unterscheiden wie zwischen Wissenschaft und Pseudowissenschaft.

Definition des Gegenstandsbereichs

Aktuelle Auffassungen über die anzustrebenden Ergebnisse des naturwissenschaftlichen Unterrichts betonen die Entwicklung eines allgemeinen Verständnisses von wichtigen naturwissenschaftlichen Konzepten und Erklärungsmodellen, von den Methoden, mit denen die Naturwissenschaften ihre Erkenntnisse überprüfen und stützen, und von den Möglichkeiten und Grenzen der Naturwissenschaften in der modernen Welt. Als wichtig wird außerdem die Fähigkeit angesehen, dieses Verständnis in realen, mit naturwissenschaftlichen Fragen verbundenen Situationen anzuwenden, in denen Behauptungen geprüft und Entscheidungen getroffen werden müssen. So besteht für Millar und Osborne (1998) das zentrale Anliegen eines modernen naturwissenschaftlichen Curriculums in der Entwicklung der „Fähigkeit, naturwissenschaftliche und technische Informationen zu lesen, sich anzueignen und in ihrer Bedeutung einzuschätzen“. Weiter heißt es bei ihnen:

„Bei diesem Ansatz liegt die Betonung weniger darauf, Naturwissenschaft praktisch zu betreiben. Es geht auch nicht darum, naturwissenschaftliches Wissen zu erzeugen oder naturwissenschaftliches Wissen für irgendeine Prüfung kurz aus dem Gedächtnis zu reproduzieren. ... Schülerinnen und Schüler sollten im naturwissenschaftlichen Unterricht angehalten werden, zu zeigen, dass sie in der Lage sind, Belege zu beurteilen, zwischen Theorien und Beobachtungen zu unterscheiden und den Grad der Gewissheit einzuschätzen, der den aufgestellten Behauptungen zugeschrieben werden kann.“ (Millar und Osborne, 1998)

Dies also sollte der Ertrag des naturwissenschaftlichen Unterrichts für alle Schülerinnen und Schüler sein. Bei einigen Schülerinnen und Schülern, jener Minderheit, die einmal die Naturwissenschaftler von morgen sein werden, werden diese grundlegenden Fähigkeiten dann durch eine intensivere Beschäftigung mit naturwissenschaftlichen Ideen vertieft und durch die Entwicklung der Fähigkeit, selbst naturwissenschaftlich zu arbeiten, erweitert.

Aus diesen Überlegungen kann gefolgert werden, dass das wesentliche Ergebnis naturwissenschaftlichen Unterrichts, das auch im Rahmen von OECD/PISA im Mittelpunkt stehen sollte, im Aufbau von *naturwissenschaftlicher Grundbildung (scientific literacy)* besteht. Dieser Begriff ist in ganz unterschiedlichen Zusammenhängen verwendet worden. Das *International Forum on Scientific and Technological Literacy for All* (UNESCO, 1993) zum Beispiel hat eine ganze Reihe von Sichtweisen vorgeschlagen. Eine davon lautet wie folgt:

„Die Fähigkeit, sich mit Verstand und Zutrauen auf angemessenen Niveaus mit der von Menschen gestalteten Welt und mit der Welt der naturwissenschaftlichen und technischen Ideen auseinanderzusetzen.“ (UNESCO, 1993)

In den vielen verschiedenen Auffassungen naturwissenschaftlicher Grundbildung (einen Überblick gibt Shamos, 1955; siehe auch Graeber und Bolte, 1997) sind immer auch Vorstellungen von den Ebenen dieser Grundbildung enthalten. So hat Bybee (1997) eine Unterscheidung von vier Ebenen vorgeschlagen. Die beiden untersten dieser Ebenen sind die „nominelle naturwissenschaftliche Grundbildung“, bei der es um die Kenntnis von Namen und Begriffen geht, und die „funktionale Grundbildung“, über die Personen verfügen, wenn sie naturwissenschaftliches Vokabular in eingeschränkten Zusammenhängen anwenden können. Diese beiden Ebenen der Grundbildung werden als zu niedrig erachtet, um als Bezugspunkte innerhalb der OECD/PISA-Rahmenkonzeption zu dienen. Die höchste der von Bybee genannten Ebenen, die „multidimensionale naturwissenschaftliche Grundbildung“, umfasst ein Verständnis der Besonderheiten von Naturwissenschaften sowie ihrer Geschichte und ihrer kulturellen Rolle auf einer Ebene, die eher für naturwissenschaftliche Experten angemessen ist als für die Allgemeinheit. Möglicherweise ist die Annahme, dass naturwissenschaftliche Grundbildung Denken auf diesem spezialisierten Niveau voraussetzt, für die Schwierigkeit verantwortlich, eine leichter zugängliche Vorstellung des Konzepts zu vermitteln. Für die Zwecke der OECD/PISA-Rahmenkonzeption scheint am ehesten Bybees dritte Ebene angemessen, die „konzeptuelle und prozedurale naturwissenschaftliche Grundbildung“.

Nachdem eine Reihe von bereits vorliegenden Beschreibungen in Erwägung gezogen wurden, definiert OECD/PISA naturwissenschaftliche Grundbildung (*scientific literacy*) wie folgt:

„Naturwissenschaftliche Grundbildung ist 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.“

Dies soll im Folgenden näher erläutert werden.

Naturwissenschaftliche Grundbildung (scientific literacy) ...

Hier ist ausdrücklich darauf hinzuweisen, dass naturwissenschaftliches Wissen (im Sinne eines Wissens über die Naturwissenschaften) und die Prozesse, mit denen dieses Wissen entwickelt wird, nicht nur wesentliche Bestandteile naturwissenschaftlicher Grundbildung darstellen, sondern dass sie innerhalb der hier verwendeten Konzeption letztlich nicht voneinander zu trennen sind. Prozesse sind, wie sogleich noch ausführlicher zu erläutern sein wird, nur dann *naturwissenschaftliche* Prozesse, wenn sie bezogen auf naturwissenschaftliche Inhalte angewendet werden. Deshalb setzt die Anwendung von naturwissenschaftlichen Prozessen notwendigerweise ein gewisses Verständnis naturwissenschaftlicher Inhalte voraus. Die im Rahmen von OECD/PISA zugrunde gelegte Auffassung naturwissenschaftlicher Grundbildung trägt dieser Verbindung von Denken und Wissen über naturwissenschaftliche Aspekte der Welt Rechnung.

... naturwissenschaftliches Wissen anzuwenden, naturwissenschaftliche Fragen zu erkennen und aus Belegen Schlussfolgerungen zu ziehen ...

Nach dieser Definition bedeutet naturwissenschaftliches Wissen weit mehr als bloßes Faktenwissen und die Kenntnis von Bezeichnungen und Begriffen. Es umfasst ein Verständnis von grundlegenden naturwissenschaftlichen Konzepten, von den Grenzen des naturwissenschaftlichen Wissens und von den Besonderheiten der Naturwissenschaften als ein von Menschen betriebenes Unterfangen. Das Erkennen naturwissenschaftlicher Fragen bezieht sich auf Fragestellungen, die sich mit naturwissenschaftlichen Untersuchungen beantworten lassen und die ein Wissen über die Naturwissenschaften und die naturwissenschaftlichen Aspekte bestimmter Themen voraussetzen. Aus Belegen Schlussfolgerungen zu ziehen verlangt, dass die zur Auswahl und Beurteilung von Informationen und Daten notwendigen Prozesse gewusst und angewendet werden. Dabei muss berücksichtigt werden, dass die vorhandenen Informationen häufig nicht ausreichen, um endgültige Schlussfolgerungen zu ziehen, und dass in solchen Fällen vorsichtig und bewusst Vermutungen anhand der verfügbaren Informationen angestellt werden müssen.

... Entscheidungen zu verstehen und zu treffen ...

Die Formulierung *Entscheidungen zu verstehen und zu treffen* weist darauf hin, dass erstens ein Verständnis der natürlichen Welt als Ziel an sich und als Voraussetzung für das Treffen von Entscheidungen wertvoll ist, und dass zweitens naturwissenschaftliches Verständnis zur Entscheidungsfindung zwar beitragen kann, diese aber selten bereits determiniert. Praktische Entscheidungen sind immer in Situationen eingebunden, die auch soziale, politische oder ökonomische Dimensionen haben, und naturwissenschaftliches Wissen wird im Zusammenhang von Wertvorstellungen angewendet, die mit diesen Dimensionen verknüpft sind. Wo über die Werte in einer Situation Einvernehmen besteht, kann der Rückgriff auf naturwissenschaftliche Belege unumstritten sein. Wo die Wertvorstellungen jedoch auseinander gehen, werden die Auswahl und die Verwendung von naturwissenschaftlichen Befunden bei der Entscheidungsfindung eher kontrovers sein.

... die natürliche Welt und die durch menschliches Handeln an ihr vorgenommenen Veränderungen ...

Die Wendung *die natürliche Welt* wird als Kurzform für das physikalische Umfeld, die Lebewesen sowie die Beziehungen zwischen ihnen gebraucht. Entscheidungen über die natürliche Welt schließen Entscheidungen ein, die mit den Naturwissenschaften und ihrer Beziehung zur eigenen Person, zur Familie, zur Gemeinde oder zu globalen Problemen zusammenhängen. Die Wendung *durch menschliches Handeln an ihr vorgenommenen Veränderungen* bezieht sich auf geplante und ungeplante Anpassungen der natürlichen Welt an menschliche Zwecke (einfache und komplexe Technologien) und auf deren Folgen.

An dieser Stelle muß hervorgehoben und später auch noch ausführlicher erklärt werden, dass der Begriff der naturwissenschaftlichen Grundbildung nicht dichotom ist. Das heißt, er ist nicht so zu verstehen, als könnte man Menschen danach einteilen, ob sie über naturwissenschaftliche Grundbildung verfügen oder nicht. Hier muß man sich vielmehr ein Kontinuum vorstellen, das von einer wenig entwickelten bis zu einer hoch entwickelten naturwissenschaftlichen Grundbildung reicht. So können beispielsweise Schülerinnen und Schüler mit einer weniger entwickelten naturwissenschaftlichen Grundbildung durchaus in der Lage sein, manche Belege zu erkennen, die für die Beurteilung einer Behauptung oder für die Untermauerung eines Arguments relevant sind, oder einfache und vertraute Situationen relativ umfassend zu bewerten. Eine höher entwickelte naturwissenschaftliche Grundbildung hingegen zeigt sich in umfassenderen Antworten und in der Fähigkeit, auch in weniger vertrauten und komplexeren Situationen Wissen anzuwenden und Behauptungen unter Rückgriff auf Belege zu beurteilen.

Organisation des Gegenstandsbereichs

Die OECD/PISA-Definition naturwissenschaftlicher Grundbildung umfasst drei Aspekte:

- *naturwissenschaftliche Prozesse*, die, eben weil sie naturwissenschaftlich sind, entsprechendes Wissen voraussetzen, wobei aber das Wissen nicht die wichtigste Voraussetzung für die erfolgreiche Bewältigung der Aufgaben sein soll
- *naturwissenschaftliche Konzepte*, deren Verständnis anhand von Anwendungsaufgaben in bestimmten Inhaltsbereichen gemessen werden soll, und
- *Situationen*, die in den Testaufgaben präsentiert werden (dieser Aspekt wird im allgemeinen Sprachgebrauch häufig als „Kontext“ bezeichnet)

Im Folgenden werden diese Aspekte einzeln behandelt. Es ist jedoch wichtig, an dieser Stelle darauf hinzuweisen, dass Messungen naturwissenschaftlicher Grundbildung immer Kombinationen aus allen drei Aspekten erfassen.

Die ersten beiden Aspekte werden sowohl zur Aufgabenkonstruktion als auch zur Charakterisierung der Schülerleistungen benutzt. Durch Berücksichtigung des dritten Aspekts wird sichergestellt, dass bei der Aufgabenentwicklung die naturwissenschaftlichen Inhalte in ein breites Spektrum relevanter Kontexte eingebettet werden.

In den folgenden Abschnitten werden die drei organisierenden Gesichtspunkte ausführlicher behandelt. Durch ihre Spezifizierung in der OECD/PISA-Rahmenkonzeption wird gewährleistet, dass die Leistungsmessung die Erträge des naturwissenschaftlichen Unterrichts umfassend abdeckt.

Naturwissenschaftliche Prozesse

Prozesse sind mentale (und manchmal physische) Handlungen, die beim Konzipieren, Erheben, Interpretieren und Anwenden von Belegen oder Daten ausgeführt werden, um Wissen oder Verständnis aufzubauen. Prozesse müssen in Bezug auf einen bestimmten Inhalt angewendet werden; es gibt keine inhaltsfreien Prozesse. Das Spektrum der Inhalte, auf die Prozesse angewendet werden können, ist breit; zu *naturwissenschaftlichen Prozessen* werden sie, wenn der Inhalt mit naturwissenschaftlichen Aspekten der Welt zu tun hat und ihre Anwendung zu einer Erweiterung des naturwissenschaftlichen Verständnisses führt.

Die Prozesse, die gewöhnlich als naturwissenschaftlich bezeichnet werden, umfassen ein breites Spektrum von Verständnis und Fähigkeiten, die erforderlich sind, um Belege aus der uns umgebenden Welt zu sammeln, zu interpretieren und Schlussfolgerungen aus ihnen zu ziehen. Zu den Prozessen, die mit dem Sammeln von Belegen verbunden sind, gehören auch forschungspraktische Prozesse wie die Planung und der Aufbau von Versuchsanordnungen, das Durchführen von Messungen, das Sammeln von Beobachtungen mit geeigneten Instrumenten usw. Die Entwicklung solcher Prozesse gehört zu den Zielen des naturwissenschaftlichen Unterrichts. Es soll damit den Schülerinnen und Schülern ermöglicht werden, zu erfahren und zu verstehen, wie naturwissenschaftliches Verständnis aufgebaut wird und, im Idealfall, was die Besonderheiten von naturwissenschaftlicher Forschung und von naturwissenschaftlichem Wissen ausmacht. Nur wenige Schülerinnen und Schüler werden diese praktischen Fertigkeiten im späteren Leben einsetzen. Für alle ist es jedoch wichtig, ein Verständnis von Prozessen und Konzepten durch praktisches Untersuchen und Experimentieren zu entwickeln. Im Übrigen gibt es gute Gründe für die Annahme, dass die traditionelle, im naturwissenschaftlichen Unterricht auch heute noch häufig anzutreffende Auffassung vom „naturwissenschaftlichen Prozess“ als induktives Schließen aus Beobachtungen im Gegensatz dazu steht, wie naturwissenschaftliches Wissen tatsächlich gewonnen wird (z.B. Ziman, 1980).

Für naturwissenschaftliche Grundbildung, wie sie hier definiert wurde, hat die Anwendung von naturwissenschaftlichem Wissen, um „aus Belegen Schlussfolgerungen zu ziehen“, höhere Priorität als die Fähigkeit, selber Belege zu sammeln. Die Fähigkeit, Belege oder Daten zu Behauptungen und Schlussfolgerungen in Beziehung zu setzen, ist wesentlich für alle Menschen, damit sie Urteile über solche Aspekte ihres Lebens fällen können, die von den Naturwissenschaften beeinflusst werden. Daraus folgt, dass alle Menschen wissen sollten, wo naturwissenschaftliches Wissen relevant ist und welche Fragen von den Naturwissenschaften beantwortet werden können und welche nicht. Weiterhin sollten alle in der Lage sein, die Gültigkeit von Belegen zu beurteilen, und zwar in Hinblick auf ihre Relevanz und auf die Art und Weise, wie sie gewonnen wurden. Am Wichtigsten ist jedoch, dass alle imstande sein sollten, Schlussfolgerungen auf die Belege zurückzubeziehen und Belege abzuwägen, die für oder gegen bestimmte, das Leben auf persönlicher, sozialer oder globaler Ebene beeinflussende Handlungen sprechen.

Die soeben vorgenommenen Unterscheidungen lassen sich dahingehend zusammenfassen, dass den Prozessen *über* Naturwissenschaften (Umgang mit naturwissenschaftlichen Erkenntnissen) höhere Bedeutung zugesprochen wird als den Prozessen *in* den Naturwissenschaften (naturwissenschaftliches Arbeiten). Bei den in Abbildung 14 aufgeführten Prozessfähigkeiten geht es daher primär um den Umgang mit den Naturwissenschaften und weniger um ihre Anwendung in den Naturwissenschaften. In allen in Abbildung 14 aufgeführten Prozessen spielt die Kenntnis naturwissenschaftlicher Konzepte eine Rolle. Bei den vier ersten Prozessen sind diese Kenntnisse notwendig, aber nicht hinreichend, da hier Wissen, das die Erhebung und Anwendung von naturwissenschaftlichen Belegen und Daten betrifft, im Vordergrund steht. Bei dem fünften Prozess ist dagegen das Verständnis naturwissenschaftlicher Konzepte selber der wesentliche Faktor.

Abbildung 14. **Ausgewählte naturwissenschaftliche Prozesse**

1. Fragestellungen erkennen, die naturwissenschaftlich untersucht werden können.
2. Belege/Nachweise identifizieren, die in einer naturwissenschaftlichen Untersuchung benötigt werden.
3. Schlussfolgerungen ziehen oder bewerten.
4. gültige Schlussfolgerungen kommunizieren.
5. Verständnis naturwissenschaftlicher Konzepte zeigen.

Diese Prozesse sollen im Folgenden genauer beschrieben werden. Darüber hinaus zeigt Abbildung 19, wie sie operationalisiert werden.

Naturwissenschaftliche Fragestellungen erkennen

Zum Erkennen naturwissenschaftlicher Fragestellungen gehört unter anderem: die Frage oder Idee zu erkennen, die in einer bestimmten Untersuchung geprüft wurde (oder hätte geprüft werden können); zu unterscheiden, welche Fragen durch naturwissenschaftliche Untersuchungen beantwortet werden können und welche nicht; oder eine Frage zu stellen, die in einer bestimmten Situation naturwissenschaftlich untersucht werden könnte.

Naturwissenschaftliche Nachweise identifizieren

Zum Identifizieren von Nachweisen, die für eine naturwissenschaftliche Untersuchung erforderlich sind, gehört es, die Informationen zu bestimmen, die für die gültige Überprüfung einer bestimmten Idee benötigt werden. Dabei kann es zum Beispiel erforderlich sein, zu bestimmen oder zu erkennen, was verglichen werden muss, welche Variablen verändert oder kontrolliert werden müssen, welche zusätzlichen Informationen benötigt werden und was getan werden muss, um relevante Daten zu erheben.

Schlussfolgerungen ziehen oder bewerten

Zum Ziehen von Schlussfolgerungen bzw. zur kritischen Bewertung solcher Schlussfolgerungen gehört unter anderem, auf der Basis naturwissenschaftlicher Belege oder Daten eine Schlussfolgerung zu formulieren bzw. aus alternativen Schlussfolgerungen diejenige auszuwählen, die zu diesen Daten passt; Gründe anzugeben, die unter Berücksichtigung der vorliegenden Daten für oder gegen eine bestimmte Schlussfolgerung sprechen; oder die Annahmen zu identifizieren, die dieser Schlussfolgerung zugrunde liegen.

Gültige Schlussfolgerungen kommunizieren

Zur Mitteilung von validen, auf vorhandene Belege und Daten gestützten Schlussfolgerungen an eine bestimmte Zielgruppe gehört unter anderem, auf der Basis der Situation und der vorliegenden Daten oder auf der Basis von zusätzlichen relevanten Informationen eine Argumentation zu entwickeln, die für das betreffende Publikum angemessen und klar formuliert ist.

Verständnis naturwissenschaftlicher Konzepte zeigen

Zum Nachweis des Verständnisses naturwissenschaftlicher Konzepte durch ihre situationsangemessene Anwendung gehört unter anderem, naturwissenschaftliche Ideen und/oder Informationen, die nicht vorgegeben wurden, heranzuziehen, um Zusammenhänge und mögliche Ursachen bestimmter Veränderungen zu erklären, um Vorhersagen über die Wirkung bestimmter Veränderungen zu treffen oder um Faktoren zu bestimmen, die ein bestimmtes Ergebnis beeinflussen.

Naturwissenschaftliches Wissen wird bei allen fünf Prozessen benötigt. Bei den ersten vier sollen diese Kenntnisse jedoch nicht die entscheidende Voraussetzung darstellen, da hier die mentalen Prozesse untersucht werden sollen, die bei der Erhebung, Beurteilung und Kommunikation von gültigen naturwissenschaftlichen Belegen ausschlaggebend sind. Bei dem fünften Prozess dagegen wird das Verständnis naturwissenschaftlicher Konzepte gemessen und stellt die eigentliche „Hürde“ dar.

Zu beachten ist, dass der Schwierigkeitsgrad der Aufgaben, die zu jedem der oben aufgeführten Prozesse gestellt werden, breit variieren kann und von den jeweiligen naturwissenschaftlichen Konzepten wie auch von den Situationen abhängt. OECD/PISA wird mit Hilfe von Rückmeldungen aus den Teilnehmerländern und den Ergebnissen des Feldtests sicher stellen, dass die für die Haupterhebung ausgewählten Items einen für 15-Jährige angemessenen Schwierigkeitsgrad aufweisen.

Konzept und Inhalt

Konzepte versetzen uns in die Lage, neuen Erfahrungen Bedeutung zu geben, indem wir sie mit dem verbinden, was wir bereits wissen. *Naturwissenschaftliche Konzepte* helfen uns, bestimmte Aspekte der natürlichen und der vom Menschen geschaffenen Welt zu verstehen. Naturwissenschaftliche Konzepte werden auf vielen verschiedenen Ebenen formuliert – von den sehr umfassenden Bezeichnungen für die Naturwissenschaft, unter denen sie in den Schulen präsentiert werden (Biologie, Physik, Geowissenschaften usw.), bis zu den langen Listen von Verallgemeinerungen, wie sie häufig in Beschreibungen von curricularen Anforderungen oder in Lehrplänen auftauchen.

Es gibt viele Möglichkeiten, naturwissenschaftliche Konzepte zu gruppieren, um das Verständnis der naturwissenschaftlichen Aspekte der Welt zu erleichtern. Manchmal sind Konzepte Bezeichnungen, die charakteristische Merkmalsbereiche für eine bestimmte Gruppe von Objekten oder Ereignissen beinhalten (z.B. „Säugetiere“, „Beschleunigung“, „Lösungsmittel“); von solchen Konzepten dürfte es mehrere Tausend geben. Konzepte können aber auch als Verallgemeinerungen bestimmter Phänomene formuliert werden (z.B. die „Gesetze“ oder Theoreme der Physik oder Chemie); von diesen gibt es viele Hunderte. Diese können außerdem als wichtige naturwissenschaftliche Themen gefaßt werden, die allgemein anwendbar sind und sich für Mess- und Berichtszwecke leichter operationalisieren lassen.

Im Rahmen von OECD/PISA wird die Auswahl der naturwissenschaftlichen Konzepte, die erfasst werden sollen, durch vier Kriterien bestimmt:

- Das erste Kriterium ist die Relevanz für alltägliche Situationen. Naturwissenschaftliche Konzepte unterscheiden sich im Grad ihrer Nützlichkeit für das tägliche Leben. Die Relativitätstheorie zum Beispiel bietet zwar eine genauere Beschreibung der Beziehungen zwischen Länge, Masse, Zeit und Geschwindigkeit, doch sind die Newtonschen Gesetze nützlicher, wenn es um die Kräfte und Bewegungen geht, mit denen wir es im Alltag zu tun haben.
- Das zweite Kriterium ist, dass die ausgewählten Konzepte und Inhalte bleibende Bedeutung für das Leben im nächsten Jahrzehnt und darüber hinaus haben sollten. Da naturwissenschaftliche Grundbildung erst im Jahr 2006 die Hauptkomponente der Erhebung sein wird, werden im ersten Zyklus von OECD/PISA solche Konzepte im Mittelpunkt stehen, die mit großer Wahrscheinlichkeit auch dann noch eine Bedeutung in den Naturwissenschaften und in der Politik haben werden.
- Die dritte Grundlage für die Auswahl ist die Relevanz der Konzepte für Situationen, die für den Nachweis naturwissenschaftlicher Grundbildung geeignet sind.
- Das vierte Kriterium ist, dass die Konzepte eine Verknüpfung mit ausgewählten naturwissenschaftlichen Prozessen gestatten. Dies wäre nicht der Fall, wenn es nur darum ginge, Bezeichnungen oder Definitionen aus dem Gedächtnis zu reproduzieren.

Abbildungen 15 und 16 zeigen das Ergebnis der Anwendung dieser Kriterien auf naturwissenschaftliche Konzepte und Inhalte. Abbildung 15 führt wichtige naturwissenschaftliche Themen zusammen mit einigen Beispielen für darauf bezogene Konzepte auf. Diese allgemeinen Konzepte sind Voraussetzung für ein Verständnis der natürlichen Welt und für das Erfassen der Bedeutung neuer Erfahrungen. Sie hängen von der Untersuchung spezifischer Phänomene und Ereignisse ab und werden aus ihnen abgeleitet, gehen aber über das hierbei gewonnene Detailwissen hinaus. Die in Abbildung 15 aufgeführten Konzepte sind Beispiele, die die Bedeutung der Themen veranschaulichen sollen; sie stellen keine vollständige Liste aller mit diesen Themen verbundenen Konzepte dar.

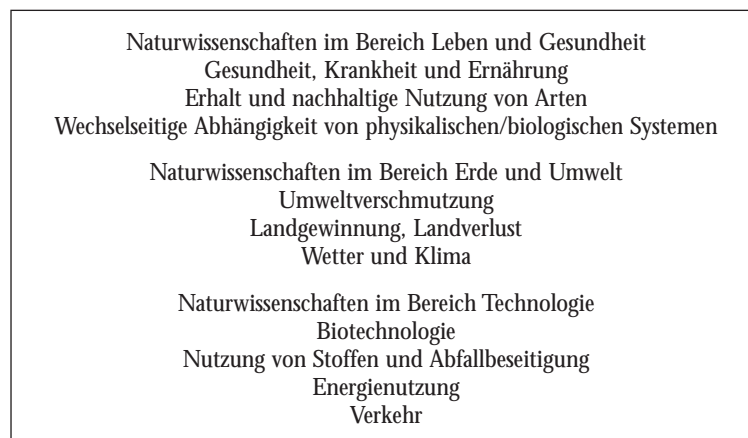
An den in Abbildung 15 als Beispiele aufgeführten Konzepten ist zu erkennen, dass sich das zu messende Wissen auf die großen Hauptbereiche der Naturwissenschaften bezieht: Physik, Chemie, Biologie, Geowissenschaften. Die Items werden nach diesen Hauptbereichen sowie nach den Themen, Anwendungsbereichen und Prozessen klassifiziert, die mit ihnen untersucht werden sollen.

Abbildung 15. **Wichtige naturwissenschaftliche Themen (mit Beispielen für zugeordnete Konzepte) für die Messung naturwissenschaftlicher Grundbildung**



Abbildung 16 zeigt diejenigen Anwendungsbereiche der Naturwissenschaften, in denen sich Fragen stellen, die Bürger von heute und morgen verstehen und zu denen sie Entscheidungen treffen müssen. An diesen Anwendungen orientiert sich die Auswahl der Aufgabeninhalte und Items. Abbildung 16 gibt also die Anwendungsbereiche an, in denen das Verständnis der in Abbildung 15 aufgeführten Konzepte gemessen wird.

Abbildung 16. **Anwendungsbereiche der Naturwissenschaften für die Messung naturwissenschaftlicher Grundbildung**



OECD/PISA wird, wie oben bereits erwähnt, wichtige Konzepte einbeziehen, die für die naturwissenschaftlichen Curricula der Teilnehmerländer relevant sind, ohne sich dabei jedoch auf einen allen nationalen Curricula gemeinsamen Nenner beschränken zu müssen. Da in der Studie der Schwerpunkt auf naturwissenschaftlicher Grundbildung (literacy) liegt, wird dies dadurch realisiert, dass ausgewählte naturwissenschaftliche Konzepte und Prozesse in wichtigen Situationen angewendet werden müssen, die die reale Welt widerspiegeln und mit naturwissenschaftlichen Ideen zu tun haben.

Situationen

Neben Prozessen und Konzepten ist das dritte Merkmal von Testaufgaben, das die Leistung beeinflusst, die Situation, in der die Probleme eingebettet sind. Dies wird häufig als der *Kontext* oder *Rahmen* von Aufgaben bezeichnet. Hier wird jedoch das Wort „*Situation*“ benutzt, um Verwechslungen mit anderen Verwendungen dieser Begriffe zu vermeiden. Da die jeweils gewählte Situation die Leistung beeinflusst, ist es wichtig, das Spektrum der Situationen für die Testaufgaben festzulegen und zu kontrollieren. Zwar ist nicht vorgesehen, Leistungen getrennt nach Situationen zu berichten. Die Situationen müssen aber dennoch genau bestimmt werden, um sicherzustellen, dass sich die Aufgaben angemessen auf die als wichtig erachteten Situationen verteilen, und dass diese Situationen, soweit es sich aufgrund der Feldtestergebnisse als notwendig erweist, über die Erhebungszyklen hinweg unter dem Gesichtspunkt ihrer internationalen Vergleichbarkeit kontrolliert werden können.

Bei der Auswahl der Situationen ist zu beachten, dass der Zweck der OECD/PISA-Erhebung im Bereich naturwissenschaftlicher Grundbildung darin besteht, die Fähigkeit von Schülern zu messen, Kompetenzen und Kenntnisse, die sie gegen Ende der Pflichtschulzeit erworben haben, auch anzuwenden. Eine Voraussetzung ist dabei, dass die Aufgaben Situationen aus dem Leben allgemein und nicht nur aus dem Leben in der Schule widerspiegeln. In der Schulsituation sind die naturwissenschaftlichen Prozesse und Konzepte oft auf das Labor oder das Klassenzimmer beschränkt, doch wird in den naturwissenschaftlichen Curricula vieler Länder zunehmend versucht, Konzepte und Prozesse anwendungsorientiert zu behandeln, also eingebettet in Zusammenhänge der außerschulischen Welt.

Realitätsnahe Situationen beinhalten Probleme, die uns als Individuen betreffen können (etwa Nahrungsmittel- und Energieverbrauch), als Mitglieder einer lokalen Gemeinschaft (etwa Trinkwasseraufbereitung oder Standortsuche für ein Kraftwerk) oder als Bürger der Welt (etwa globale Erwärmung, Abnahme der Artenvielfalt). All diese Situationen sind im Spektrum der Testaufgaben von OECD/PISA vertreten. Für bestimmte Themen geeignet sind weiterhin historische Situationen, mit deren Hilfe das Verständnis von Fortschritten naturwissenschaftlichen Wissens untersucht werden kann. Im Rahmen von OECD/PISA betreffen die Items also Angelegenheiten, die sich auf die eigene Person und Familie (persönlich), auf die Gemeinschaft (öffentlich) oder auf die ganze Welt (global) beziehen. Weiterhin werden die Entwicklung naturwissenschaftlichen Wissens und der Einfluss dieses Wissens auf gesellschaftliche Entscheidungen angesprochen (historische Relevanz).

Im Rahmen einer internationalen Studie ist es wichtig, die für die Items verwendeten Situationen unter Berücksichtigung des Lebens und der Interessen von Schülern in allen Teilnehmerländern auszuwählen. Außerdem sollten sie zur Messung von naturwissenschaftlichen Prozessen und Konzepten geeignet sein. Bei der Entwicklung und Auswahl der Aufgaben hat die Sensibilität für kulturelle Unterschiede nicht nur wegen der Validität der Messung hohe Priorität, sondern auch weil Rücksicht auf die unterschiedlichen Wertvorstellungen und Traditionen der Teilnehmerländer genommen werden muss. Mit Hilfe der Ergebnisse aus den Feldtests soll sichergestellt werden, dass die für die Testaufgaben ausgewählten Situationen sowohl für alle Teilnehmerländer relevant und geeignet sind als auch eine Verbindung von naturwissenschaftlichem Wissen und Anwendung naturwissenschaftlicher Prozesse beinhalten.

Mit der Einbettung der Items in diese Art von Situationen verfolgt OECD/PISA das Ziel, die Anwendung von Wissen zu messen, das die Schülerinnen und Schüler im Rahmen ihres naturwissenschaftlichen Unterrichts erworben haben dürften (auch wenn bestimmte Kenntnisse aus anderen Fächern oder außerschulischen Quellen stammen können). Obwohl jedoch das Wissen, das dabei gefordert ist, zum Lehrstoff gehört, wird es im Rahmen von OECD/PISA als angewandtes, auf realitätsnahe Situationen übertragenes Wissen gemessen. Auf diese Weise soll herausgefunden werden, ob der Wissenserwerb über ein isoliertes Faktenwissen hinausgegangen ist und zur Entwicklung von naturwissenschaftlicher Grundbildung (scientific literacy) geführt hat. Dies wird später anhand von Beispielitems verdeutlicht werden.

Aufgabenmerkmale

Gemäß der OECD/PISA-Definition naturwissenschaftlicher Grundbildung sind für die Bearbeitung jeder Testaufgabe die Anwendung eines oder mehrerer der in Abbildung 14 aufgeführten Prozesse und, wie ebenfalls bereits dargestellt, ein bestimmtes Maß an naturwissenschaftlichem Wissen erforderlich. Die Aufgaben bestehen aus einer Reihe von Fragen (Items), die sich auf ein bestimmtes, die Situation definierendes Stimulusmaterial beziehen. Manche Aufgaben können neben Items, die naturwissenschaftliche Grundbildung erfassen, auch Items zur Messung von Lesekompetenz und mathematischer Grundbildung enthalten.

Beispiele für Items zur Erfassung von naturwissenschaftlichen Prozessen

Die folgenden Items, die sich zur Erfassung einiger der Prozesse eignen könnten, sollen die operationale Bedeutung dieser Prozesse verdeutlichen. Die beiden ersten Prozesse werden mit zwei Fragen in der Aufgabe „Stopp diese Bazille!“ gemessen. In dieser Aufgabe lesen die Schülerinnen und Schüler einen kurzen Text, der einen Auszug über die Geschichte der Impfung enthält. Der Auszug und zwei Beispielfragen sind in Abbildung 17 dargestellt.

Abbildung 17. **Naturwissenschaftliche Grundbildung, Beispiel 1**

Bereits im 11. Jahrhundert haben chinesische Ärzte das Immunsystem manipuliert. Sie bliesen pulverisierten Schorf eines Windpockenkranken in die Nasenlöcher ihrer Patienten und konnten so häufig eine milde Form von Windpocken herbeiführen, die eine spätere, schwerere Erkrankung verhinderte. Anfang des 18. Jahrhunderts rieb man sich mit getrocknetem Schorf ein, um sich vor der Krankheit zu schützen. Diese primitiven Praktiken wurden in England und in den amerikanischen Kolonien eingeführt. 1771 und 1772 ritzte Zabdiel Boylston, ein Arzt in Boston, bei seinem sechsjährigen Sohn sowie bei 285 anderen Leuten die Haut ein und rieb Eiter von Windpockenschorf in die Wunde. Bis auf sechs Patienten überlebten alle.

Beispielitem 1: Welche Idee könnte Zabdiel Boylston getestet haben?

Beispielitem 2: Nenne zwei weitere Informationen, die du brauchen würdest, um zu entscheiden, wie erfolgreich Boylstons Ansatz war.

Beispielitem 1 erfordert eine frei formulierte Antwort, die je nach der Anzahl der darin enthaltenen relevanten Details die Punktwerte 2, 1 oder 0 erhalten kann. (Die folgende Antwort würde zum Beispiel zwei Punkte bekommen: „Durch Einritzen der Haut und Einbringen von Eiter direkt in die Blutbahn werden die Chancen erhöht, dass sich Abwehrstoffe gegen Windpocken bilden.“) Das Item erfasst Prozess 1 – *naturwissenschaftliche Fragestellungen erkennen* –, wobei Kenntnisse aus der *Humanbiologie* im Bereich *Leben und Gesundheit* angewendet werden müssen.

Beispielitem 2 kann ebenfalls die Punktwerte 2, 1 oder 0 erhalten, je nachdem, wie viele der relevanten Informationen genannt werden (die Überlebensrate ohne Boylstons Behandlung und die Frage, ob seine Patienten außerhalb dieser Behandlung mit Windpocken zu tun hatten). Das Item erfasst Prozess 2 – *Naturwissenschaftliche Nachweise identifizieren* – und erfordert Kenntnisse der *Humanbiologie* und ihrer Anwendung im Bereich *Leben und Gesundheit*.

Die vier Items in Abbildung 18 gehören zu einer Aufgabe, deren Stimulusmaterial ein Text über Peter Cairney ist, der für den *Australian Road Research Board* arbeitet.

Beispielitem 3 erfasst Prozess 2 – *Naturwissenschaftliche Nachweise identifizieren* –, wobei Kenntnisse über *Kräfte und Bewegung* im Bereich *Technologie* angewendet werden müssen.

Beispielitem 4 erfasst Prozess 3 – *Schlussfolgerungen ziehen oder bewerten* –, wobei ebenfalls Kenntnisse über *Kräfte und Bewegung* im Bereich *Technologie* angewendet werden müssen. Für die Zustimmung oder Ablehnung an sich werden keine Punkte vergeben, sondern für die angegebenen Gründe dieser Zustimmung oder Ablehnung, die mit der

gegebenen Information in Einklang stehen müssen (zum Beispiel: Zustimmung, weil die Gefahr von Zusammenstößen geringer ist, wenn der Verkehr nah an den Rändern der Strasse bleibt, auch wenn er insgesamt schneller wird; wird er schneller, besteht weniger Anreiz zum Überholen. Oder: Ablehnung, weil bei schneller werdendem Verkehr und gleich bleibendem Abstand in Notfällen möglicherweise nicht genug Platz für Brems- und Ausweichmanöver bleibt).

Beispielitem 5 erfasst Prozess 5 – *Verständnis naturwissenschaftlicher Konzepte zeigen* – mit Fragen zu *Kräften und Bewegung* im Bereich *Technologie*. Es erfordert eine frei formulierte Antwort, die die Punktwerte 2, 1 oder 0 erhalten kann, je nach dem, ob einer oder beide der signifikanten Punkte genannt werden: a) ein Fahrzeug, das schneller fährt, hat eine größere Schubkraft, und infolgedessen ist mehr Kraft erforderlich, um es zum Stehen zu bringen; b) ein mit höherer Geschwindigkeit fahrendes Fahrzeug, das die Geschwindigkeit verringert, legt in der gleichen Zeit eine weitere Strecke zurück als ein langsames Fahrzeug.

Beispielitem 6 erfasst ebenfalls Prozess 5 – *Verständnis naturwissenschaftlicher Konzepte zeigen* – mit Fragen zu *Kräften und Bewegung* im Bereich *Technologie*. Es ist ein einfaches Multiple-Choice-Item mit nur einer richtigen Antwort (*Alternative b*), die mit einem Punkt bewertet wird.

Abbildung 18. Naturwissenschaftliches Verständnis, Beispiel 2

... Peter sammelt Informationen auch noch auf andere Art, nämlich mit einer Fernsehkamera, die an einer 13 Meter langen Stange angebracht ist und mit der er den Verkehr auf einer schmalen Strasse filmt. Den Bildern können die Forscher zum Beispiel entnehmen, wie schnell der Verkehr ist, wieviel Abstand die Fahrzeuge voneinander halten und welchen Teil der Strasse die Fahrzeuge nutzen. Nach einer gewissen Zeit werden dann Linien auf die Strasse gemalt, um Fahrbahnen zu markieren. Dann können die Forscher mit Hilfe der Fernsehkameras sehen, ob der Verkehr jetzt anders verläuft. Ist der Verkehr schneller oder langsamer geworden? Sind die Abstände zwischen den Fahrzeugen größer oder kleiner als vorher? Fahren die Autofahrer jetzt, wo es Linien gibt, dichter am Rand oder eher in der Mitte der Strasse? Nachdem Peter diese Dinge weiß, kann er Ratschläge dazu geben, ob auf schmalen Strassen Fahrbahnen mit Linien markiert werden sollten oder nicht.

Beispielitem 3: Um sicherzugehen, dass er gute Ratschläge gibt, könnte Peter, außer die schmale Strasse zu filmen, auch noch andere Informationen sammeln. Welche der folgenden Informationen würden ihm helfen, noch sicherer zu sein, dass seine Ratschläge zu den Auswirkungen von Fahrbahnmarkierungen auf schmalen Strassen richtig sind?

- | | |
|---|---------|
| a) Die gleiche Information für andere schmale Strassen sammeln. | Ja/Nein |
| b) Die gleiche Information für breite Strassen sammeln. | Ja/Nein |
| c) Die Unfallzahlen vor und nach Anbringen der Fahrbahnmarkierungen prüfen. | Ja/Nein |
| d) Die Zahl der Autos prüfen, die vor und nach Anbringen der Fahrbahnmarkierungen die Strasse benutzen. | Ja/Nein |

Bewertung: Ja bei a) und c) Nein bei b) und d) (2 Punkte)
 Ja bei a) Nein bei b), c) und d) (1 Punkt)
 Alle anderen Kombinationen (0 Punkte)

Beispielitem 4: Nimm an, Peter findet heraus, dass sich der Verkehr in einem bestimmten Abschnitt der schmalen Strasse ändert, wie in der folgenden Tabelle angegeben.

Geschwindigkeit	Verkehr ist schneller geworden
Position	Verkehr hält sich dichter an den Rändern der Strasse
Abstand	Unverändert

Aufgrund dieser Ergebnisse wurde beschlossen, dass auf allen schmalen Strassen die Fahrbahnen durch Begrenzungslinien markiert werden sollten.

Meinst du, dass dies die beste Entscheidung war?

Nenne deine Gründe, warum du zustimmst oder nicht zustimmst.

Ich stimme zu

Ich stimme nicht zu

Begründung: _____

Beispielitem 5: Autofahrern wird geraten, wenn sie schnell fahren, grösseren Abstand zu dem Fahrzeug vor ihnen zu halten, als wenn sie langsam fahren, weil schnellere Fahrzeuge länger brauchen, um zum Stehen zu kommen.

Erkläre, warum ein schnelleres Fahrzeug länger braucht, um zum Stehen zu kommen, als ein langsames.

Begründung: _____

Beispielitem 6: Peter sieht auf seinem Fernsehschirm, wie ein mit 45 km/Std. fahrendes Auto (A) von einem anderen Auto (B) überholt wird, das 60 km/Std. fährt. Aus der Sicht einer Person, die in Auto (A) sitzt, wie schnell scheint Auto (B) zu fahren?

- a) 0 km/Std.
- b) 15 km/Std.
- c) 45 km/Std.
- d) 60 km/Std.
- e) 105 km/Std.

Zur Beantwortung all dieser Fragen müssen die Schülerinnen und Schüler Wissen aktivieren, das sie im naturwissenschaftlichen Unterricht erworben haben, und dieses Wissen in einer neuen Situation anwenden. Bei Items, die nicht in erster Linie konzeptuelles Verständnis messen sollen, stellt das erforderliche Wissen nicht die Hauptschwierigkeit (Haupttürde) dar, und der Erfolg sollte von der Fähigkeit im Umgang mit dem Prozess abhängen, um den es geht. Ist jedoch der Hauptzweck des Items, wie in den Beispielen 5 und 6, die Messung konzeptuellen Verständnisses, besteht der Prozess im Nachweisen dieses Verständnisses.

Erhebungsstruktur

Wie die Beispiele zeigen, wird eine Aufgabe in der Regel aus mehreren, an ein und dasselbe Stimulusmaterial anknüpfenden Items bestehen. Dabei können die Items innerhalb einer Aufgabe verschiedene Prozesse oder naturwissenschaftliche Konzepte messen, wobei jedoch jedes Item immer einen der in Abbildung 14 aufgeführten naturwissenschaftlichen Prozesse erfasst.

Diese Struktur dient unter anderem dazu, die Aufgaben so realistisch wie möglich zu machen und die Komplexität realitätsnaher Situationen widerzuspiegeln. Ein weiterer Grund für die Wahl der Struktur ist die effizientere Nutzung der Testzeit: Die Zeit, die die Schülerinnen und Schüler zum „Einstieg“ in das mit einer neuen Situation angesprochene Thema benötigen, verkürzt sich, wenn zu jeder Situation nicht nur eine, sondern mehrere Fragen gestellt werden. Bei der Entwicklung der OECD/PISA-Aufgaben wird darauf geachtet, dass jeder einzelne Punkt, der innerhalb einer solchen Aufgabe erreicht werden kann, von allen anderen innerhalb der Aufgabe zu erzielenden Punkten unabhängig ist. Wichtig ist außerdem, Verzerrungen zu minimieren, die durch Verwendung einer geringeren Anzahl verschiedener Situationen entstehen können.

Alle Aufgaben werden aus mehreren – maximal acht – Items bestehen, die jeweils unabhängig zu bewerten sind. Fast alle, vielleicht sogar alle Aufgaben enthalten sowohl Items, bei denen es um die Kenntnis und das Verständnis eines bestimmten Konzepts geht (siehe Beispiele 5 und 6), als auch Items, bei denen einer oder mehrere Prozesse angewendet werden müssen, die mit der Erhebung und Verwendung von Belegen und Daten in einer naturwissenschaftlichen Untersuchung verbunden sind (siehe Beispiele 1 bis 4). OECD/PISA wird, wie bereits erwähnt, keine praktischen Aufgaben enthalten, zumindest nicht in den Jahren 2000 und 2003, in denen naturwissenschaftliche Grundbildung eine Nebenkategorie ist.

Tabelle 10 zeigt das für den Test angestrebte Gleichgewicht zwischen den Prozessen in Form ihrer prozentualen Anteile an den Gesamtwerten. Für die OECD/PISA-Erhebung im Jahr 2006, in der die Naturwissenschaften den Schwerpunkt bilden werden, kann diese Verteilung noch revidiert werden.

Tabelle 10. **Empfohlene Verteilung der zu erzielenden Punkte auf die naturwissenschaftlichen Prozesse**

Naturwissenschaftliche Prozesse	% der zu erzielenden Punkte
Naturwissenschaftliche Fragestellungen erkennen	10-15
Naturwissenschaftliche Nachweise identifizieren	15-20
Schlussfolgerungen ziehen oder bewerten	15-20
Schlussfolgerungen kommunizieren	10-15
Verständnis naturwissenschaftlicher Konzepte zeigen	40-50

Es kann passieren, dass durch die Inhalte in manchen Aufgaben eine Verlagerung des Schwerpunktes auf die Messung konzeptuellen Verständnisses erfolgt (Prozess 5). Bei anderen Aufgaben wird jedoch dafür eine Verschiebung in die entgegengesetzte Richtung erfolgen. Soweit möglich, werden Items zur Messung der Prozesse 1 bis 4 und Items zur Messung von Prozess 5 in allen Aufgaben vorkommen, um erstens das Ziel zu erreichen, dass alle wichtigen naturwissenschaftlichen Konzepte, die die Schülerinnen und Schüler im naturwissenschaftlichen Unterricht oder auch außerhalb der Schule erworben haben dürften, erfasst werden, und um zweitens zu berücksichtigen, daß die Fähigkeit, Prozesse anzuwenden, sehr stark von der Situation abhängt, in der diese vorkommen (so dass die Prozesse in Bezug auf eine ganze Reihe verschiedener Konzepte gemessen werden müssen). Wie in den

Zielen von OECD/PISA definiert, sind sowohl das konzeptuelle Verständnis als auch die Verbindung naturwissenschaftlichen Wissens mit der Fähigkeit, anhand vorliegender Informationen Schlussfolgerungen zu ziehen, wertvolle Lernergebnisse. Die angestrebte, ungefähr gleichmäßige Verteilung der erreichbaren Punkte auf diese beiden wichtigen Lernergebnisse soll diesen Zielen dienen.

Es wurde bereits darauf hingewiesen, dass es bei allen Itemtypen um die Anwendung von naturwissenschaftlichen Konzepten geht, die die Schülerinnen und Schüler als Teil des naturwissenschaftlichen Lehrstoffs in der Schule erworben haben dürften. Die OECD/PISA-Items unterscheiden sich von manchen – aber keineswegs allen – in der Schule benutzten naturwissenschaftlichen Tests insofern, als diese Konzepte in realitätsnahen Situationen angewendet werden müssen. Auch die Fähigkeit, aus Belegen Schlussfolgerungen zu ziehen, gehört zu den Zielen vieler naturwissenschaftlichen Curricula. Im OECD/PISA-Test sollen die Prozesse aber in Kontexten angewendet werden, die über Situationen innerhalb des Schullabors oder Klassenzimmers hinausgehen. Inwieweit dies neu für die Schülerinnen und Schüler ist, wird davon abhängen, in welchem Umfang realitätsnahe Anwendungen bereits Teil ihres Curriculums sind.

Tabelle 11 zeigt, dass die Verteilung der im Test zu erzielenden Punkte auf die drei Hauptgruppen der Anwendungsbereiche so gleichmäßig wie möglich sein wird.

Tabelle 11. Empfohlene Verteilung der zu erzielenden Punkte auf die Anwendungsbereiche

Anwendungsbereiche der Naturwissenschaften	% der zu erzielenden Punkte
Naturwissenschaften im Bereich Leben und Gesundheit	30-40
Naturwissenschaften im Bereich Erde und Umwelt	30-40
Naturwissenschaften im Bereich Technologie	30-40

Ebenso wird im Rahmen von OECD/PISA eine gleichmäßige Verteilung der Items auf die vier genannten Situationen angestrebt: persönlich, gemeinschaftsbezogen, global und historisch.

Die in den Aufgaben enthaltenen Situationen werden mit Hilfe von Stimulusmaterial in Gestalt eines kurzen Textes oder der zu einer Tabelle, einem Diagramm oder einem Graphen gehörenden Beschriftung definiert. Die Items werden aus einem Komplex einzeln zu bewertender Fragen bestehen, die die in Abbildung 19 aufgeführten Arten von Antworten erfordern. Zu beachten ist, dass diese auch die Fähigkeit umfassen, Verständnis naturwissenschaftlicher Konzepte durch ihre Anwendung in den Anwendungsbereichen nachzuweisen. Die erforderlichen Antworten beziehen sich also auf die im Stimulus vorgegebenen Situationen und Anwendungsbereiche.

Bevor die Vortests durchgeführt und ihre Ergebnisse analysiert worden sind, können keine verbindlichen Aussagen darüber gemacht werden, wie einheitlich die Aufgaben zum Beispiel in Bezug auf die Zahl der in ihnen enthaltenen Items sein sollten, ob die Items sich auf mehr als einen Anwendungsbereich beziehen sollten oder wie Items mit verschiedenen Formaten angeordnet werden sollten. Zur Zeit ist Folgendes geplant:

- Alle Aufgaben werden aus mehreren Items bestehen. Sie werden Items enthalten, mit denen ein oder mehrere naturwissenschaftliche Prozesse (Abb. 14), Kenntnisse naturwissenschaftlicher Konzepte (Abb. 15) und Kenntnisse in Bezug auf einen oder mehrere Anwendungsbereiche der Naturwissenschaften (Abb. 16) erfasst werden und die mit Papier und Bleistift zu beantworten sind (schriftlich oder als Zeichnung).
- Die meisten Aufgaben werden in schriftlicher Form gestellt, zumindest im Rahmen der Erhebung im Jahr 2000. Für die Erhebung im Jahr 2006, in der die Naturwissenschaften den Hauptbereich bilden werden, wird auch der Einsatz von Stimulusmaterial in anderen Formen erwogen.
- Einige Aufgaben werden Items enthalten, bei denen es auch um Lesekompetenz und/oder mathematische Grundbildung geht und die einen Beitrag zu den Messungen in diesen Bereichen leisten sollen. Es wird jedoch im Bereich naturwissenschaftlicher Grundbildung weder Items geben, bei denen lediglich Informationen aus dem Stimulusmaterial wiederzugeben sind, noch Items, für die isoliertes Faktenwissen aus dem Gedächtnis reproduziert werden muss.

Zur Erfassung des ganzen Spektrums der Fähigkeiten und des Verständnisses, wie sie in dieser Rahmenkonzeption definiert wurden, sind verschiedene Antwortformate erforderlich. So können zum Beispiel Multiple-Choice-Items für eine valide Messung von Prozessen entwickelt werden, bei denen es um Wiedererkennen oder Auswählen geht. Zur Messung der Beurteilungs- und Kommunikationsfähigkeit dürften valide und authentische Ergebnisse jedoch eher mit frei formulierten Antworten zu erzielen sein. In vielen Fällen wird es aber vom Inhalt des Items abhängen, welches Format am besten geeignet ist.

Abbildung 19. **Itemtypen für die Messung naturwissenschaftlicher Grundbildung**

Vorgabe	Verlangte Antwort
Naturwissenschaftliche Fragestellungen erkennen	
Darstellung einer Untersuchung oder eines Verfahrens, bei der oder dem Daten erhoben oder Vergleiche durchgeführt wurden.	Die Frage oder Idee auswählen oder formulieren, die getestet wurde (oder hätte getestet werden können).
Beschreibung einer Situation, in der Fragen naturwissenschaftlich untersucht werden könnten.	Eine Frage formulieren, die naturwissenschaftlich untersucht werden könnte.
Mehrere Fragen/Ideen, die sich aus der vorgegebenen Situation ergeben oder für sie relevant sind.	Diejenige(n) Frage(n)/Idee(n) auswählen, die durch naturwissenschaftliche Untersuchungen beantwortet/getestet werden könnte(n).
Naturwissenschaftliche Nachweise identifizieren	
Eine Idee oder Hypothese, die in der Frage oder in dem zu testenden Stimulusmaterial dargelegt wird.	Informationen auswählen oder anführen, die benötigt werden, um die Idee zu testen oder eine auf ihr beruhende Vorhersage zu machen. Diese Informationen können folgende Fragen betreffen: a) Welche Dinge verglichen werden sollten. b) Welche Variablen verändert oder kontrolliert werden sollten. c) Welche zusätzlichen Informationen benötigt werden. d) Welche Schritte unternommen werden sollten, damit relevante Daten erhoben werden können.
Schlussfolgerungen ziehen oder bewerten	
Daten (Testergebnisse, Beobachtungen), aus denen Schlussfolgerungen gezogen werden können.	Schlussfolgerungen formulieren, die zu den Daten passen.
Daten (Testergebnisse, Beobachtungen) und Schlussfolgerungen, die aus ihnen gezogen wurden.	Diejenige Schlussfolgerung auswählen, die zu den Daten passt, und eine Erklärung geben.
Daten (Testergebnisse, Beobachtungen) und eine Schlussfolgerung, die aus ihnen gezogen wurde.	Gründe dafür angeben, warum die vorliegenden Daten die Schlussfolgerung stützen oder nicht stützen, oder beurteilen, inwieweit die Schlussfolgerung verlässlich ist.
Schlussfolgerungen kommunizieren	
Eine Situation, aus der (verschiedene) Schlussfolgerungen gezogen werden können oder in der Informationen zusammen gezogen werden müssen, um eine Schlussfolgerung zu stützen, sowie Bestimmung eines Publikums.	Eine Argumentation entwickeln, die für das gegebene Publikum klar formuliert ist und sich auf relevante Belege/Daten aus dem Stimulusmaterial stützt.
Verständnis naturwissenschaftlicher Konzepte zeigen	
Eine Situation, in der eine Vorhersage, Erklärung oder Information verlangt wird.	Eine Vorhersage oder Erklärung oder Zusatzinformation formulieren oder auswählen, die auf dem Verständnis eines naturwissenschaftlichen Konzepts oder auf Information beruht, die nicht in der Frage oder dem Stimulusmaterial enthalten ist.

Skalentwicklung

Skalen und Subskalen

Die Entwicklung von Leistungsskalen für Schülerinnen und Schüler ist eine wesentliche Voraussetzung für die Verwirklichung der Ziele von OECD/PISA. Der Prozess der Skalentwicklung ist insofern iterativ, als erste Entwürfe, die auf bereits vorliegenden Erfahrungen und Untersuchungen im Bereich Lernen und kognitive Entwicklung in den Naturwissenschaften basieren, auf Grundlage der beim Feldtest von OECD/PISA gewonnenen Daten weiterentwickelt werden sollen.

Bereits vorhandene Forschungsergebnisse und frühere Erfahrungen deuten darauf hin, dass eine Skala identifiziert werden kann, die die Entwicklung naturwissenschaftlicher Grundbildung abbildet. Diese Entwicklung reicht

von der Fähigkeit, relativ leicht begreifbare naturwissenschaftliche Konzepte zu benutzen und in vertrauten Situationen Folgendes zu tun:

- Fragen zu erkennen, die durch naturwissenschaftliche Untersuchungen entschieden bzw. nicht entschieden werden können;
- Informationen zu identifizieren, die vorliegen müssen, damit in Situationen, in denen eine Variable verändert und eine andere kontrolliert werden muss, eine Behauptung geprüft oder ein Problem untersucht werden kann;
- zu erklären, warum sich Schlussfolgerungen oder Behauptungen in Situationen nicht aufrecht erhalten lassen, in denen eine Variable kontrolliert werden müsste, aber nicht kontrolliert werden kann;
- die wichtigsten Verknüpfungspunkte zwischen Schlussfolgerungen und Belegen so darzustellen, dass sie von anderen verstanden werden können;
- mit Hilfe von relativ leicht zu begreifenden Konzepten Vorhersagen zu machen und Erklärungen zu geben;

bis zur Fähigkeit, kognitiv anspruchsvollere Konzepte anzuwenden und in komplexeren Situationen Folgendes zu tun:

- zu verstehen, dass jedes naturwissenschaftliche Verständnis vorläufig ist und dass das Testen von Theorien zu einer Revision und Verbesserung dieses Verständnisses führen kann;
- zu bestimmen, welche Informationen gesammelt werden müssen und unter welchen Bedingungen sie gesammelt werden sollten, um auf ihrer Grundlage in komplexen Situationen eine Erklärung zu überprüfen oder ein Problem zu untersuchen;
- die Angemessenheit der zur Unterstützung einer Behauptung oder eines Arguments herangezogenen Informationen kritisch zu prüfen bzw. in Fällen, in denen es keine einfache Kausalbeziehung gibt, auf der Basis der vorhandenen Information für oder gegen eine Erklärung oder Schlussfolgerung zu argumentieren;
- unter Bezugnahme auf naturwissenschaftliche Kenntnisse und vorhandene Daten oder Informationen ein gut aufgebautes Argument für oder gegen eine bestimmte Schlussfolgerung vorzulegen;
- auf der Grundlage eines Verständnisses von relativ komplexen und abstrakten naturwissenschaftlichen Konzepten Vorhersagen zu machen und Erklärungen zu geben.

Im Einzelnen wird sich die Skala für naturwissenschaftliche Grundbildung aus Analysen der Feldtestdaten ergeben. Diese Ergebnisse werden zeigen, welche Items sich zusammenfassen lassen und wie sie sich auf unterschiedliche Punkte der Skala verteilen. Die empirischen Daten sollen genutzt werden, um die hier vorgeschlagene progressive Skala naturwissenschaftlicher Grundbildung auf der Grundlage eigener Urteile sowie bereits vorliegender Erkenntnisse über kognitive Entwicklung zu testen.

Im Jahr 2006, wenn genügend Testzeit für eine breite Erfassung der naturwissenschaftlichen Konzepte und Anwendungsbereiche vorhanden sein wird, wird möglicherweise zusätzlich eine Subskala für das Verständnis von naturwissenschaftlichen Konzepten (Prozess 5) verwendet, das anhand deren Anwendung in vorgegebenen Situationen erfasst wird. Auf einer solchen Skala wird die Entwicklung **vom** Nachweis eines richtigen, aber unvollständigen Verständnisses von eher leichter zu begreifenden Konzepten **bis zum** Nachweis eines umfassenderen Verständnisses von eher komplexeren Konzepten dargestellt.

Im Jahr 2006 werden genügend Informationen über alle in Abbildung 14 aufgeführten naturwissenschaftlichen Prozesse vorliegen, so dass es eventuell möglich sein wird, Subskalen für Leistungen in einzelnen Prozessen oder in den Hauptbereichen der Naturwissenschaften zu bilden und zu berichten. Auch dies wird wieder von statistischen, konzeptuellen und schulpolitischen Überlegungen abhängen. Sollte sich eine Berichterstattung mit Hilfe solcher Subskalen als möglich erweisen, würden damit die Teilnehmerländer in die Lage versetzt, die im Einzelnen erzielten Ergebnisse ihres naturwissenschaftlichen Unterrichts mit den von ihnen für wünschenswert erachteten Zielen zu vergleichen.

Berichte über den Inhalt verschiedener Items und über falsche Antworten, die gegeben wurden, stellen eine wichtige Ergänzung zu den Itemstatistiken dar. Zu erwarten ist, dass sich entsprechende auf den Inhalt bezogene Kategorien aus dem Feldtest ergeben werden und den verschiedenen Typen der von den Schülern tatsächlich gegebenen Antworten zugeordnet werden können. Eine Berichterstattung über einige Typen von Antworten auf bestimmte Items wird ausserdem nötig sein, um die Skala bzw. die Skalen zu erläutern und ihnen sinnvolle Bezeichnungen zu geben. Hierfür müssen einige der in OECD/PISA verwendeten Items freigegeben werden.

Weitere Berichtsebenen wären wünschenswert und könnten nach der Haupterhebung zur naturwissenschaftlichen Grundbildung im Jahr 2006 möglich werden. Eine davon beträfe Leistungen in Itemgruppen, die über mehrere Aufgaben verteilt sind und sich auf die verschiedenen Anwendungsbereiche der Naturwissenschaften beziehen. Diese Informationen wären nützlich, um einzuschätzen, ob aktuelle Fragestellungen ausreichend und effizient berücksichtigt werden.

Sonstiges

Verknüpfungen mit anderen Bereichen

Wenn die Informationen für eine Testaufgabe zur naturwissenschaftlichen Grundbildung in Form eines längeren schriftlichen Textes vorgelegt werden, können gleichzeitig auch Aspekte der Lesekompetenz gemessen werden. Ebenso kann, wenn das Material aus Tabellen, Diagrammen, Graphen usw. besteht, zugleich die Fähigkeit gemessen werden, Informationen dieser Art zu lesen. Mit Items, die ein gewisses Maß an Umgang mit Zahlen verlangen, schließlich, können bestimmte Aspekte mathematischer Grundbildung gemessen werden. Aufgaben dieser Art werden als integrierte Items im Test enthalten sein. Mit anderen Aufgaben sollen nur naturwissenschaftliche Prozesse gemessen werden, zu denen das Ziehen von Schlussfolgerungen aus Belegen und der Nachweis des Verständnisses naturwissenschaftlicher Konzepte gehören.

Eine psychometrische Verknüpfung zwischen OECD/PISA und IEA/TIMSS wird aus den bereits im Kapitel über mathematische Grundbildung erläuterten Gründen nicht möglich sein.

Haupterhebungen und Nebenerhebungen im Bereich Naturwissenschaften

Die Erhebungen naturwissenschaftlicher Grundbildung in den Jahren 2000 und 2003, in denen die Naturwissenschaften nur eine Nebekomponente darstellen, werden die Grundlage für Vergleiche im Zeitverlauf bilden. Die beschränkte Zahl der Testaufgaben in den Jahren 2000 und 2003 bedeutet, dass es weniger Aufgaben zu den einzelnen Anwendungsbereichen der Naturwissenschaften geben wird als im Jahr 2006 (dies gilt obwohl das Erhebungsdesign unterschiedliche Zusammenstellungen von Items für unterschiedliche Untergruppen von Schülern vorsieht). Die „Nebenerhebungen“ zur naturwissenschaftlichen Grundbildung werden alle in Abbildung 14 aufgeführten Prozesse sowie einige Konzepte und Anwendungsbereiche aus Abbildungen 15 und 16 abdecken. Im Jahr 2006, dem Jahr in dem naturwissenschaftliche Grundbildung Hauptkomponente sein wird, wird dann eine weitaus breitere Erfassung der naturwissenschaftlichen Konzepte und Anwendungsbereiche möglich sein.

LITERATURHINWEISE

- ALMOND, R.G. und MISLEVY, R.J. (1998),
Graphical Models and Computerized Adaptive Testing, TOEFL Technical Report No.14, Educational Testing Service, Princeton, NJ, March.
- BAKER, L. (1991),
"Metacognition, reading and science education", in C. M. Santa and D. E. Alvermann (Hrsg.), *Science Learning: Processes and Applications*, International Reading Association, Newark, DE, S. 2-13.
- BENNETT, R.E. (1993),
"On the meanings of constructed response", in R.E. Bennett (Hrsg.), *Construction vs. Choice in Cognitive Measurement: Issues in Constructed Response, Performance Testing, and Portfolio Assessment*, Lawrence Erlbaum Associates, Hillsdale, NJ, S. 1-27.
- BINKLEY, M. und LINNAKYLÄ, P. (1997),
"Teaching reading in the United States and Finland", in M. Binkley, K. Rust und T. Williams (Hrsg.), *Reading Literacy in an International Perspective*, U.S. Department of Education, Washington, DC.
- BRUNER, J. (1990),
Acts of Meaning, Harvard University Press, Cambridge, MA.
- BYBEE, R.W. (1997),
"Towards an understanding of scientific literacy", in W. Grabe and C. Bolte (Hrsg.), *Scientific Literacy – An International Symposium*, IPN, Kiel.
- Council of Europe (1996),
Modern Languages: Learning, Teaching, Assessment. A Common European Framework of Reference, CC LANG (95) 5 Rev. IV, Strasbourg.
- Council of Ministers of Education, Canada (1994),
Student Achievement Indicators Program: Reading and Writing, Toronto.
- de LANGE, J. und VERHAGE, H. (1992),
Data Visualization, Sunburst, Pleasantville, NY.
- de LANGE, J. (1987),
Mathematics, Insight and Meaning, OW and OC, Utrecht.
- DEVLIN, K. (1994, 1997),
Mathematics, The Science of Patterns, Scientific American Library, New York.
- DOLE, J., DUFFY, G., ROEHLER, L. und PEARSON, P. (1991),
"Moving from the old to the new: Research on reading comprehension instruction", *Review of Educational Research*, 16 (2), S. 239-264.
- EHRlich, M.F. (1996),
"Metacognitive monitoring in the processing of anaphoric devices in skilled and less-skilled comprehenders", in C. Cornoldi und J. Oakhill (Hrsg.), *Reading Comprehension Difficulties: Processes and Interventions*, Lawrence Erlbaum Associates, Mahwah, NJ, S. 221-249.
- EHRlich, M. F., KURTZ-COSTES, B. und LORIDANT, C. (1993),
"Cognitive and motivational determinants of reading comprehension in good and poor readers", *Journal of Reading Behavior*, 25, S. 365-381.
- EINSTEIN, A. (1933),
"Preface to M. Plank", *Where is Science Going?*, Allen and Unwin, London.
- ELLEY, W.B. (1992),
How in the World do Students Read?, International Association for the Evaluation of Educational Achievement, The Hague.
- FREDERICKSON, N. (1984),
"The real test bias", *American Psychologist*, 39, S. 193-202.0

- FREUDENTHAL, H. (1973),
Mathematics as an Educational Task, Reidel, Dordrecht.
- FREUDENTHAL, H. (1983),
Didactical Phenomenology of Mathematical Structures, Reidel, Dordrecht.
- GRAEBER, W. und BOLTE, C (Eds) (1997),
Scientific Literacy – An International Symposium, IPN, Kiel.
- GRONLUND, N.E. (1968),
Constructing Achievement Tests, Prentice Hall, Englewood Cliffs.
- GRÜNBAUM, B. (1985),
"Geometry strikes again", *Mathematics Magazine*, 58 (1), S. 12-18.
- HAWKING, S.W. (1988),
A Brief History of Time, hier zitiert nach der deutschen Ausgabe: Eine kurze Geschichte der Zeit, Reinbek bei Hamburg, Rowohlt, 1988.
- HUBBARD, R. (1989),
"Notes from the underground: Unofficial literacy in one sixth grade", *Anthropology and Education Quarterly*, 20, S. 291-307.
- JONES, S. (1995),
"The practice(s) of literacy", in *Literacy, Economy and Society: Results of the First International Adult Literacy Survey*, OECD and Statistics Canada, Paris and Ottawa, S. 87-113.
- KIRSCH, I. (1995),
"Literacy performance on three scales: Definitions and results", in *Literacy, Economy and Society: Results of the First International Adult Literacy Survey*, OECD and Statistics Canada, Paris and Ottawa, S. 27-53.
- KIRSCH, I.S. und MOSENTHAL, P.B. (1989-1991),
"Understanding documents. A monthly column", *Journal of Reading*, International Reading Association, Newark, DE.
- KIRSCH, I. S. und MOSENTHAL, P. B. (1994),
"Interpreting the IEA reading literacy scales", in M. Binkley, K. Rust und M. Winglee (Hrsg.), *Methodological Issues in Comparative Educational Studies: The Case of the IEA Reading Literacy Study*, U.S. Department of Education, National Center for Education Statistics, Washington, DC, S. 135-192.
- KIRSCH, I., JUNGEBLUT, A. und MOSENTHAL, P. B. (1998),
"The measurement of adult literacy", in T. S. Murray, I. S. Kirsch, und L. Jenkins (Hrsg.), *Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey*, U.S. Department of Education, National Center for Education Statistics, Washington, DC.
- LANGER, J. (1995),
Envisioning Literature, International Reading Association, Newark, DE.
- LINNAKYLÄ, P. (1992),
"Recent trends in reading literacy research in Finland", in P. Belanger, C. Winter und A. Sutton (Hrsg.), *Literacy and Basic Education in Europe on the Eve of the 21st Century*, Council of Europe, Strasbourg, S. 129-135.
- LUNDBERG, I. (1991),
"Reading as an individual and social skill", in I. Lundberg und T. Høien (Hrsg.), *Literacy in a World of Change*, Center for Reading Research/UNESCO, Stavanger.
- MACCARTHEY, S.J. und RAPHAEL, T.E. (1989),
Alternative Perspectives of Reading/Writing Connections, College for Education, Institute for Research on Teaching, Occasional paper #130, Michigan State University.
- MILLAR, R. und OSBORNE, J. (1998),
Beyond 2000: Science Education for the Future, King's College London School of Education, London.
- MYERS, M. und PARIS, S.G. (1978),
"Children's metacognitive knowledge about reading", *Journal of Educational Psychology*, 70, S. 680-690.

- PARIS, S., WASIK, B. und TURNER, J. (1991),
 "The development of strategic readers", in R. Barr, M. Kamil und P. Mosenthal (Hrsg.), *Handbook of Reading Research, Vol. II*, Longman, New York.
- SENECHAL, M. (1990),
 "Shape", in L.A. Steen (Hrsg.), *On the Shoulders of the Giant – New Approaches to Numeracy*, National Academy Press, Washington D.C., S. 139-182.
- SHAFER, M.C. und ROMBERG, T. A. (in press),
 "Assessment in Classroom that Promote Understanding", in E. Fennema und T. A. Romberg (Hrsg.), *Mathematics Classroom that Promote Understanding*, Erlbaum, Mahwah, NJ.
- SHAMOS, M.H. (1995),
The Myth of Scientific Literacy, Rutgers University Press, New Brunswick.
- SMITH, M.C. (1996),
 "Differences in adults' reading practices and literacy proficiencies", *Reading Research Quarterly*, 31, S.196-219.
- STICHT, T.G. (Hrsg.) (1975),
Reading for Working: A Functional Literacy Anthology, Human Resources Research Organization, Alexandria, VA.
- STIGGINS, R.J. (1982),
 "An analysis of the dimensions of job-related reading", *Reading World*, 82, S. 237-247.
- STREEFLAND, L. (1990),
Fractions in Realistic Mathematics Education, A Paradigm of Developmental Research, Reidel Dordrecht.
- STUART, I. (1990),
 "Change", in: L.A. Steen (Hrsg.), *On the Shoulders of the Giant – New Approaches to Numeracy*, National Academy Press, Washington D.C., S. 183-218.
- TAUBE, K. und MEJDING, J. (1997),
 "A nine-country study: What were the differences between the low and high performing students in the IEA Reading Literacy Study?", in M. Binkley, K. Rust und T. Williams (Hrsg.), *Reading Literacy in the International Perspectives*, U.S. Department of Education, National Center for Education Statistics, Washington, DC, S. 63-100.
- TRAUB, R.E. (1993),
 "On the equivalence of the traits assessed by multiple-choice and constructed-response tests", in R. E. Bennett (Hrsg.), *Construction vs. Choice in Cognitive Measurement: Issues in Constructed Response, Performance Testing, and Portfolio Assessment*, Lawrence Erlbaum Associates, Hillsdale, NJ, S. 29-44.
- TRAVERS, K.J. und WESTBURY, I. (1989),
The IEA Study of Mathematics, Vol. 1, Analysis of mathematics curricula, Pergamon Press, Oxford.
- TREFFERS, A. (1986),
Three Dimensions, Reidel, Dordrecht.
- TREFFERS, A. und GOFFREE, F. (1985),
 "Rational analysis of realistic mathematics education", in L. Streefland (Hrsg.), *Proceedings of the Ninth International Conference for the Psychology of Mathematics Education (PME)*, OW and OC, Utrecht, S. 79-122.
- UNESCO (1993),
International Forum on Scientific and Technological Literacy for All, Final Report, UNESCO, Paris.
- WARD, W.C., DUPREE, D. und CARLSON, S.B. (1987),
A Comparison of Free-response and Multiple-choice Questions in the Assessment of Reading Comprehension (RR-87-20), Educational Testing Service, Princeton, NJ.
- WERLICH, E. (1976),
A Text Grammar of English, Quelle & Meyer, Heidelberg.
- ZIMAN, J.M. (1980),
Teaching and Learning about Science and Society, Cambridge University Press.

MITGLIEDER DER FUNKTIONALEN EXPERTENGRUPPEN

Lesen

Irwin Kirsch, Vorsitzender
Educational Testing Service
Princeton, New Jersey
USA

Marilyn Binkley
National Center for Education Statistics
Washington, DC
USA

Alan Davies
University of Edinburgh
Schottland
Vereinigtes Königreich

Stan Jones
Statistics Canada
Nova Scotia
Kanada

John de Jong
CITO, National Institute for Educational Measurement
Arnhem
Niederlande

Dominique Lafontaine
Universität Liège
Liège
Belgien

Pirjo Linnakylä
Universität Jyväskylä
Jyväskylä
Finnland

Martine Rémond
Institut National de Recherche Pédagogique
Paris
Frankreich

Wolfgang Schneider
Universität Würzburg
Würzburg
Deutschland

Ryo Watanabe
National Institute for Educational Research
Tokyo
Japan

Mathematik

Jan de Lange, Vorsitzender
Universität Utrecht
Utrecht
Niederlande

Raimondo Bolletta
Centro Europeo dell'Educazione
Frascati
Italien

Sean Close
St. Patricks College
Dublin
Irland

Maria Luisa Moreno
Instituto Nacional de Calidad y Evaluación (INCE)
Madrid
Spanien

Mogens Niss
Universität Roskilde
Roskilde
Dänemark

Kyung Mee Park
Korea Institute of Curriculum and Evaluation
Seoul
Korea

Thomas Romberg
University of Wisconsin-Madison
Madison, Wisconsin
USA

Peter Schüller
Bundesministerium für Bildung und Kultur
Wien
Österreich

Naturwissenschaften

Wynne Harlen, Vorsitzende
The Scottish Council for Research in Education
Edinburgh
Vereinigtes Königreich

Peter Fensham
Monash University
Melbourne
Australien

Raul Gagliardi
Genf
Schweiz

Donghee Shin
Korea Institute of Curriculum and Evaluation
Seoul
Korea

Svein Lie
Universität Oslo
Oslo
Norwegen

Manfred Prenzel
Institut für die Pädagogik der Naturwissenschaften (IPN)
Kiel
Deutschland

Senta Raizen
National Center for Improving Science Education
Washington, DC
USA

Elizabeth Stage
University of California
Oakland, California
USA

ÜBERLEGUNGEN FÜR ZUKÜNFTIGE ERHEBUNGSZYKLEN VON OECD/PISA

Erfahrungen und Einstellungen, die für die Erhebung naturwissenschaftlicher Grundbildung relevant sind

Neben den kognitiven Ergebnissen sollten auch Informationen zu den affektiven Ergebnissen des naturwissenschaftlichen Unterrichts erhoben werden. Für OECD/PISA sind ganz unterschiedliche Arten von Informationen relevant, die die Teilnahme der Schülerinnen und Schüler an naturwissenschaftlichen Aktivitäten sowie ihre Ansichten über den Wert solcher Aktivitäten für sie selbst und für die Gesellschaft allgemein betreffen. Da die Art der Aufgaben, die im Rahmen der OECD/PISA-Erhebungen verwendet werden, für manche Schülerinnen und Schüler neu sein könnte, ist es wichtig, zu erheben, wie vertraut ihnen Inhalt und Form der Aufgaben und Items sind. Diese Informationen können in manchen Fällen als erklärende Variablen dienen, in anderen Fällen die Daten zu den Ergebnissen des naturwissenschaftlichen Unterrichts bei Schülern und Schülerinnen signifikant ergänzen.

Daher ist geplant, in künftigen Zyklen folgende Informationen entweder mit dem Schülerfragebogen oder mit einem Testpaket zu erheben:

- Teilnahme an naturwissenschaftlichen Aktivitäten innerhalb und außerhalb der Schule (z. B. außerschulische naturwissenschaftliche Aktivitäten wie das Lesen naturwissenschaftlicher Zeitschriften, das Sehen von naturwissenschaftlichen Sendungen im Fernsehen und die Mitarbeit bei Aktivitäten innerhalb der Gemeinde, etwa in Umweltschutzorganisationen)
- Einschätzung des Nutzens der innerhalb und außerhalb der Schule erworbenen naturwissenschaftlichen Kenntnisse für persönliche bzw. die Gemeinschaft betreffende Entscheidungen
- Urteile über die Rolle, die die Naturwissenschaften bei der Entstehung oder Lösung von Problemen spielen
- Lernerfahrungen in Bezug auf die Themen, die die Grundlage der Testaufgaben bilden
- Vertrautheit mit der Form der Aufgaben und Items.

Mit Ausnahme des vierten Aspekts – Lernerfahrungen in Bezug auf die Themen, die die Grundlage der Testaufgaben bilden – sollten diese Informationen mit dem allgemeinen Schülerfragebogen erhoben werden. Der vierte Aspekt könnte in den Tests zu naturwissenschaftlicher Grundbildung als letzte Frage aufgenommen werden. Sie würde aus einer Liste der im Test vorkommenden Items bestehen, und die Schülerinnen und Schüler müssten jeweils angeben, ob sie schon häufiger, selten oder nie etwas über die mit diesem Item angesprochene Thematik gehört haben. Die Frage sollte erstens in Bezug auf die Schule und zweitens in Bezug auf außerschulische Kontexte beantwortet werden.

Alternative Itemformate für Mathematik

Wenn im Jahr 2003 mathematische Grundbildung Schwerpunkt von OECD/PISA wird, kann das Spektrum der für den Test verwendeten Aufgaben unter Umständen erweitert werden. Dies wäre insbesondere dann wichtig, wenn in künftigen Zyklen Aufgaben der Kompetenzklasse 3 stärker berücksichtigt werden sollen.

Aufgaben mit längeren, frei formulierten Antworten

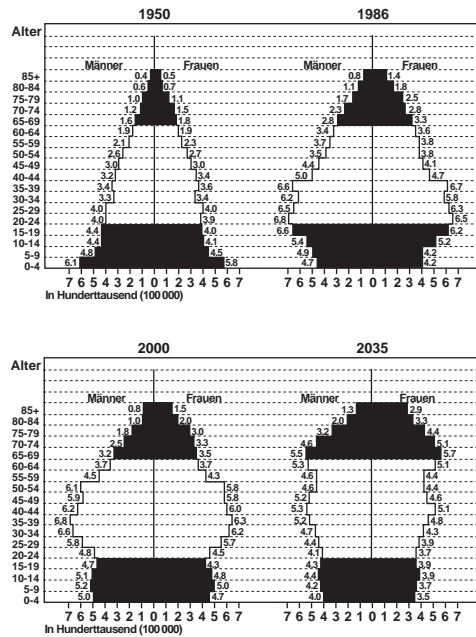
Eine Möglichkeit bestünde in der Verwendung von Aufgaben mit längeren, frei formulierten Antworten (extended-response essay tasks). Nach Gronlund (1968) sind derartige Aufgaben zwar ungeeignet, um Lernergebnisse im Sinne von Wissen zu messen, sie bieten jedoch bei den Antworten jene Freiheit, die nötig ist, um komplexe Lernergebnisse zu erfassen. Zu diesen komplexen Lernergebnissen gehören die Fähigkeit, schöpferisch zu sein, zu organisieren, zu integrieren, zu formulieren, sowie sonstige Handlungs- und Verhaltensweisen, bei denen Ideen entwickelt und in eine Synthese gebracht werden müssen.

Aufgaben mit längeren, frei formulierten Antworten geben den Schülern die Möglichkeit, tiefer in ein komplexes Problem einzudringen. Im Gegensatz zu einem Format, bei dem eine Reihe von Items an einen gemeinsamen Stimulus anknüpfen, können die Schülerinnen und Schüler hier bei der Untersuchung des Problems eigene Antworten entwickeln. Dieses Format eignet sich daher besonders zur Erfassung von Kompetenzklasse 3.

Abb. 20 zeigt ein Beispiel für eine Aufgabe mit einer längeren, frei formulierten Antwort.

Abbildung 20. **Mathematikaufgabe mit einer längeren, frei formulierten Antwort**

Die vier Pyramiden zeigen die Bevölkerung der Niederlande.
Die beiden ersten zeigen die tatsächliche Bevölkerung in den Jahren 1950 und 1986.
Die beiden letzten zeigen die erwartete Bevölkerung für die Jahre 2000 und 2035.



Beschreibe deine Überlegungen zu der Frage, inwieweit die Informationen aus diesen Bevölkerungspyramiden darauf schließen lassen, dass die niederländische Bevölkerung älter wird; benutze, wo es angebracht ist, verschiedenartige grafische Darstellungen, um deine Argumente zu unterstützen.

Mündliche Aufgaben

In einigen Ländern sind oder waren mündliche Tests allgemein üblich, und zwar sogar als Teil von allgemeinen nationalen Prüfungen. Es gibt verschiedene Formen von mündlichen Tests, wie zum Beispiel die folgenden:

- Gespräch über bestimmte, den Schülerinnen und Schülern bekannte mathematische Themen
- Gespräch über ein Thema, das den Schülerinnen und Schülern vor dem Gespräch als Hausaufgabe aufgegeben wurde
- Gespräch über ein nach der Prüfung als Hausaufgabe aufgegebenes Thema.

Das Format des mündlichen Tests wird häufig benutzt, um Prozesse höherer Ordnung zu operationalisieren.

Zwei-Stufen-Aufgaben

Eine Kombination unterschiedlicher Aufgabenformate kann als Zwei-Stufen-Aufgabe bezeichnet werden. Ein typisches Beispiel wäre eine schriftliche Aufgabe, an die sich eine mündliche Aufgabe zum gleichen Thema anschließt. Zwei-Stufen-Aufgaben verbinden die Vorteile von traditionellen schriftlichen Tests mit fester Zeitvorgabe mit den Möglichkeiten, die offene Aufgaben bieten.

Nachstehend wird eine Zwei-Stufen-Aufgabe beschrieben. Die erste Stufe besteht aus einem schriftlichen Test mit folgenden Merkmalen:

- der Test wird mit allen Schülern gleichzeitig durchgeführt
- der Test muss von allen Schülerinnen und Schülern innerhalb einer bestimmten Zeit bearbeitet werden

- der Test soll nicht in erster Linie ermitteln, was die Schülerinnen und Schüler wissen, sondern was sie nicht wissen
- Grundlage des Tests sind in der Regel Aktivitäten auf „niedrigeren“ Ebenen wie Rechnen und Basisverständnis
- der Test besteht aus offenen Fragen
- die Benotung ist so objektiv wie möglich

Auf der zweiten Stufe werden diejenigen Elemente ergänzt, die in der ersten Stufe fehlen. Die zweite Stufe hat folgende Merkmale:

- der Test wird ohne Zeitbegrenzung durchgeführt
- der Test kann zu Hause durchgeführt werden
- der Test konzentriert sich mehr darauf, was die Schülerinnen und Schüler wissen, als darauf, was sie nicht wissen
- der Test berücksichtigt insbesondere Aktivitäten auf höheren Ebenen wie Interpretation, Reflexion, Kommunikation usw.
- der Test hat eine offenere Struktur und besteht aus Fragen, für die lange Antworten bis hin zu kleinen Abhandlungen erforderlich sind
- die Bewertung kann schwierig und weniger objektiv sein

Produktionsitems

Testverfahren sollten den Schülern die Möglichkeit geben, ihre Fähigkeiten nachzuweisen, und als integraler Bestandteil des Lehr- und Lernprozesses betrachtet werden. Die Verwendung von eigenen Produktionen der Schülerinnen und Schüler ist nicht neu, und es liegen bereits umfangreiche Erfahrungen mit dieser Form von Test vor. Treffers (1987) führte die Unterscheidung zwischen Konstruktion und Produktion ein. Die freie Konstruktion lässt den Schülerinnen und Schülern die größte Freiheit, um eigene Gedanken und Fähigkeiten nachzuweisen. Sie umfasst unter anderem folgende Formate:

- Das Lösen offener Aufgaben, die zur Produktion unterschiedlicher Antworten anregen, weil verschiedene Lösungen möglich sind. Oft sind die möglichen Lösungen auf unterschiedlichen Ebenen der Mathematisierung angesiedelt.
- Das Lösen unvollständiger Aufgaben, bei denen die Schülerinnen und Schüler zunächst bestimmte Daten und Bezüge suchen und sich erarbeiten müssen.
- Das Entwickeln eigener (leichter, mittlerer, schwieriger) Aufgaben, die in Form eines Testbogens oder Aufgabenhefts zu einem bestimmten Thema für die nächste Schülerkohorte aufgeschrieben werden (Streefland, 1990).

Diese und andere mögliche Formate sollten im Hinblick auf ihre mögliche Verwendung im Rahmen von OECD/PISA in den kommenden Jahren weiter untersucht werden.

Wie gut sind Schülerinnen und Schüler auf die Herausforderungen der Zukunft vorbereitet? Verfügen sie über das Wissen und die Fähigkeiten, die sie benötigen, um als verantwortliche Bürger aktiv am gesellschaftlichen Leben teilzunehmen und um lebenslang zu lernen? Eltern, Schülerinnen und Schüler, Lehrkräfte, Schulverwaltungen und die allgemeine Öffentlichkeit haben Anspruch darauf, zu erfahren, inwieweit Heranwachsende die notwendigen Voraussetzungen auch tatsächlich erwerben. Mit Hilfe von internationalen Indikatoren kann beschrieben werden, welcher Leistungsstand in verschiedenen Staaten erreicht wird und von anderen angestrebt werden könnte. Weiterhin können die Indikatoren Orientierungshilfen für die Gestaltung von Unterricht und für das Lernverhalten von Schülerinnen und Schülern geben sowie Einsichten in die Stärken und Schwächen eines Curriculums vermitteln.

Die 29 Mitgliedsstaaten der OECD haben, zusammen mit einigen weiteren Ländern, das OECD Programme for International Student Assessment (PISA) initiiert, um auf regelmäßiger Basis internationale Indikatoren für Schülerleistungen zu erheben. Ziel von PISA ist es, zu erfassen, inwieweit Schülerinnen und Schüler gegen Ende der Pflichtschulzeit einige der zentralen Voraussetzungen für eine volle Teilhabe am gesellschaftlichen Leben erworben haben.

Eine neue Rahmenkonzeption für die Erfassung von Wissen und Fähigkeiten, der erste Band der PISA-Serie, beschreibt die konzeptuelle Rahmenkonzeption, die dem ersten Zyklus der PISA-Untersuchung (PISA 2000) zugrunde liegt. Es werden darin die Bereiche Lesekompetenz (reading literacy), mathematische Grundbildung (mathematical literacy) und naturwissenschaftliche Grundbildung (scientific literacy) definiert, die den Kern von PISA bilden. Diese Bereiche umfassen Wissensinhalte, die Schülerinnen und Schüler erwerben müssen, Prozesse, die sie ausführen können sollten und Kontexte, in denen das Wissen und die Prozesse angewendet werden. Weiterhin beschreibt die Rahmenkonzeption die Methoden, mit denen sichergestellt wird, dass die eingesetzten Aufgaben in allen Teilnehmerstaaten valide sind, relevante Eigenschaften präzise erfassen und auf authentischen, lebensnahen Situationen basieren.

Das OECD Programme for International Student Assessment (PISA)

PISA ist ein kooperativer Prozess, der die wissenschaftliche Kompetenz aus den Teilnehmerstaaten zusammenführt und gemeinschaftlich, auf der Basis gemeinsamer politischer Interessen, von den Regierungen dieser Staaten gesteuert wird. Die folgenden zentralen Merkmale von PISA machen die Besonderheit des Versuchs aus, „Literacy“ zu erfassen:

- Die Langfristigkeit des Projekts: Da sich PISA über das gesamte nächste Jahrzehnt erstreckt, eröffnet das Projekt den Teilnehmerstaaten die Möglichkeit, ihre Fortschritte kontinuierlich zu beobachten. Zukünftige Bände dieser Serie werden sich auf die Tests konzentrieren, deren Ergebnisse alle drei Jahre aktualisiert und publiziert werden sollen.
- Die getestete Altersgruppe: 15-Jährige. Leistungsmessungen gegen Ende der Pflichtschulzeit liefern wichtige Hinweise auf die Leistungsstärken und –schwächen von Bildungssystemen.
- Der „Literacy“-Ansatz: In PISA werden Wissen, Fähigkeiten und Kompetenzen nicht auf der Basis von Lehrplänen definiert und getestet, sondern ausgehend davon, was junge Menschen für ihr zukünftiges Leben voraussichtlich brauchen.
- Die weltweite Beteiligung: Insgesamt nehmen 32 Staaten an PISA teil; 28 Mitgliedsstaaten der OECD sowie Brasilien, China, Lettland und Russland