

Martin Baethge / Frank Achtenhagen /
Lena Arends / Edvin Babic / Volker
Baethge-Kinsky / Susanne Weber

Berufsbildungs-PISA

Machbarkeitsstudie

Pädagogik

Franz Steiner Verlag

Martin Baethge / Frank Achtenhagen u. a.
Berufsbildungs-PISA

Martin Baethge / Frank Achtenhagen /
Lena Arends / Edwin Babic /
Volker Baethge-Kinsky / Susanne Weber

Berufsbildungs-PISA

Machbarkeitsstudie



Franz Steiner Verlag 2006

Bibliografische Information der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte
bibliografische Daten sind im Internet über
<<http://dnb.ddb.de>> abrufbar.

ISBN-10: 3-515-08947-0
ISBN-13: 978-3-515-08947-0



ISO 9706

Jede Verwertung des Werkes außerhalb der
Grenzen des Urheberrechtsgesetzes ist unzulässig
und strafbar. Dies gilt insbesondere für Übersetzung,
Nachdruck, Mikroverfilmung oder vergleichbare
Verfahren sowie für die Speicherung in Datenver-
arbeitungsanlagen. © 2006 by Franz Steiner Verlag
GmbH, Stuttgart. Gedruckt auf säurefreiem,
alterungsbeständigem Papier.
Druck: Printservice Decker & Bokor, München.
Printed in Germany

INHALTSVERZEICHNIS

Vorbemerkung	7
1. Das Konzept der Machbarkeitsstudie für ein Berufsbildungs-PISA.....	9
1.1. Bedeutung, Perspektiven und Durchführung der Studie	9
1.2. Das Berufsbildungskonzept der Machbarkeitsstudie	12
2. Kompetenzkonzepte in der Berufsbildung – Definition und Ansätze.....	15
2.1. Zur Problematik des Kompetenzbegriffs.....	18
2.2. Schwierigkeiten der Operationalisierung und Messung von Kompetenzen	28
2.2.1. Ermittlung von Kompetenzen auf der Grundlage von externen Tätigkeiten.....	30
2.2.2. Ermittlung von Kompetenzen auf der Grundlage interner Bedingungen.....	33
2.3. Kompetenzen im Spannungsverhältnis von nationaler Spezifik und internationaler Vergleichbarkeit	35
2.4. Arbeitsdefinition für ein Konzept beruflicher Kompetenz.....	38
2.5. Vorschlag für die Datenerhebung.....	49
2.6. Messformate und Skalierung: Überlegungen zum Messmodell.....	61
3. Institutionelle und individuelle Einflussfaktoren für die Qualität beruflicher Bildungsprozesse.....	67
3.1. Kontextbedingungen in internationalen Vergleichsstudien zu Bildungssystemen und -prozessen	68
3.2. Indikatoren zur Beschreibung institutioneller und individueller Bedingungen von Berufsbildungsprozessen – zu einem Modell von Ausbildungsproduktions- und -verwertungsvariablen	77
3.2.1. Das Bildungs- und Beschäftigungssystem als Ganzes (Makroebene).....	81
3.2.2. Bildungseinrichtungen	87
3.2.3. Einzelne Bildungsteilnehmer und Lernende.....	91
3.3. Datenverfügbarkeit.....	97

4. Das Problem der vertikalen und horizontalen Vergleichbarkeit von Berufsbildungsstrukturen: Zur Untersuchungsanlage und Samplekonstruktion	101
4.1. Vertikale und horizontale Vergleichbarkeit und das Problem der Definition von beruflicher Erstausbildung.....	102
4.1.1. Vertikale Vergleichbarkeit.....	102
4.1.2. Horizontale Vergleichbarkeit.....	110
4.2. Konsequenzen für die Sample-Konstruktion	114
5. Untersuchungsanlage für ein Berufsbildungs-PISA – alternative Modelle...	125
5.1. Zur Ausgangslage.....	125
5.2. Das Forschungs-Design einer Längsschnittsstudie	127
5.3. Das Forschungs-Design einer Querschnittsstudie.....	132
6. Zum weiteren Vorgehen	137
7. Literaturverzeichnis	139
Anhang	I

VORBEMERKUNG

Das hier vorgelegte Buch basiert auf einem Forschungsprojekt, das das Bundesministerium für Wirtschaft und Arbeit (BMWA) nach einer Ausschreibung an die beiden Hauptautoren und die von ihnen repräsentierten Institute (Soziologisches Forschungsinstitut an der Universität Göttingen – SOFI [federführend] – und das Seminar für Wirtschaftspädagogik der Universität Göttingen) vergeben hat.

Unser Konzept der Machbarkeitsstudie sah vor, dass ein Berufsbildungs-PISA nicht nur allgemeine Grundkompetenzen messen, sondern auch in der Ausbildung erworbene Kompetenzzuwächse und die Verwertung von Kompetenzen auf dem Arbeitsmarkt und in der Arbeit erfassen sollte. Hierin und in der Einbeziehung von berufsfachlichen Kompetenzen unterscheidet sich ein Berufsbildungs- von bisher durchgeführten Schul-PISA. Ob und wie ein so weit gesteckter Untersuchungsrahmen realisierbar sein würde, war in der Machbarkeitsstudie zu prüfen. Daran sind ihre Resultate zu messen.

Die Fertigstellung der Machbarkeitsstudie im engen Zeitraum von einem Jahr war nur möglich, weil das Forschungsteam vielfältige Unterstützung erfuhr. An erster Stelle ist hier der internationale Kolleginnen- und Kollegenkreis zu nennen, der aus 17 Ländern zu zwei mehrtägigen internationalen Workshops in Göttingen zusammenkam (vgl. Teilnehmerliste im Anhang). Auch die interministerielle Arbeitsgruppe von BMWA, BMBF und BiBB, aus der mehrere Mitglieder an den Workshops teilnahmen, war für die Arbeit sehr hilfreich. Ein gleiches gilt für die Anregungen, die wir von weiteren Kolleginnen und Kollegen in Deutschland erfahren haben; hervorzuheben sind hier Steffen Kühnel vom Sozialwissenschaftlichen Methodenzentrum der Universität Göttingen und Jürgen Rost vom IPN in Kiel. Ihnen allen möchten wir an dieser Stelle danken. Eingeschlossen in den Dank ist auch Ministerialrat Walter Hoffmann (BMWA) für seine verlässliche und kompetente Begleitung des Projekts.

Ohne das hohe Engagement von Erika Beller, Barbara Dehne und Ingelore Stahn aus dem Sekretariat des SOFI sowie Christel Schikora und Heike Jachinke aus dem Seminar für Wirtschaftspädagogik wären weder die Workshops so harmonisch und ertragreich abgelaufen noch die diversen Textfassungen so akkurat präsentiert worden. Auch ihnen gilt unser Dank.

Wir hoffen, dass die Machbarkeits-Studie „Berufsbildungs-PISA“ die Diskussion über die Gestaltung der Berufsbildung gerade unter dem Vorzeichen des Europäischen Qualifikationsrahmens (EQF) beleben hilft – und dass es nicht bei einer Machbarkeitsstudie bleibt.

Für die Autoren
Martin Baethge
Frank Achtenhagen

Göttingen, 3. Juli 2006

1. DAS KONZEPT DER MACHBARKEITSSTUDIE FÜR EIN BERUFSBILDUNGS-PISA

1.1. BEDEUTUNG, PERSPEKTIVEN UND DURCHFÜHRUNG DER STUDIE

Der politische und ökonomische Sinn einer „internationalen Vergleichsstudie zur beruflichen Bildung“, die als „large scale assessment“ durchgeführt werden soll, liegt auf der Hand: Die fortschreitende Internationalisierung der ökonomischen Austauschbeziehungen auf den Güter- und Arbeitsmärkten sowie das ökonomische, politische und soziale Zusammenwachsen Europas unter der Bedingung steigender Wissensintensität der Arbeitsprozesse weisen den Bildungs- und dabei besonders den Berufsbildungssystemen eine erhöhte Bedeutung sowohl für die internationale Innovations- und Wettbewerbsfähigkeit der Unternehmen als auch für die Sicherung der Kompetenzen für (internationale) berufliche Mobilität und eigenständige Lebensführung zu. Die Europäische Union hat sich in der „Lissabon-Strategie“ ehrgeizige ökonomische und soziale Ziele gesetzt: Unter der strategischen Zielperspektive, die Europäische Union „zum wettbewerbsfähigsten und dynamischsten wissensbasierten Wirtschaftsraum der Welt zu machen“ spielt die verbesserte Partizipation an Bildung und Wissen und verstärkte Beteiligung am Arbeitsmarkt eine entscheidende Rolle (vgl. Europäische Sozialagenda, 2001). Sollen diese Ziele erreicht werden, dann müssen die Berufsbildungssysteme den Arbeitskräften neben hohen fachlichen Qualifikationen jene Kompetenzen vermitteln, die sie zu grenzüberschreitender beruflicher Mobilität und internationaler Kooperation befähigen.

Die Europäische Union hat in den letzten Jahren ihre Bemühungen intensiviert, mit der Entwicklung eines Europäischen Qualifizierungsrahmens (EQF) und eines europäischen Leistungspunktesystems (ECVET) mehr Transparenz über die Berufsbildung in Europa zu schaffen und die Gleichstellung von beruflichen Abschlüssen zu regeln. Die Erklärung von Kopenhagen vom 30. November 2002 und das Kommuniqué von Maastricht vom 14. Dezember 2004 stehen für diese forcierten Bemühungen, die ihren elaborierten Ausdruck z. B. in den acht Referenzniveaus des EQF gefunden haben (vgl. Kommission der Europ. Gemeinschaften, 2005; Sellin, 2005).

Gleichstellungsregelungen stellen eine notwendige *normative Definition* zur Klassifizierung von Ausbildungsgängen und -abschlüssen dar. Sollen sie ihren Hauptzweck, Glaubwürdigkeit und Vertrauen bei den Ausbildungseinrichtungen sowie institutionellen und individuellen Arbeitsmarktakteuren zu schaffen (Kommission der Europ. Gemeinschaften, 2005, S. 6 f.), erfüllen, muss man wissen, welche Kompetenzen in den unterschiedlichen europäischen Berufsbildungsprozessen tatsächlich erworben werden. Bei der Vielfalt der Berufsbildungsformen in

Europa kann solches Wissen auch dazu beitragen, dass die Länder in der Berufsbildung voneinander lernen.

Hierzu kann eine Vergleichsstudie zur Berufsbildung, die analog des „Programme for an International Student assessment“ (PISA) als „large scale assessment“ angelegt sein soll, einen wichtigen Beitrag leisten. Large scale assessment heißt, dass die Studie valide, reliabel und in objektiver Weise messen soll, welche Kompetenzen junge Menschen in ihrer beruflichen Ausbildung in den verschiedenen Ländern erwerben können, und wie man die Messergebnisse mit den mikro- und makrostrukturellen Bedingungen, unter denen die Kompetenzen erworben, genutzt und verwertet werden, verknüpfen kann. Gerade die Klärung der Zusammenhänge zwischen in Ausbildung erworbenen Kompetenzen und Arbeits- und Arbeitsmarkteffekten für Individuen und Unternehmen verleiht einer Berufsbildungsvergleichsstudie im Vergleich zum Schul-PISA eine besondere Perspektive. Zu dieser gehört auch, dass auf diesem Weg die Zusammenhänge zwischen fachlichen Qualifikationen und allgemeinen kognitiven Kompetenzen in ihrer Bedeutung für die individuelle Lebensgestaltung und Berufsplanung besser aufgeklärt werden können.

In Anlehnung an das Schul-PISA benutzen wir im Folgenden die Kurzformel „Berufsbildungs-PISA“ oder in der internationalen Kommunikation „PISA-VET“ (Programme for an International Student Assessment in Vocational Education and Training).

Machbarkeitsstudie bedeutet die Klärung

- der methodischen Voraussetzungen und Probleme, die bei einem international vergleichenden large scale assessment zu lösen sind;
- der Möglichkeiten und Probleme eines internationalen Vergleichs von Berufsbildungssystemen bzw. -institutionen;
- der Untersuchungsanlage (Sample-Bildung [Längsschnitts- oder Querschnittsstudie]);
- der einzubeziehenden Kompetenz-Dimensionen, insbesondere des Verhältnisses von fachlichen und fachübergreifenden allgemeinen Kompetenzen;
- des Verhältnisses von Kompetenzen zu Übergängen in Arbeit und Beschäftigung sowie zum Arbeitsmarktverhalten;
- der relevanten individuellen Hintergrundvariablen (sozio-ökonomische und lernbiografische Variable);
- der organisatorischen und finanziellen Probleme der Durchführung.

Es bestand Einvernehmen zwischen Auftraggeber und Verfassern, dass das Erstellen der Machbarkeitsstudie noch nicht die Vorlage der Instrumente für die Empirie heißen kann, da beiden Seiten bewusst ist, dass die Erstellung valider und reliabler Kompetenzmessverfahren einen hohen Entwicklungsaufwand erfordert, der im Rahmen der zur Verfügung stehenden Zeit (ein Jahr) nicht zu leisten war. Was zu leisten war und geleistet worden ist, ist das Aufzeigen von Möglichkeiten, zu solchen Messverfahren zu gelangen (vgl. Kapitel 2).

Zur Durchführung der Machbarkeitsstudie

Im Zentrum der Machbarkeits-Studie steht die Aufarbeitung relevanter Beiträge der nationalen und internationalen Forschung zur Kompetenzmessung in der beruflichen Bildung, zur internationalen Komparatistik von Berufsbildungssystemen und zur methodischen Anlage von large scale assessment-Studien. Ein „Berufsbildungs-PISA“ erweist sich in wenigstens zwei Aspekten gegenüber einem auf die allgemeinbildenden Schulen zielenden PISA als komplizierter: Zum einen darin, dass berufliche Bildung immer domänenspezifische Fachkompetenzen und nicht nur allgemeine Kompetenzen messen muss. Zum anderen variieren die Berufsbildungssysteme nach Eintrittsalter der Auszubildenden und Dauer der Ausbildung gravierend, so dass die Vergleichbarkeit schwierig ist. Entsprechend den Besonderheiten der Berufsbildung sollten bei der Aufarbeitung des Forschungsstands folgende Aspekte diskutiert werden.

Im Hinblick auf die Messung von Kompetenzen:

- (a) Diskussion des Kompetenzbegriffs und verschiedener Konzepte zur Kompetenzmessung vor dem Hintergrund ihrer kulturellen und disziplinären Einbettung: Identifikation von Kernelementen verschiedener Kompetenzbegriffe und ihrer Operationalisierung.
- (b) Vorschlag für eine Arbeitsdefinition des Kompetenzbegriffs und Perspektiven seiner Operationalisierung im Rahmen eines internationalen Vergleichs von Berufsbildungssystemen und -prozessen.
- (c) Erörterung relevanter Berufsfelder und Kompetenzbereiche für eine internationale Vergleichsstudie zur beruflichen Bildung.
- (d) Analyse geeigneter Konzepte für die Messung relevanter Domänen.

Bezüglich der mikro- und makro-strukturellen Aspekte und der institutionellen Bedingungen müssen folgende Probleme geklärt werden:

- (a) Vergleichbarkeit der Ausbildungsinhalte und der Niveaus der Ausbildungsziele (z. B. ISCED-Klassifikation).
- (b) Erörterung der institutionellen Einflussfaktoren für die Qualität von Ausbildungsprozessen.
- (c) Bestimmung sozio-ökonomischer und bildungsbiografischer Variablen zur empirischen Erfassung individueller Ausbildungsbedingungen.
- (d) Einigung auf Berufsfelder für das Sample (mit der notwendigen Berücksichtigung international vergleichbarer Berufsklassifikationen).

Ein zweiter Schwerpunkt bei der Erarbeitung der Machbarkeits-Studie lag bei der Vorbereitung, Durchführung und Auswertung zweier internationaler Experten-Workshops zum Thema. Diese Workshops fanden unter dem Titel PISA-VET vom 18. bis 20. November 2004 und vom 28. bis 30. April 2005 in Göttingen mit reger Beteiligung statt. Insgesamt nahmen 22 ausländische Expertinnen und Experten aus 13 Ländern (Australien, USA sowie aus 11 europäischen Ländern – ohne Deutschland¹) und zwei Vertreter von europäischen Berufsbildungs-(for-

1 Vgl. Teilnehmer-Verzeichnis der beiden Workshops im Anhang.

schungs-)institutionen (CEDEFOP und European Training Foundation – Turin²) teil. In den intensiven Diskussionen der Workshops konnten die Probleme eines Berufsbildungs-PISA und Lösungsmöglichkeiten aus der Perspektive der unterschiedlichen Länder in ganz anderer Weise ausgeleuchtet werden, als dies durch Literatur-Studium möglich gewesen wäre. Zugleich entstand auf diesem Weg ein Kompetenz- und Forschungsnetzwerk, auf das bei Umsetzung der Machbarkeitsstudie und der Erschließung von Ressourcen in den Ländern für eine Hauptstudie zurückgegriffen werden kann. Als Gewinn betrachten wir es auch, dass für Berufsbildung verantwortliche Experten aus BMWA und BMBF sowie aus dem Bundesinstitut für Berufsbildungsforschung (BIBB) an den Workshops teilnahmen und sich einen unmittelbaren Eindruck von der internationalen Diskussion verschaffen konnten.²

Die Ergebnisse von Literatur-Studium und Workshop-Expertise sind in den folgenden Kapiteln des Berichts in der Weise zusammengefasst, dass zunächst der Forschungsstand problemorientiert resümiert wird und auf dieser Grundlage Perspektiven für ein Berufsbildungs-PISA entwickelt werden. Das zweite Kapitel setzt sich mit Kompetenz-Konzepten und -Messverfahren auseinander; das dritte Kapitel ist den mikro- und makrostrukturellen Einflussfaktoren für berufliche Bildungsprozesse gewidmet; das vierte Kapitel behandelt die Probleme der Sample-Konstruktion eines international vergleichenden large scale assessment. Die Ergebnisse dieser drei Kapitel werden im fünften Kapitel zu alternativen Modellen der Untersuchungsanlage für ein Berufsbildungs-PISA verdichtet. Den Schluss bilden Überlegungen zum weiteren Procedere.

1.2. DAS BERUFSBILDUNGSKONZEPT DER MACHBARKEITSSTUDIE

Ein internationaler Vergleich der Leistungsfähigkeit beruflicher Bildung via Kompetenzmessung bedarf eines gemeinsamen Verständnisses von den Zielen der Berufsbildung. Ein solches gemeinsames Verständnis kann nicht unterstellt, es muss konsensuell wissenschaftlich, gegebenenfalls auch politisch hergestellt werden. Es macht für eine Vergleichsstudie einen Unterschied, ob man das Ziel beruflicher Bildung eng auf die Fähigkeit hin definiert, die funktionalen Anforderungen eines spezifischen Arbeitsplatzes erfüllen zu können, oder einer weiteren Zieldefinition folgt, nach der Berufsbildung neben funktionalen Qualifikationen auch allgemei-

2 Ergänzend zur internationalen Expertise haben wir auch zusätzliches wissenschaftliches Know how im nationalen Rahmen aktiviert. Die beiden Hauptverantwortlichen für die Machbarkeits-Studie haben am 04.10.2004 das Konzept vor der Sektion Berufs- und Wirtschaftspädagogik der Deutschen Gesellschaft für Erziehungswissenschaft in Siegen und am 11.10.2004 vor der Sektion „Learning and Professional Development“ der European Association for Research on Learning and Instruction vorgestellt und in einer ausführlichen Diskussion erörtert. Ähnlich hilfreich war ein Workshop am 4. April in Bonn mit Wissenschaftlerinnen und Wissenschaftlern des BIBB.

nere Kompetenzen zu vermitteln und einen Beitrag zur Entwicklung der Persönlichkeit zu leisten hat.

In Anlehnung an die aktuelle wissenschaftliche Debatte in Deutschland haben wir auf der Systemebene drei zentrale Zieldimensionen definiert, denen Ausbildungssysteme sich stellen müssen und die dementsprechend auch den Bezugspunkt für die Definition von Kompetenzen darstellen, die in der Berufsbildung vermittelt werden müssten:

- (1) die Entwicklung der individuellen beruflichen Regulationsfähigkeit – unter einer individuellen Nutzerperspektive und dem zentralen Aspekt der personalen Autonomie;
- (2) die Sicherung der Humanressourcen einer Gesellschaft und
- (3) die Gewährleistung gesellschaftlicher Teilhabe und Chancengleichheit

Um diese Zieldimensionen angemessen diskutieren, kritisieren und weiterentwickeln zu können, werden sie im Folgenden weiter ausdifferenziert und operationalisiert:

Ad (1): Die individuelle Regulationsfähigkeit (persönliche und biografische Nutzerperspektive) zielt auf das Vermögen des Individuums, sein Verhalten und Verhältnis zur Umwelt, die eigene Biografie und das Leben in der Gemeinschaft selbständig zu gestalten; individuelle Regulationsfähigkeit meint neben den rein fachlichen auch solche Kompetenzen, die auf Selbstorganisationsfähigkeit, Problemlösungsverhalten, Kommunikativität und Reflexivität (Metakognition) zielen. Wichtig ist, dass das Individuum stets in dem Zusammenhang von individueller Zielsetzung und Anstrengung einerseits sowie förderlichen und hindernden Umweltbedingungen andererseits gesehen wird, über welche die Spielräume zum Aufbau und zur Stabilisierung von personaler und beruflicher Identität definiert werden.

Ad (2): Die Sicherstellung der Humanressourcen (gesellschaftliche Bedarfs- als auch individuelle ökonomische Nutzerperspektive); hier sind alle Aspekte des Ausbildungssystems angesprochen, welche die individuelle Handlungsfähigkeit in der Erwerbsarbeit und auf dem Arbeitsmarkt fördern sowie den quantitativen und qualitativen Arbeitskräftebedarf sichern helfen. Angesprochen ist die Angemessenheit der Berufsbildungsprozesse im Hinblick auf die Entwicklung im Beschäftigungssystem – wie sie auch im Konzept der Mega-Trends gefasst werden (vgl. Achtenhagen, Nijhof & Raffe, 1995; Achtenhagen & Grubb, 2001; Baethge, Lanfer & Buß, 2003). Man kann bei dieser Betrachtung einen eher quantitativen sowie einen eher qualitativen Aspekt unterscheiden. Unter einer eher quantitativen Perspektive geht es vor allem um die Vermeidung von Disproportionalitäten zwischen Berufsbildungs- und Beschäftigungssystem; hierzu gehören zum Beispiel die Vermeidung zu geringer Ausbildungsangebote in stark nachgefragten oder einer Überausbildung in wenig nachgefragten oder wenig zukunftssträchtigen Berufsfeldern. Unter einer eher qualitativen Perspektive geht es vor allem um eine angemessene Vorbereitung auf jeweilig fachlich definierte Tätigkeitsfelder bis hin zur Ausbildung von Wissen und Fähigkeiten, die eine flexible Aufnahme von Tä-

tigkeiten in unterschiedlichen Arbeitsmarktsektoren erlauben. In der einschlägigen Diskussion wird diese Thematik unter den Aspekten der „Spezialisierung“ vs. „Generalisierung“ behandelt; hier liegen auch Schwerpunkte für die Bereiche der beruflichen Weiterbildung und des lebenslangen Lernens.

Ad (3): Die Gewährleistung gesellschaftlicher Teilhabe und Chancengleichheit (Integrations- bzw. Inklusionsperspektive); hierunter fallen alle Aspekte, welche das Verhältnis von Ausbildung und sozialer Strukturierung kennzeichnen; sie beziehen sich auf den Beitrag, den berufliche Ausbildung zum einen zur Verringerung der Herkunftsabhängigkeit von Erwerbs-, Bildungs- und Lebenschancen, zum anderen zur gesellschaftlichen Integration und dazu leistet, dass junge Menschen an der Gestaltung der sozialen und politischen Gemeinschaft teilnehmen können.

Wir haben diese Zieldimensionen in beiden internationalen Workshops diskutiert und über sie als generelle Perspektiven für berufliche Bildung Einvernehmen unter den Experten gefunden. Ein weiter Begriff von Berufsbildung, für den neben den funktionalen Aspekten der Beherrschung von Arbeitsplatzanforderungen auch allgemeine Kompetenzen der individuellen Handlungsfähigkeit auf dem Arbeitsmarkt, der Selbstorganisationsfähigkeit in Bezug auf die eigene Berufsbiografie und lebenslanges Lernen sowie der Teilhabe am Gemeinschaftsleben konstitutiv sind, stößt auch in der internationalen Diskussion auf Konsens. Allerdings war allen Teilnehmern auch bewusst, dass ein in dieser Weise weites Konzept von Berufsbildung Konsequenzen für die in die Untersuchung aufzunehmenden Kompetenzkategorien haben und nicht leicht zu operationalisieren sein würde (vgl. dazu Abschnitt 2.1).

Dementsprechend wurde angemahnt, große Sorgfalt auf die Frage der Operationalisierung und der Verknüpfung der unterschiedlichen Kompetenzarten zu legen und im Auge zu behalten, dass alle Kompetenzen – funktionale wie fachübergreifende – auf verschiedenen Wegen vermittelt werden können und nicht allein der institutionalisierten beruflichen Bildung zuzuschreiben sind. In diesem Zusammenhang ist auch auf das so genannte informelle Lernen zu verweisen, über das Arbeiter und Angestellte Kompetenzen außerhalb von organisierten Lehr-Lernprozessen erwerben.

2. KOMPETENZKONZEPTE IN DER BERUFSBILDUNG – DEFINITIONEN UND ANSÄTZE

Eine der zentralen Fragestellungen im Rahmen dieser Machbarkeitsstudie ist, wie geeignete Instrumente für einen internationalen Berufsbildungsvergleich gefunden werden können. Die Komplexität des Kompetenzkonzepts wurde sowohl bei unseren Recherchen der aktuellen internationalen Literatur zur Kompetenzmessung und zu Kompetenzen allgemein als auch aufgrund der Ergebnisse des ersten und zweiten Workshops deutlich: Es existiert zurzeit weder ein einheitlicher, allgemein akzeptierter Kompetenzbegriff, noch gibt es einen „Königsweg“ der Operationalisierung und Messung von Kompetenzen.

Entsprechend der jeweiligen wissenschaftlichen und kulturellen Ausrichtung findet man eine unüberschaubare Vielzahl unterschiedlicher Ansätze. Norris (1991) unterscheidet beispielsweise drei grundlegende Forschungstraditionen zum Kompetenzkonzept seit dem zweiten Weltkrieg: Die hauptsächlich in den USA vorherrschende behaviouristische, die in England verbreitete und hauptsächlich auf Managementprobleme ausgerichtete Hervorhebung allgemeiner („generic“) Ansätze und die vor allem von der Linguistik her angestoßene kognitive Tradition. Weinert hebt sogar neun verschiedene theoretische Konzepte zur Definition und Erfassung von Kompetenz hervor: allgemeine kognitive Fähigkeiten, spezielle kognitive Fähigkeiten, Kompetenz-Performanz-Modell, modifiziertes Kompetenz-Performanz-Modell, motivierte Handlungstendenzen, objektive und subjektive Selbstkonzepte, Handlungskompetenz, Schlüsselkompetenz und Meta-Kompetenz (Weinert 2001, S. 6–14).

Das Problem der Kompetenzmessung stellt sich in der Berufsbildung ungleich komplizierter dar als im Bereich der allgemeinbildenden Schulen. Während eine international vergleichende Schulleistungsuntersuchung wie TIMSS („Third International Mathematics and Science Study“) auf die Untersuchung von Mathematik- und Naturwissenschaftsleistungen in der Grundschule und Sekundarstufe I beschränkt war, und die PISA-Studie („Programme for International Student Assessment“) sich auf „literacy“ und „numeracy“ konzentrierte, wird die Komplexität bei einem Berufsbildungs-PISA u. a. sowohl durch die Erfassung von Aspekten der Bewährung am Arbeitsplatz und auf dem Arbeitsmarkt als auch durch die Relevanz von funktionalen Kompetenzaspekten (z. B. motorische Fähigkeiten, Geschicklichkeit) erhöht. Schulleistungsuntersuchungen wie TIMSS und PISA können auf eine langjährige Forschungstradition zurückgreifen und sich gleichsam auf ein „Weltcurriculum“ in Mathematik und Naturwissenschaften einigen. In einer vergleichbar methodisch komfortablen Situation befindet sich ein Berufsbildungs-PISA nicht. Die Fülle und Heterogenität beruflicher Spezialisierungen bereits innerhalb einer Gesellschaft und noch mehr deren vielfältige Ausprägungen

in verschiedenen Gesellschaften machen die Suche nach einheitlichen Kompetenzstandards fast zu einer Sisyphus-Arbeit. Hinzu kommt die Fülle der Kompetenz-Dimensionen, die in einem Berufsbildungs-PISA zu berücksichtigen sind: Es sind nicht allein fachliche, sondern ebenso kognitive Basiskompetenzen wie beim PISA für allgemeinbildende Schulen und berufsübergreifende Aspekte der employability, die in Rechnung zu stellen und zu erfassen sind. Eine Schneise durch diese Komplexität zu schlagen, sie nicht einfach nur dezisionistisch zu reduzieren, ist eine der Aufgaben der Machbarkeitsstudie. Gerade die Frage, wie sich das Verhältnis zwischen allgemeinen kognitiven und fachspezifischen Kompetenzen im Laufe von Ausbildung und Arbeitsleben entwickelt, gewinnt heute unter der Bedingung lebenslangen Lernens zentrale Bedeutung.

Im Falle der Berufsbildung liegen dazu nur wenige bzw. gar keine Vorarbeiten zu Vorstellungen von den Merkmalen bzw. der Entwicklung beruflicher Expertise auf der internationalen Ebene vor³ und es fehlt ein allgemeiner internationaler Konsens über fachliche Standards und Niveaus. Anfänge dazu lassen sich für Europa durch Bemühungen der EU erkennen, im Rahmen des „Kopenhagen-Prozess“ zu einem einheitlichen europäischen Qualifikationsrahmen (EQF) und Leistungspunkte-System (ECVET) zu kommen (vgl. Kommission der Europ. Gemeinschaften, 2005; Sellin, 2005). Ein Berufsbildungs-PISA könnte diese Ansätze wissenschaftlich unterstützen.

Bedingt durch die Differenzen zwischen den jeweiligen Arbeitsmärkten und Kulturen muss im Kontext eines Berufsbildungs-PISA der Stellenwert der curricularen Ziele überdacht werden; denn die meisten Berufsbildungssysteme kennen nicht das System formaler berufsqualifizierender Abschlüsse, wie das etwa in Deutschland der Fall ist. Sie bilden von daher lediglich eine – durchaus nicht immer zwingend erforderliche – Grundlage dafür, in den Arbeitsmarkt einzutreten. Diese Fragen des Vorhandenseins der Vergleichbarkeit der curricularen Ziele und Inhalte werden für die weitere Arbeit des Berufsbildungs-PISA eine zentrale Rolle zu spielen haben – sowohl politisch-institutionell und lehr-lerntheoretisch als auch messmethodisch.

Den drei Zielen (vgl. Kapitel 1) entsprechend hat ein Berufsbildungs-PISA drei Kompetenz-Ebenen zu thematisieren und miteinander zu verknüpfen:

1. Die Ebene *allgemeiner kognitiver Grundkompetenz* des Lesens, Schreibens und Rechnens sowie der Problemlösefähigkeit, die in bisherigen Kompetenz-Studien für allgemeinbildende Schulen bearbeitet worden sind und die für die Bewältigung jedweder Lebens- und Arbeitssituation unabdingbar sind.
2. Die Ebene *berufsübergreifender arbeitsbezogener Kompetenz*, die auf das Agieren und Interagieren in Erwerbsorganisationen und auf dem Arbeitsmarkt

3 Als Anfänge für die Entwicklung allgemeiner fachlicher Standards sind auf nationaler Ebene in den USA die Allgemeinen Standards für Wirtschaftliche Bildung (National Council on Economic Education, 1997) sowie international die Allgemeinen Standards für Technische Bildung („Standards for Technological Literacy“ und „Advanced Excellence in Technological Literacy“) der International Technology Education Association (2002) anzusehen.

bezogen sind und heute häufig mit dem Begriff „employability“ zu fassen versucht werden. Zu ihnen zählen Kenntnisse über die Funktionsweise von Organisationen und Arbeitsmärkten ebenso wie Fähigkeiten zum Interagieren in sozial heterogenen Gruppen, zu Selbstorganisation und zur interaktiven Nutzung von technischen Hilfsmitteln. In der Literatur werden solche Kompetenzen oft unter „generic skills“ oder „Schlüsselqualifikationen“ subsumiert. (In der Realität gibt es fließende Übergänge zu allgemeinen wie auch zu berufsfachlichen Kompetenzen, z. B. „Problemlösen“ oder „strategisches Wissen“.)

3. Die Ebene *berufsfachlicher Kompetenzen* umfasst alle Fähigkeiten und Qualifikationen, die für die erfolgreiche Ausfüllung spezifischer Berufsrollen und der Bewältigung von Arbeitsvollzügen erforderlich sind.

Für unseren Zusammenhang geht es darum, wie die verschiedenen Kompetenzebenen angemessen auf die berufliche Aus- und Weiterbildung bezogen werden können. Dabei spielen die Definitionsversuche für die mittlere Ebene, d. h. der Bezug auf „skills for employability“, sowie für die Ebene der berufsfachlichen Kompetenzen eine besondere Rolle. Wir verfolgen die Diskussionen um die Ausprägungen allgemeiner Kompetenzen, wie sie im Anschluss an die Schul-PISA-Studien geführt wurden, hier nicht, da es zu dieser Thematik einen elaborierten Diskussionsstand gibt (vgl. Abbildung 5.2).

Die seit vielen Jahren andauernden und mit unterschiedlichen Zielen geführten nationalen Diskussionen um „generic skills“⁴ können grob in drei Forschungssträngen klassifiziert werden (vgl. Smith & Comyn, 2003; Kearns, 2001):

- die amerikanisch geprägte Diskussion um breit angelegte „generic skills“ in den Bereichen Allgemeinbildung, Persönlichkeitseigenschaften, Werte, Lernstrategien sowie allgemein auf den Arbeitsplatz bezogene Skills (wie z. B. Kommunizieren von Ideen und Informationen, Problemlösen, etc.);
- die anglo-australische Diskussion, die sich eher auf ein eng definiertes Set von „generic skills“ ähnlicher Art bezieht, ohne jedoch auf Persönlichkeitseigenschaften, Werte etc. einzugehen und
- schließlich ein europäisch geprägter Ansatz, der breiter angelegte soziale und arbeitsplatzbezogene Aspekte im Rahmen des lebenslangen Lernens (wie z. B. „citizenship“) focussiert.

Die unterschiedlichen kulturell geprägten Vorstellungen zu den entscheidenden „generic skills“, um in der heutigen Gesellschaft im Beruf sowie im Leben allgemein erfolgreich zu sein, wurden u. a. in dem von der OECD durchgeführten DeSeCo Projekt zusammenzufügen versucht. Im Rahmen des Projekts wurden

- 4 Aufgrund der Fülle von Terminologien in verschiedenen Ländern (Australien: Key competencies; Neuseeland: Essential skills; USA: SCANS & workplace know-how; UK: Core/common skills & employability skills; Frankreich: Transferable competencies; Deutschland: Key qualifications; Dänemark: Process independent qualifications; Italien: Transversal competencies; Niederlande: Core competencies; Kanada: Strategy for prosperity; Südafrika: Critical cross field outcomes; Singapur: Critical enabling skills training; Finland: Framework for evaluating educational outcomes; Norwegen: Core curriculum; Schweiz: Transdisciplinary goals), wird hier die englische Bezeichnung „generic skills“ verwendet.

Ansätze aus 12 verschiedenen OECD Ländern im Hinblick auf ihre Gemeinsamkeiten analysiert. Dabei konnte festgestellt werden, dass bestimmte „generic skills“ in allen Ländern (Muttersprache; Kommunikation; Sozialkompetenz/Kooperation/Teamwork; Information/Problemlösen; Autonomie/Selbstverwaltung; Handlungsorientierung/Entscheidungsfindung; Werte/Ethik) betont wurden, während in anderen Ländern nur einige „generic skills“ genannt wurden (Kreativität/Ausdruck/Ästhetik; Fremdsprachen; Kulturelle Identität und Tradition; Religion; Politik; Ökologisches Bewusstsein/Naturbewusstsein; körperliche Fitness/Gesundheit) (vgl. Smith & Comyn, 2003). Zugleich zeigen andere Studien, dass „generic skills“ erst Bedeutung als Teil eines Ganzen innerhalb eines beruflichen Kontextes (wie z. B. innerhalb der Domäne Elektrotechnik) erhalten (vgl. Nijhof 1998).

Im Folgenden wird vor allem die Komplexität der Kompetenzkonzepte und möglicher Operationalisierungsansätze für eine international vergleichende Studie dargestellt und in Bezug auf jeweilige Vor- und Nachteile unter dem Aspekt ihrer Bedeutung für ein Berufsbildungs-PISA erörtert.

2.1. ZUR PROBLEMATIK DES KOMPETENZBEGRIFFS

Hinsichtlich des Begriffs der Kompetenz finden wir in unterschiedlichen Ländern und wissenschaftlichen Disziplinen eine Vielfalt von differierenden Definitionen (vgl. die Beispiele in Abbildung 2–1).

Gemeinsam ist den Kompetenz-Definitionen der Bezug auf Handlungsfähigkeit. Ihre Unterschiede reflektieren zumeist kulturell und historisch bedingte sprachliche Unterschiede, die zu terminologischen Verwirrungen führen. Diese Tatsache wurde auch auf dem 2. Workshop von vielen Teilnehmern bestätigt. So machte Alan Brown (Great Britain) deutlich, dass für das Vereinigte Königreich der Begriff ‚Competence‘⁵ im Gegensatz zu Deutschland viel stärker negativ belegt sei und ‚methodological competence‘ (Methodenkompetenz) eine ganz andere Bedeutung hätte. Ähnlich betonte Frank Coffield (Great Britain), dass aus historischer Perspektive der Kompetenzbegriff in Deutschland im Gegensatz zu England als wissenschaftliche Kategorie akzeptiert würde.

Anhand einer Analyse der einschlägigen Fachliteratur kommt Ellström (1997) zu dem Schluss, dass es unmöglich sei, eine einheitliche Theorie oder Definition zu formulieren, welche die Vielfältigkeit unterschiedlicher Ansätze in Einklang bringen könne. Norris (1991) begründet das „theoretische Chaos“ damit, dass die stillschweigend zugrundegelegten Bedeutungen des Kompetenzbegriffs die Notwendigkeit einer präzisen Definition und Operationalisierung des Konzepts überschatteten. Die Folge davon sei, dass das vermeintlich Einfache grundlegend kompliziert werde (S. 331). Auf beiden Workshops in Göttingen wurde die Wichtig-

5 Zur Vermeidung von linguistischen Problemen werden die englischen Begriffe im Zusammenhang mit dem Kompetenzkonzept zugrundegelegt

keit einer grundsätzlichen Klärung semantischer Differenzen des Kompetenzkonzepts hervorgehoben und erörtert, wieweit eine weniger wissenschaftlich und mehr pragmatisch orientierte Kompetenzdefinition, wie sie beispielsweise das CEDEFOP vorschlägt (vgl. Abbildung 2–1), einen Weg abgeben könnte.

Abbildung 2–1: Beispiele für Kompetenzdefinitionen

„Die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert, 2002, S. 27 f.).
„The proven/demonstrated – and individual – capacity to use know-how, skills, qualifications or knowledge in order to meet usual – and changing – occupational requirements“ (Bjornavold, 2000, S. 208).
„The ability to successfully meet complex demands in a particular context through the mobilization of psychosocial prerequisites (including both cognitive and non-cognitive aspect“ (Rychen & Salganik, 2003, S. 43).
„An ability that extends beyond the possession of knowledge and skills. It includes: i) cognitive competence involving the use of theory and concepts, as well as informal tacit knowledge gained experientially; ii) functional competence (skills or know-how), those things that a person should be able to do when they work in a given area; iii) personal competence involving knowing how to conduct oneself in a specific situation; and iv) ethical competence involving the possession of certain personal and professional value“ (Lisbon-to-Copenhagen-to-Maastrich Consortium Partners, 2004, S. 89)
„The ability to apply knowledge, know-how and skills in an habitual and/ or changing work situation“ (CEDEFOP, 2003, S. 39).

Mit ähnlichen Problemen haben auch die derzeitigen kompetenzorientierten Anstrengungen auf der EU-Ebene zu kämpfen, wo Kompetenzen als Instrument für Transparenz und Flexibilität eingeführt werden. Im Zuge der Entwicklung eines gemeinsamen Europäischen Qualifikationsrahmens (EQF) sollen gemeinsame Bezugsebenen in einer kompetenzorientierten Matrix, in die sowohl komplette Ausbildungsgänge als auch einzelne Kompetenzmodule eingeordnet werden können, hergestellt werden. Die vertikale Dimension besteht aus Niveauangaben („Levels“) (verantwortlich: QCA (Qualifications and Curriculum Authority, London)), die in der horizontalen Ebene mit Hilfe von Deskriptoren hinsichtlich Art und

Umfang der erworbenen Kompetenzen differenziert werden (Vorschlag: Winterton, Delamare-LeDeist & Stringfellow, 2005). Hierzu wird als Ergänzung zum bestehenden European Credit Transfer System (ECTS) für die Höhere Bildung ein modularisiertes und kompetenzorientiertes Anrechnungssystem für Berufsbildung geschaffen: das European Credit Transfer System for Vocational Education and Training (ECVET). Die Struktur der Matrix ist in Abbildung 2–2 dargestellt:

Abbildung 2–2: Struktur der Matrix ECVET

Level	Cognitive Competence (Knowledge)	Functional Competence (Skills)	Social and Meta-Competence (Behaviours and Attitudes)
Level 8			
Level 7			
Level 6			
Level 5			
Level 4			
Level 3			
Level 2			
Level 1			

Quelle: Winterton, Delamare Le-Deist & Stringfellow, 2005, S. 42.

In dem von Winterton, Delamare Le-Deist & Stringfellow (2005) vorgelegten Bericht für die Unterstützung der Entwicklung eines ECVET wird für die horizontale Dimension ein Prototyp für eine Typologie von Kompetenzbereichen präsentiert, der auf einem Vergleich von kompetenzorientierten Ansätzen in der Berufsbildung in verschiedenen europäischen Ländern beruht. Als gemeinsame Schnittmengen wurden „Knowledge“ (Kenntnisse), „Skills“ (Fertigkeiten) und „Behaviours and Attitudes“ (Verhaltensweisen und Einstellungen) identifiziert (vgl. Abbildung 2–2).⁶

6 Die Spaltenbezeichnungen der Matrix sind von der Kommission der Europäischen Gemeinschaften (2005) nach vielfacher Kritik weiter ausdifferenziert worden in „Knowledge“ (Kenntnisse), „Skills“ (Fertigkeiten) und „Personal and Professional Competence“ (persönliche und fachliche Kompetenz), diese letztere Kategorie ist weiter unterteilt in „Autonomy and Responsibility“ (Selbständigkeit und Verantwortung), „Learning Competence“ (Lernkompetenz), „Communication and Social Competence“ (Kommunikationskompetenz und soziale Kompetenz), und „Professional and Vocational Competence“ (fachliche und berufliche Kompetenz). Dennoch bestehen die genannten Probleme der Überscheidungen der Spaltenbezeichnungen und den damit verbundenen Schwierigkeiten der Operationalisierung sowie die

Hinsichtlich der Anzahl der „Level“ einigte sich die QCA auf acht Niveaustufen, da diese im weitesten Sinne mit den Vorschlägen von ISCED (International Standard Classification of Education), des CEDEFOP (European Centre for the Development of Vocational Training) und ISCO (International Standard Classification of Occupations) kompatibel seien. In der von Winterton, Delamare Le-Deist & Stringfellow präsentierten Matrix soll die Komplexität der Niveaus von Level 1, welches Lernergebnisse der verpflichtenden allgemeinen Schulbildung umfasst, bis Level 8, welches diejenigen Qualifikationen abdeckt, die auf hochspezialisierte Experten in Führungspositionen mit komplexen Aufgabenbereichen zutreffen, abgebildet werden (Winterton, Delamare Le-Deist & Stringfellow, 2005, S. 42).

Die derzeitige Diskussion dreht sich um die Frage, ob es sinnvoll sei, Sub-Levels zu integrieren, um die Flexibilität zu erhöhen und Differenzen zwischen verschiedenen nationalen Ausbildungen zu berücksichtigen. Ohne Sub-Levels wären beispielsweise sowohl die dänische als auch die englische Ausbildung auf der vierten Stufe einzuordnen, obwohl die dänische Ausbildung bedeutend breiter und inhaltlich tiefergehend angelegt ist als die englische. Die Problematik der angemessenen Klassifizierung wird auch in einer Studie von Fust-Bleil & Ebner (2005) bei einem Leistungsvergleich zwischen dem englisch-walisischen berufsqualifizierenden Abschluss „GNVQ Advanced Business“ mit dem deutschen Berufsabschluss „Industriekaufmann/Industriekauffrau“ deutlich. Auf der Grundlage eines 5-Stufen Schemas (vgl. EUR-Lex, 1985) kommen die Autoren zu dem Ergebnis, dass der deutsche Berufsabschluss mindestens auf – wenn nicht sogar über – dem Niveau des Englisch-Walisischen Abschlusses eingeordnet werden müsste und nicht wie vorgesehen unterhalb auf Stufe 2. Die Spitzenverbände der deutschen Wirtschaft schlagen in ihrer Stellungnahme zum EQF und ECVET im Hinblick auf die angemessene Einordnung der deutschen anerkannten Ausbildungsberufe sechs Niveaustufen vor, auf denen, sowohl für die Berufs- als auch für die Hochschulbildung, Kompetenzprofile von beruflichen Domänen zugeordnet werden sollen (vgl. Kuratorium der deutschen Wirtschaft für Berufsbildung, 2005).

Die Entwicklung der Niveaus basiert auf bereits existierenden Nationalen Qualifikationsrahmen in anderen EU-Mitgliedsstaaten, insbesondere auf denen von Irland und Schottland. Der schottische Rahmen (vgl. Scottish Credit and Qualifications Framework, 2003) umfasst 12 Stufen, während der irische 10 beinhaltet

mangelnde Einbeziehung zentraler Ziele für die berufliche Bildung weiterhin. In Weiterführung der Vorschläge von Winterton, Delamare Le-Deist & Stringfellow (2005) hat die Kommission der Europäischen Gemeinschaften (2005) ein Schema entwickelt, das versucht die acht Stufen des Europäischen Qualifikationsrahmens nach Lernergebnissen zu operationalisieren (vgl. Anhang Kapitel 2, Dokument 6). Eine grundsätzliche Schwierigkeit besteht darin, dass die Operationalisierungsvorschläge weitgehend auf semantischen Verschiedenheiten zwischen Verben, Adjektiven und Adverbien setzen. Damit bleiben sie auf einem Assoziationsniveau; für empirische Überprüfungen bzw. einen Einsatz in einem large-scale assessment wären umfangreiche Anpassungen notwendig.

(vgl. National Framework of Qualifications, 2003). Beide Vorschläge sind von der Absicht geprägt, die jeweiligen Bildungssysteme einem hohen internationalen Standard entsprechend neu zu konzipieren. Die Frage, wie diese vorhandenen Konzepte in den Europäischen Qualifikationsrahmen integriert werden könnten, muss in weiteren Diskussionen anhand eines Vergleichs der horizontalen Deskriptoren geklärt werden. Sowohl der schottische als auch der irische Qualifikationsrahmen integrieren die berufliche und höhere Bildung.

Die Autoren ziehen das Fazit, dass in vielen EU-Mitgliedsstaaten eine Tendenz zu output-orientierten und kompetenzbasierten Berufsbildungssystemen bestehe, was auf ein allgemeines Interesse an multidimensionalen Rahmenwerken schließen lasse. Hinsichtlich der drei Kompetenzbereiche (Knowledge, Skills, Behaviours and Attitudes) bestehe grundsätzliche Übereinstimmung in einzelnen EU-Mitgliedsstaaten. Abweichungen ließen sich auf terminologische Gründe zurückführen; allerdings existierten in anderen Fällen fundamentale und konzeptionelle Differenzen. Die Herausforderung bestehe in der Entwicklung einer konsistenten und einheitlichen Typologie für die drei Kompetenzbereiche, die in der Lage sei, die vorhandenen Differenzen widerspruchsfrei zu operationalisieren und einander zuzuordnen. Als Vorteile einer Einordnung nach dem „Level“-System werden folgende Punkte hervorgehoben:

- Es ist möglich, eine präzise wissenschaftliche Sprache für die Kommunikation über Kompetenzen und Leistungserbringung zu entwickeln.
- Outcome-basierte Qualifikationen haben ein hohes Niveau an Validität, Glaubwürdigkeit und Nutzbarkeit, da sie auf Inhalten der Arbeitsprozesse basieren.
- Für die Zielsetzungen von modernen nationalen Berufsbildungssystemen ist es ausreichend, wünschenswerte Outcomes festzulegen, da diese sowohl die Basis für die Anerkennung von Kompetenzen als auch für den Unterricht im Sinne von erwünschten Lernergebnissen sind.
- Bedeutende Vorteile kommen durch ein Assessment, welches unabhängig von der Art und Weise, dem Ort und der Dauer des Trainings ist und damit die Möglichkeit einer Akkreditierung, sowohl für Experten als auch für Auszubildende, bietet.
- Einheitliche Qualifikationen ermöglichen eine genauere und flexiblere Gestaltung der Anerkennung von Kompetenzen, da verschiedene Kombinationen von Einheiten für die Zusammenstellung von Qualifikationen verwendet werden können (Oates, 2004, S. 56).

Auf der anderen Seite gibt Oates gegenüber diesen Annahmen zu bedenken, dass sie bei der Formulierung der National Vocational Qualifications ein zu sehr restringiertes Menschenbild nahegelegt haben.

Hinsichtlich des Kompetenzbegriffs stehen die Autoren vor den gleichen Schwierigkeiten, die wir einleitend herausgestellt haben: Dieser Begriff wird in unterschiedlichen Wissenschaftsdisziplinen und Praxisbereichen in den verschiedenen Ländern so vielfältig, schillernd und uneinheitlich gebraucht, dass man bei

seinem Gebrauch auf ausgearbeitete und theoretisch begründete Konzepte zurückgreifen muss, will man ihn in produktiver Absicht verwenden. Problematisch sind auch die zugrundegelegten, formalen Definitionen, die häufig auf einer sehr abstrakten Ebene, wie z. B. die für Deutschland herangezogene Definition der Kulturministerkonferenz (KMK) von 1996, angesiedelt sind und daher oft keine konkreten Operationalisierungen bereitstellen (vgl. KMK, 1996). Da Winterton, Delamare-LeDeist & Stringfellow keine eindeutigen Ziele für die berufliche Aus- und Weiterbildung vorgeben, haben sie Probleme, die vielfältigen, durchaus widersprüchlichen Auffassungen davon, was denn eigentlich unter „Kompetenz“ zu verstehen sei, zu sortieren und angemessen zu klassifizieren. Insbesondere trifft dieses für die Kategorien „Social and meta-competence“ zu (vgl. Abbildungen 2–2 und 2–3), in der Sozialkompetenz und allgemeine Fähigkeiten, die wir im Folgenden unter dem Terminus „Selbstkompetenz“ subsumieren, gemischt sind (vgl. hierzu die Ausführungen in Kapitel 2.4). Damit werden Antriebe und Reflexionen, die auch für die Bereiche der „Cognitive Competence“ und der „Functional Competence“ von Wichtigkeit sind, nicht hinreichend erfasst und operationalisiert.

Die Schwierigkeiten der Autoren, die Spaltenbezeichnungen stimmig zu fassen, deuten bereits auf Probleme der theoretischen Konsistenz und der dahinter stehenden Klarheit von Zielsetzungen für die berufliche Bildung hin. Die immer wieder vorgetragenen Hinweise auf eine Übereinstimmung mit entsprechenden Vorstellungen auf der EU-Ebene können nicht darüber hinwegtäuschen. Wir halten den prinzipiellen Rückgriff auf das Kompetenzkonzept ebenfalls für notwendig, sind aber der Auffassung, dass hierzu weitere Spezifikationen erforderlich sind, die auch die internationale Diskussion um die Operationalisierung von Wissensarten und psychischen Dispositionen, vor allem aber eine sowohl aus ökonomischer als auch aus pädagogischer Sicht stringente Fassung des Kompetenzkonzepts mit berücksichtigen.

Im Folgenden soll der Vorschlag von Winterton, Delamare-Le Deist & Stringfellow (2005) mit solchen aus der deutschen Literatur verglichen werden, die sich ebenfalls auf Fragen der Kompetenzdefinition und -messung beziehen. Die Hauptprobleme des Papiers von Winterton, Delamare-Le Deist & Stringfellow sind zum einen im Fehlen einer Zielbegründung für Maßnahmen der beruflichen Bildung – bzw. in einer unreflektierten Akzeptanz einer Anpassung des Individuums an Arbeitsplatzanforderungen – und zum anderen in der eher bruchstückhaften Begründung der Kategorienbildung zu sehen.

Abbildung 2–3: Ansätze zur Klassifizierung von Kompetenzen in der Berufsbildung in den EU-Mitgliedsstaaten

Land	Cognitive Competence (knowledge) Wissen	Functional Competence (skills) Fertigkeiten & Fähigkeiten	Social and Meta- Competence (Bahviours & Attitudes) Interaktion, Motivation, Einstellung, Metakognition
Österreich	Sachkompetenz		Personalkompetenz
		Sozialkompetenz	
Finland	Knowledge	Working Methods	Core competencies
	Safety		
Frankreich	Savoir	Savoir-faire	Savoir être
Deutschland	Fachkompetenz		Personalkompetenz
		Methodenkompetenz	
	Sachkompetenz	Sozialkompetenz	
Ungarn	Knowledge	Application of knowledge	Professional attitudes & behaviour
Ireland	Knowledge	Know-how and skill	Competence Role, Insight & Learning
	Competence – Context		
Niederlande	Profession-specific methods		Establishing & main- taining contacts, co- operation, teamwork
			Development of an individual team
	Professional competences for functioning in the context of work organisations		
Portugal	Competências cognitivas	Competências funcionais	Competências sociais
England & Wales	Underpinning knowledge	Functional competence	Social competence
Schottland	Knowledge & understanding	Practice (applied knowledge)	Autonomy, accountability, working with others
	Generic cognitive skills	Communication, Numeracy, IT skills	

Quelle: Winterton, Delamare-Le Deist & Stringfellow, 2005, S. 45–46.

In Deutschland haben Erpenbeck & von Rosenstiel (2003a; 2003b) ein „Handbuch Kompetenzmessung“ herausgegeben, bei dem das „Erkennen, Verstehen und Bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis“ im Mittelpunkt steht. Unter einer primär psychologischen Sichtweise geht es hierbei um den Abgleich einer

- *subjektzentrierten Eigenschaftsdiagnostik*: „Was für Eigenschaften hat die Person?“, einer
- *handlungsorientierten Verhaltensdiagnostik*: „Was tut die Person?“, sowie einer
- *Ergebnisdiagnostik*: „Welche Leistungen erbringt die Person?“ (2003b, S. XXIX).

Unterschieden wird, ob es sich bei den Ergebnissen um solche aus konvergent-anforderungsorientierten oder divergent-selbstorganisativen Handlungs- und Tätigkeitssituationen handelt. Die ersteren sind „direkt auf die Erfüllung äußerer Anforderungen, Vorgaben, Ziele gerichtet“, die zweiten sind als kreativ sowie als teilweise oder gänzlich ziel- und ergebnisoffen zu charakterisieren (S. XXVIII). Gemäß diesen Kategorisierungen werden dann unterschieden:

- *Qualifikationen* als „klar zu umreißende Komplexe von Kenntnissen, Fertigkeiten und Fähigkeiten, über die Personen bei der Ausübung beruflicher Tätigkeiten verfügen müssen“. Sie werden als *handlungsorientiert* und dabei als *konvergent-anforderungsorientiert* gefasst. Sie lassen sich für gewöhnlich eindeutig operationalisieren und sind von daher akkreditierungsfähig (S. XXIX).
- *Fähigkeiten* werden als „verfestigte Systeme verallgemeinerter psycho-physischer Handlungsprozesse“ gefasst und gelten damit als handlungszentriert. „Sie können sich gleichermaßen auf *konvergent-anforderungsorientierte* wie auf *divergent-selbstorganisative* Handlungssituationen beziehen“.
- *Kompetenzen* „bezeichnen ... Selbstorganisationsdispositionen psychischen und physischen Handelns, wobei unter Dispositionen die bis zu einem bestimmten Handlungszeitpunkt entwickelten inneren Voraussetzungen zur Regulation der Tätigkeit verstanden werden. Damit umfassen Dispositionen nicht nur individuelle Anlagen, sondern auch Entwicklungsergebnisse“. Sie „sind folglich eindeutig *handlungszentriert* und primär auf *divergent-selbstorganisative* Handlungssituationen bezogen“ (S. XXIX).

Diese Unterscheidung hat Konsequenzen für Messoperationen: Anhand der Messung von psychischen Eigenschaften kann man Indizien „für das Vorhandensein einer Selbstorganisationsdisposition und damit einer Kompetenz“ erheben – „ob sie tatsächlich vorhanden ist, lässt sich erst im Handlungszusammenhang beurteilen“ (S. XXIX).

Auf jeden Fall ist in diesem Zusammenhang auf die Bedeutung von Wissensbeständen und Fertigkeiten- wie Fähigkeitsbündeln zu verweisen. Auf Anregung der OECD, die ebenfalls versucht, das Kompetenzkonzept neu zu fassen, schlägt Weinert (2002, S. 27/28) vor, „unter Kompetenzen sind die bei Individuen verfü-

baren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten“ zu verstehen, „um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“.

In einer Erläuterung nennt Weinert dann acht Kompetenzen – nach der Häufigkeit ihrer Nennung geordnet: Beherrschung der Muttersprache (mündlich und schriftlich), mathematisches Wissen; Lesekompetenz; Beherrschung wenigstens einer Fremdsprache; Medienkompetenz; Lernstrategien; soziale Kompetenzen; divergentes, kritisches Denken. Diese Liste entspricht in hohem Maße dem „Lehrplan des Abendlandes“ (Dolch, 1965), der auf den „Königsweg“ der Bildung, nämlich den Besuch allgemeiner Schulen bis hin zur Universität bezogen ist (vgl. hierzu auch Achtenhagen, 1998). Der Bezug zu beruflichen Bildungsprozessen ist nicht explizit ausgearbeitet, die Erweiterung um spezifische Fachkompetenzen erscheint uns aber möglich. Weinerts Ausführungen sind für unsere Überlegungen wichtig, weil er die Interdependenzen der Kompetenzkomponenten hervorhebt und dabei auch die Bedeutung einer Einbeziehung von Motiven und Einstellungen sowie metakognitiven Fähigkeiten betont.

Diese Sichtweise ist auch für die Diskussion um notwendige Entwicklungen der Berufsbildung wichtig. Oates (2004, S. 59) beklagt, dass bei der Durchsetzung der NVQ in England und Wales zu schnell der „formation“-Aspekt durch den „competence“-Aspekt verdrängt wurde.

Aus den Ausführungen von Oates wird deutlich, dass ein rigide eingegrenztes Kompetenzkonzept wünschenswerte Aspekte der Berufsbildung, wie sie über unsere grundlegende Zielsetzung zum Ausdruck kommen, nicht hinreichend berücksichtigt. Hier liegt auch der Grund dafür, dass englische Experten – so etwa A. Brown und F. Coffield auf dem zweiten Workshop – auf den pejorativen Gebrauch von „competence“ im Englischen verwiesen haben. Für unseren Vorschlag haben wir aus dieser Diskussion die Konsequenz gezogen, auf ein breiter angelegtes, pädagogisch-anthropologisch begründetes Kompetenzkonzept zurückzugreifen, das über den Bereich der Selbstkompetenz auch individuelle Einstellungen und Motive einbezieht.

In der deutschen Expertise „Zur Entwicklung nationaler Bildungsstandards“ werden Kompetenzen schulbezogen gefasst: Sie beschreiben „solche Fähigkeiten der Subjekte, die auch der Bildungsbegriff gemeint und unterstellt hatte: Erworben, also nicht von Natur aus gegebene Fähigkeiten, die an und in bestimmten Dimensionen der gesellschaftlichen Wirklichkeit erfahren wurden und zu ihrer Gestaltung geeignet sind, Fähigkeiten zudem, die der lebenslangen Kultivierung, Steigerung und Verfeinerung zugänglich sind, so, dass sie sich intern graduieren lassen, z. B. von der grundlegenden zur erweiterten Allgemeinbildung; aber auch Fähigkeiten, die einen Prozess des Selbstlernens eröffnen, weil man auf Fähigkeiten zielt, die nicht allein aufgaben- und prozessgebunden erworben werden, sondern ablösbar von der Ursprungssituation, zukunftsfähig und problemoffen“

(Klieme et al., 2003, S. 65). Dass eine Ausrichtung auch auf den berufsbildenden Bereich erforderlich ist, wird nicht bestritten. Beziehungen zur Berufsbildung aber, die in Deutschland ja ebenfalls zur Pflichtschulbildung gehört, werden nicht hergestellt. Allerdings hätte eine Verbesserung der Leistungen zum Abschluss der Sekundarstufe I sicher wesentliche Erleichterungen für die Lehr-Lernprozesse im beruflichen Bereich im Gefolge. Für unsere Überlegungen ist wichtig, dass ein Berufsbildungs-PISA von vergleichbaren Überlegungen zur Persönlichkeitsförderung auszugehen hat, wie das für die Allgemeinbildung postuliert wird.

Abbildung 2–4: Vergleich des englischen „formation“ – und „competence“-Konzepts

Formation	Competence
formative assessment to provide feedback for learners in order for them to enhance their performance	summative assessment to provide a transparent statement of what someone can do, for certification, selection, etc.
open, unstructured information on performance which reflects the specific composition and meaning of achievement, motivation, etc. of an individual	tightly structured, centralised information on performance which provides a common language for describing competence
emphasis on inputs relating to effective, structured use of learning processes and settings, with assumptions that processes of learning affect the nature of emerging competence	emphasis on outcomes with an emphasis on inferring competence from performance, on leaving open the mode, duration, and location of learning

Quelle: Oates, (2004), S. 59.

Die hier dargestellten drei deutschen Positionen reichen von einer primär psychologischen Sicht bei Erpenbeck & von Rosenstiel bis hin zu einer mehr inhaltsbezogenen Argumentation der Expertenkommission für die Entwicklung von Bildungsstandards, die zur Definition von Outputs für die Fächer Deutsch, Mathematik und erste Fremdsprache am Ende der Sekundarstufe I überleitet. Auf Weinerts Position beziehen sich beide Autorengruppen.

Gegenüber dem Ansatz von Winterton, Delamare-Le Deist & Stringfellow ist die Spezifikation der Wissensdimension bei den deutschen Ansätzen hervorzuheben, die bei Erpenbeck & von Rosenstiel mit ihrer psychologischen Ausrichtung nur schwach durchscheint, jedoch bei Weinert, vor allem aber bei der Expertenkommission sehr viel stärker durchschlägt. Winterton, Delamare-Le Deist & Stringfellow heben zwar auf verschiedene Inhaltsbereiche ab, bleiben aber mit ihren Hinweisen auf Branchen oder Berufsfelder eher blass. An genau dieser Stel-

le meinen wir, die Erfassung einer Kompetenzentwicklung im beruflichen Bereich ansetzen zu müssen – im Sinn einer Output-Orientierung.

2.2. SCHWIERIGKEITEN DER OPERATIONALISIERUNG UND MESSUNG VON KOMPETENZEN

Betrachtet man die Bestandteile verschiedener Kompetenzdefinitionen, können, unabhängig von linguistischen Differenzen, bestimmte gemeinsame Elemente als Bestandteile vieler Definitionen identifiziert werden: Die Anwendung/der Einsatz eines individuellen Ensembles von (Performanz-)Voraussetzungen (Wissen, Fähigkeiten, Fertigkeiten, Strategien, Routinen, Emotionen, Einstellungen, Motivationen, Bereitschaften etc.), um bestimmte Probleme/Aufgaben/Situationen in verschiedenen Kontexten zu bewältigen/zu lösen.

Der „Kompetenzwirrwarr“ zeigt sich in der Vielfalt individueller Voraussetzungen und führt zu einer Vermischung der Ebenen, vor allem wenn es um berufliche Spezifität geht (vgl. z. B. Oates, 2004). Die Vielfältigkeit nicht-beobachtbarer individueller Bedingungen, die in verschiedenen Situationen zur Anwendung kommen können, wurde auf den Workshops deutlich. Abbildung 2–5 zeigt exemplarisch die auf dem Workshop diskutierten Bereiche interner Bedingungen.

Im Rahmen einer internationalen Vergleichsstudie, die die drei einleitend beschriebenen übergeordneten Ziele der Berufsbildung – Entwicklung individueller beruflicher Regulationsfähigkeit (personale Autonomie); Sicherung der Humanressourcen einer Gesellschaft; Gewährleistung gesellschaftlicher Teilhabe und Chancengleichheit – zugrunde legt, ist eine der größten Herausforderungen die Konkretisierung ausgewählter Kompetenzelemente im Hinblick auf diese Ziele. Insbesondere auf dem zweiten Workshop wurde betont, dass man zwar viele Kompetenzelemente auflisten könnte, es aber viel wichtiger sei, sich bei der Komplexitätsreduktion auf eine Auswahl von Elementen zu konzentrieren, die man in Relation zu den drei Zielen der Berufsbildung setzen kann. Damit geht der vorgeschlagene Ansatz deutlich über eine pragmatisch orientierte Eignungsdiagnostik hinaus.

Anhand der Unterscheidung zwischen nicht-beobachtbaren individuellen Voraussetzungen und beobachtbarem Verhalten in verschiedenen Situationen wird deutlich, dass zwei entgegengesetzte Betrachtungsweisen oder Ansätze der Kompetenzmessung unterschieden werden können. Kompetenz als Potential für das Verhalten in verschiedenen Situationen bezieht sich auf interne Bedingungen (z. B. Wissen, Können, Motivation, Werte, Metakognition), die für Außenstehende nicht beobachtbar sind, während kompetentes Verhalten in einer Situation der Beobachtung durch Dritte zugänglich ist. Zur Operationalisierung und Messung von beruflichen Kompetenzen kann daher – wie in Abbildung 2–5 veranschaulicht – zwischen den beiden folgenden Möglichkeiten unterschieden werden:

Abbildung 2–5: Alternative Operationalisierungs- und Messungswege von beruflichen Kompetenzen

Interne Voraussetzungen/individuelle Fähigkeiten Person	Bedingungen /externe Performanz Verhalten Situation
Deklaratives Wissen Bestimmte Prozeduren Allgemeine Prozeduren Metakognitive Fähigkeiten Strategien & Fertigkeiten Werte Einstellungen Motivation Selbstwirksamkeit, Selbstbeherrschung Angstgefühl Big Five*	Performanz Lösung von Arbeitsaufgaben in verschiedenen Situationen
nicht direkt beobachtbar	beobachtbar

* Fünf unabhängige Faktoren zur Beschreibung und Messung der Persönlichkeit (Extraversion, Soziale Verträglichkeit, Gewissenhaftigkeit, Neurotizismus, Intellekt)

- Bestimmung *externer Tätigkeiten*: Situationen bzw. Aufgaben und Anforderungen, die im derzeitigen und vermutlich zukünftigen Beruf/Berufsfeld international charakteristisch sind;
- Bestimmung *interner Bedingungen*: z. B. Wissen, Können, Motivation, Werte, Metakognition etc., die für die Expertise in einem bestimmten Beruf/Berufsfeld kennzeichnend sind.

Diese Alternativen sollen unter dem Aspekt ihrer Bedeutung für ein Berufsbildungs-PISA im Hinblick auf ihre jeweiligen Vor- und Nachteile erörtert werden. Lang-von Wins (2004, S. 587) hat darauf aufmerksam gemacht, dass sich bei dieser Betrachtungsweise der „wissenschaftlich-theoretische Anspruch einer hohen Abstraktion und der anwendungsorientierte Aspekt eines möglichst hohen Pragmatismus“ gegenüberstünden.

2.2.1. Ermittlung von Kompetenzen auf der Grundlage von externen Tätigkeiten

Bei diesem Ansatz der Kompetenzerfassung sind typische, international vergleichbare berufliche Tätigkeiten und Aufgaben zu ermitteln, die für bestimmte Berufsfelder kennzeichnend sind. Eine Möglichkeit wäre, Schnittmengen für ausgewählte Berufsbilder anhand formaler Beschreibungen von Situationen und Anforderungen zu bestimmen. Beispielsweise ließen sich kompetenzorientierte Anforderungsbeschreibungen von angestrebten Outputs wie bei den National Vocational Qualifications (NVQs) in England zugrundelegen. Preiß (2004) hat einen Vergleich der Strukturen zur Beschreibung von Kompetenzen im kaufmännischen Bereich für Deutschland, England und Frankreich durchgeführt, Fulst-Bleil & Ebner (2005) für Deutschland und Wales. Dieses Vorgehen, das man als ersten Schritt für das Verfahren der Ermittlung von Kompetenzen auf der Basis externer Tätigkeiten betrachten kann, ist allerdings mit erheblichen Problemen verbunden: Zum einen existieren linguistische Unterschiede hinsichtlich bestimmter Begriffe (Woher gewinnt man die Sicherheit, dass mit einer bestimmten Tätigkeitsbeschreibung in den verglichenen Ländern tatsächlich Identisches gemeint ist?); darüber hinaus ist es möglich, dass erlernte Inhalte später auf dem Arbeitsmarkt gar nicht benötigt werden, und schließlich könnten verschiedene Abschnitte der Ausbildung (Erstausbildung, Weiterbildung) betroffen sein. Entscheidend ist auch, ob ein eher modulares oder ein eher ganzheitliches Verständnis der Berufsausbildung zugrunde gelegt wird. Dieser Punkt betrifft vor allem die Ganzheitlichkeit des Verstehens beruflicher Arbeitszusammenhänge.

Die Probleme verschärfen sich, je mehr Länder in einen Vergleich einbezogen werden. Selbst wenn man von weltweit gültigen gemeinsamen Anforderungskernen für bestimmte Berufsbilder (z. B. IT, Hotelberufe, Kfz-Mechatroniker) ausginge, wäre im Einzelfall genau zu prüfen, ob dabei nicht zu sehr von nationalen Besonderheiten in der Definition des jeweiligen Tätigkeitszuschnittes abgesehen würde. Selbst beispielsweise für den Kfz-Bereich, für den große Übereinstimmungen behauptet werden, lassen sich auf der nationalen Ebene erhebliche Unterschiede ausmachen.

Der Vorteil des Verfahrens der Kompetenzbestimmung über externe Tätigkeiten liegt im Vergleich zur Entwicklung von Konzepten zur Bestimmung interner Bedingungen, z. B. „Technical Literacy“, in der kürzeren Entwicklungsdauer. Dieser Vorteil wurde von einigen Teilnehmern des 2. Workshops betont: Um die zeitintensive Entwicklungsarbeit zu vermeiden, sollte man zunächst von typischen, international vergleichbaren beruflichen Tätigkeiten auszugehen und diese anschließend in den einzelnen Ländern empirisch überprüfen. Tim Oates (England) macht allerdings die Schwierigkeit dieser Vorgehensweise sehr deutlich und weist auf die Gefahr hin, die Komplexität von Kompetenz zu verkennen, und betont, dass Kompetenz durch Beobachtung der Performanz von Individuen in verschiedenen Situationen abgeleitet werden sollte, um diesen Fehler zu vermeiden (vgl. Oates, 2004, S. 60).

Eine konkretere Variante des externen Verfahrens im Vergleich zu Berufs- und Tätigkeitsprofilen liegt in Arbeits- und Anforderungsanalysen, für die allerdings hervorzuheben ist, dass sie relativ zeit- und arbeitsaufwändig sind. Im Allgemeinen wird ein dreistufiges Vorgehen gewählt (Lang-von Wins, 2004, S. 591):

- „Analyse der Auftrags- und Erfüllungsbedingungen einer Arbeitstätigkeit,
- Identifikation und Beschreibung von Tätigkeitsdimensionen,
- Analyse der Tätigkeit anhand dieser Dimensionen, mit dem Ziel, leistungs- und qualifizierungsdifferenzierende Variablen zu finden“.

Mit diesem Vorgehen hängt auch das Verfahren der Arbeitsproben zusammen, die „der Logik einer kontrollierten Nachbildung erfolgskritischer Situationen“ folgen, „innerhalb derer das Verhalten der zu beurteilenden Person beobachtet und bewertet werden kann“ (ebenda S. 597). Hervorgehoben wird dabei der hohe Anspruch an eine angemessene Konstruktion solcher Arbeitsproben – woraus auch eine Befürwortung von simulationsorientierten Prognosekonzepten in Kombination mit biographischen Methoden resultiert (ebenda S. 613; vgl. auch die Messverfahren in Erpenbeck & von Rosenstiel, 2003a, die Beispiele in Arbeitsgemeinschaft Betriebliche Weiterbildungsforschung, 2005 sowie der Bericht zu PISA-Deutschland: Wirth & Klieme, 2003).

Die am weitesten elaborierte und konkreteste Variante zur Ermittlung von Kompetenzen, die auf die Zusammenfassung ähnlicher Arbeitstätigkeiten bzw. Verhaltensweisen, die der Erfüllung der wichtigsten Arbeitsaufgaben zugrunde liegen, hin ausgerichtet ist, wurde auf den internationalen Workshops von Wim Nijhof (Niederlande) ins Spiel gebracht. Das zugrundegelegte Instrument stammt aus dem Kontext einer arbeitsmarktbezogenen (O*NET) Datenbank aus den USA (Jeanneret & Boreman, 2002). Zentraler Bestandteil der O*NET-Datenbank sind Zusammenfassungen vergleichbarer Aktivitäten bzw. Verhaltensweisen, die der Erfüllung der wichtigsten Arbeitsaufgaben zugrunde liegen und als „Generalized Work Activities (GWA)“ bezeichnet werden. Eine Taxonomie mit 42 Konstrukten stellt die Grundlage zur Beurteilung der Anforderungen an Arbeitsaktivitäten für die Mehrheit der Berufsfelder weltweit dar. Die Datenbasis beruht auf der Annahme, dass Arbeitsverhalten nicht notwendigerweise mit spezifischen Aufgaben und Techniken verbunden sein muss und auf einer abstrakten Ebene zu „GWA“ zusammengefasst werden kann. Das bedeutet nicht, dass sich alle „GWA“ auf der gleichen Abstraktionsebene befinden müssen; es kann auch Abstufungen geben, die aber niemals so speziell sind, dass sie nur mit bestimmten einzelnen Berufen verbunden werden können.

Der Entwicklung von „Generalized Work Activities (GWA)“ geht eine lange Forschungstradition voraus. In einer Vielzahl von Ansätzen wurde bereits versucht, die Dimensionalität und Struktur von Arbeit auf der Grundlage von Verhalten zu identifizieren. Jeanneret (1969) untersucht beispielsweise die Hypothese, dass Strukturen der menschlichen Arbeit sich in Form von Dimensionen identifizieren lassen. „GWA“ sind definiert als eine Zusammenfassung von ähnlichen Arbeitstätigkeiten/Verhaltensformen („Activities“), die der Erfüllung bestimmter

Hauptarbeitsfunktionen zugrunde liegen. Sie basieren auf drei Prinzipien: sie müssen erstens breit genug sein, um für möglichst viele Jobs angewendet werden zu können; sie müssen zweitens auf analytischen Arbeitsforschungen beruhen und drittens charakteristisch für die unterliegende Arbeitsstruktur sein.⁷

Toolsema (2003) setzt „GWA“ als Instrument zur Ermittlung von Kompetenzen ein, indem er Arbeitstätigkeiten („Activities“) als Kompetenzindikatoren auf einer höheren Abstraktionsebene, die den Zweck und das Ziel der Kompetenzen angeben, interpretiert. Auf der Grundlage eines Vergleichs von verschiedenen Studien differenziert er sechs Kompetenzbereiche: soziale, partizipative, kognitive, physische/technische Fähigkeit, Lernfähigkeit und Arbeitsfähigkeit, und verknüpft diese mit den „GWA“. Das Instrument ist für Messungen kognitiver, affektiver und meta-kognitiver Aspekte geeignet, und ermöglicht, unterschiedliche Ebenen von Allgemeinheit abzudecken, womit internationale Vergleichbarkeit gewährleistet werden könnte.

Der Vorteil dieses Ansatzes wird darin gesehen, dass über die Bestimmung von externen Situationen anhand von Arbeitstätigkeiten Kompetenzen auf einem berufsübergreifenden Abstraktionsniveau („Working Activities“) für ausgewählte Berufsfelder abgeleitet werden können. Dem entspricht das forschungspragmatische Argument, dass Individuen am Arbeitsplatz nicht alle Kompetenzen benötigen, die sie in formellen Lernprozessen gelernt haben, weswegen es auch nicht notwendig sei, die Gesamtheit der angestrebten Kompetenzen zu erfassen. Im Hinblick auf eine Komplexitätsreduktion eines Berufsbildungs-PISA könnte daraus folgen, sich auf diejenigen Kompetenzen zu konzentrieren, die am Arbeitsplatz wirklich benötigt würden (so vor allem Wim Nijhof auf den internationalen Workshops). Gegen ein solches Vorgehen lassen sich zwei Einwände vorbringen: Zum einen besteht die Gefahr einer zu eng an konkreten Arbeitsverrichtungen ausgerichteten Definition von beruflicher Kompetenz. Zum anderen ist forschungsmethodisch ein Zirkularitätsproblem gegeben: Wie sollen ohne Arbeits- und Anforderungsanalysen die „wirklich benötigten“ Kompetenzen, die dann als Maßstab gelten würden, zuvor erhoben werden. Es ist daher zu fragen, ob ein solcher Ansatz nicht auch sorgfältige – und zusätzlich zeitaufwändige – Beobachtungen erfordert, mit deren Hilfe direkt Performanz erfasst wird – oder ob die Messung über Simulationen und Tests, ergänzt durch biographische Auskünfte, eher indirekt erfolgen sollte.

Für ein Berufsbildungs-PISA muss gründlich überlegt werden, ob eine Beschränkung auf Performanzmerkmale für eine Kompetenzmessung methodisch sinnvoll und vertretbar ist oder ob in die zeitintensivere Entwicklungsarbeit von Konzepten zur Bestimmung interner Bedingungen investiert werden sollte.

7 Eine ausführliche Darstellung der Struktur sowie Beispiele für „GWA“ befinden sich im Anhang.

2.2.2. Ermittlung von Kompetenzen auf der Grundlage interner Bedingungen

Kompetenzmessung auf der Grundlage der jeweils individuellen Potentiale bzw. internen Handlungsvoraussetzungen verläuft im large-scale assessment über Testverfahren. Für ein Berufsbildungs-PISA sind vergleichbar mit den Konzeptualisierungen in den internationalen Schulleistungsuntersuchungen wie PISA und TIMSS (z. B. „Literacy“, „Numeracy“) ebenfalls international akzeptierte Konzeptualisierungen zur Struktur und Entwicklung beruflichen Wissens und Könnens in den verschiedenen Wissens- und Anwendungsdomänen notwendig. Bisherige Konzepte müssten daher zunächst daraufhin überprüft werden, ob sie für die Erfassung von Expertise in der Berufsbildung geeignet sind bzw. zu modifizieren oder weiterzuentwickeln sind. Des Weiteren muss entschieden werden, welche Kompetenzbereiche erfasst und wie Niveaustufen der Kompetenz bestimmt werden können. Im folgenden Abschnitt soll die Struktur von bestehenden Konzepten am Beispiel der in den USA entwickelten und in Deutschland übernommenen und modifizierten Standards für technische Bildung („Standards for Technological Literacy“) dargestellt werden (vergleiche hierzu auch die Stellungnahme des VDI (Verein Deutscher Ingenieure) in den vorgelegten „Bildungsstandards im Fach Technik für den mittleren Schulabschluss“ (vgl. Verein Deutscher Ingenieure, 2004).

Die amerikanischen Standards sind für die allgemeine technische Bildung vom 2. bis zum 12. Schuljahr konzipiert. Insgesamt gibt es 20 Standards, denen jeweils grundlegende Lerninhalte zugeordnet werden. Es werden zwei Klassen von Standards unterschieden, die Schüler beherrschen sollten:

- Kognitive Standards: Wie funktioniert Technik? Welchen Stellenwert nimmt sie in der heutigen Welt ein?
- Prozess-Standards: Welche Fähigkeiten sollten Schüler erlangen, um sich in technisch gefragten Umwelten bewegen zu können.

Im Hinblick auf diese Standards werden fünf Hauptkategorien hervorgehoben:

- Charakter der Technik,
- Technik und Gesellschaft,
- Konstruktion und Gestaltung technischer Produkte,
- Fähigkeiten für das Leben in einer technischen Welt und
- die technische Welt.

Die Standards sind folgendermaßen strukturiert: Zunächst wird der jeweilige Standard in einem Satz formuliert; daraufhin wird die Intention des Standards erläutert; anschließend werden Materialien für die jeweilige Schulstufe vorgestellt, anhand derer die Standards erläutert sowie Vorschläge gemacht werden, wie sie sich im Technikraum umsetzen lassen. Die Ausführungen zu jeder Schulstufe werden ergänzt durch eine Reihe von Leistungskriterien, welche zugleich die besonderen Kenntnisse und Fähigkeiten zu ihrer Erfüllung auflisten. Jedes Kriterium wird weiter spezifiziert.

Für den kaufmännischen Bereich sind als vergleichbares Konzept die in den USA entwickelten Standards für wirtschaftliche Bildung („National Content Standards in Economics“) für die Jahrgangsstufen 4–12 zu nennen (vgl. National Council on Economic Education, 1997). Insgesamt gibt es 20 inhaltsbezogene Standards, mit denen einerseits Schlüsselkompetenzen („Key Skills“) und andererseits Wissenskompetenz erreicht werden sollen. Für das Ende des 12. Schuljahrs werden fünf Schlüsselkompetenzen angestrebt: Wirtschaftliche Probleme, Alternativen, Gewinne und Kosten erkennen; Anreize der Arbeit in einer wirtschaftlichen Situation analysieren; Konsequenzen von Veränderungen wirtschaftlicher und politischer Bedingungen untersuchen; Sammeln und Ordnen wirtschaftlicher Hinweise; Gewinne mit Verlusten vergleichen. Hinzu kommen drei verschiedene wirtschaftliche Wissensbereiche:

- Verständnis von grundlegenden wirtschaftlichen Konzepten sowie die Fähigkeit der logischen Schlussfolgerung bei wirtschaftlichen Fragen, die das Leben als Arbeiter, Konsument und Staatsbürger betreffen, damit Fehler vermieden werden können, wie sie bei Personen auftreten, die wirtschaftliche Fragestellungen nicht verstehen;
- Wissen über einschlägige Fakten der amerikanischen Wirtschaft (z. B. Arbeitslosenrate, Inflation, Zinsrate); Wissen darüber, dass Wirtschaftswissenschaftler verschiedene, aber auch gemeinsame Ansichten haben können.
- Für die Schulstufen 4–12 werden die folgenden 20 Wissensbereiche mit den jeweiligen Benchmarks für die Schulstufen 4, 8 und 12 bereitgestellt: Knappheit der Güter; Grenzkosten, -gewinn; Verteilung von Produkten und Dienstleistungen; Funktion von Anreizsystemen; Ertrag durch Tausch; Spezialisierung und Tausch; Märkte, Preise und Quantitätsdeterminanten; Funktion von Preisen auf Märkten; Funktion von Wettbewerb; Funktion von wirtschaftlichen Institutionen; Funktion des Geldes; Funktion von Zinsraten; Funktion von Ressourcen zur Bestimmung von Einkommen; Gewinn und der Unternehmer; Wachstum; Funktion der Regierung; Evaluation von Regierungsprogrammen mit Hilfe der Kosten und Verlustrechnung; Makroökonomie des Einkommens, der Beschäftigung, der Preise; Arbeitslosigkeit und Inflation; Geld und Finanzpolitik.

An der Auflistung der amerikanischen Standards für zwei – freilich sehr umfangreiche – berufliche Wissensbereiche mag deutlich geworden sein, wie viel Arbeit und Aufwand bei einem Berufsbildungs-PISA zum einen in die internationale Konsensbildung über Standards, zum anderen in die Entwicklung von Indikatoren und Verfahren zur Messung der entsprechenden Kompetenzen zu investieren ist.

Wie bereits in der Einleitung erwähnt, erhöht sich die Komplexität bei einem Berufsbildungs-PISA durch den Stellenwert von funktionalen Kompetenzaspekten und durch die Komponente des Arbeitsmarktverhaltens. Im Hinblick auf die spätere Bewährung auf dem Arbeitsmarkt und die Notwendigkeit eines lebenslangen Lernens kommen Probleme der Veränderung und Relevanz von den in Ausbildungsprozessen erlernten Wissens- und Fähigkeitsstrukturen hinzu. Bei Eintritt in

das Arbeitsleben und der Entwicklung beruflicher Expertise verändern sich die Struktur der erlernten Ausbildungsinhalte. Eine entscheidende Rolle spielen ebenfalls der informelle Erwerb von Wissen und Fähigkeiten, aber auch Einstellungen und Motive. In Kapitel 2.5. stellen wir ein entsprechendes Modell dar.

2.3. KOMPETENZEN IM SPANNUNGSVERHÄLTNIS VON NATIONALER SPEZIFIK UND INTERNATIONALER VERGLEICHBARKEIT

Neben den Möglichkeiten zur Ermittlung von Kompetenzen soll nun die Problematik des Spannungsfeldes einer internationalen Vergleichsstudie zwischen den Polen berufsspezifisch und berufsübergreifend dargestellt werden. Die Komplexität und Vielschichtigkeit dieser Ebenen werden im folgenden Abschnitt anhand der Beiträge aus dem Workshop sowie einiger Beispiele aus der Literatur erläutert, um einerseits die Problematik, aber auch das Ziel dieser internationalen Vergleichsstudie weiter zu veranschaulichen.

In den Diskussionen der Workshops wurde die Schwierigkeit, einen Konsens über die Aggregationsebene einer Vergleichsstudie zu finden, mehrfach deutlich: Je spezifischer bestimmte berufliche Tätigkeiten definiert sind, desto schwieriger ist die internationale Vergleichbarkeit. Einigkeit bestand bei den internationalen Experten darüber, dass ein geeigneter Ansatz einerseits in der Lage sein muss, unterschiedliche berufliche Domänen voneinander abzugrenzen, und andererseits internationale Vergleichbarkeit zu gewährleisten: Während der Erfolg eines Verkäufers z. B. maßgeblich von seinen Fähigkeiten des Anpreisens seiner Ware abhängen mag, erklärt sich der Erfolg der Kindergärtnerin möglicherweise vor allem über die eigene Durchsetzungsfähigkeit (Kanning, 2003, S. 18). Da es zudem auch noch national unterschiedliche Ausprägungen von Kompetenzfacetten und Situationsanforderungen gibt, müsste ein Abstraktionsniveau oberhalb der jeweiligen nationalen Ebene gewährleistet sein.

Auf der Basis von verschiedenen Techniken von Arbeitsanalysen macht Oates (2004) die Vielfältigkeit der Aggregationsebenen bei berufsbezogenen Kompetenzbeschreibungen deutlich. Er unterscheidet fünf verschiedene Beschreibungsebenen:

- „*Generic*“ beschreibt transferfähige Kompetenzen auf einer allgemeinen Ebene („*Key Skills*“), die nicht auf das Verhalten in speziellen beruflichen Kontexten bezogen sind;
- „*Occupational*“ bezieht sich auf Beschreibungen auf der Berufsebene auf einem hohen Abstraktionsniveau mit der Intention, für möglichst viele Arbeitskontexte anwendbar zu sein;
- „*Task-specific but independent of specific jobs*“ bezieht sich auf Beschreibungen von bestimmten Aufgaben (z. B. Dachdecken), ohne die Art und Weise, wie die Aufgabe speziell organisiert werden könnte, näher zu beschreiben.

Bei der Entwicklung von Nationalen Standards (wie z. B. die NVQs in England) wird dieses Level häufig zugrunde gelegt;

- „*Job-specific, enterprise-specific*“ bezieht sich auf Beschreibungen der Art und Weise, wie eine bestimmte Aufgabe in einem bestimmten System der Arbeitsorganisation ausgeführt wird. Häufig fallen in diese Kategorie Arbeitsplatzbeschreibungen, Arbeitsprozessanalysen etc.;
- „*Person-specific*“ bezieht sich auf Beschreibungen der Art und Weise, wie Aufgaben/Arbeitstätigkeiten von Individuen in einem bestimmten System ausgeführt werden. Diese Beschreibungen beruhen auf Selbstbeurteilungsverfahren und eignen sich für die Beurteilung und Messung von Ausbildungsprozessen in der Wahrnehmung von Personen (vgl. Oates, 2004, S. 62).

Die Vielschichtigkeit von möglichen Beschreibungen macht deutlich, wie wichtig eine grundlegende Klärung der Ebenen im Hinblick auf die beabsichtigte Zielsetzung ist. Würde man beispielsweise arbeitsplatzspezifische („job-specific“) Beschreibungen für den Zweck der Entwicklung nationaler Standards einsetzen, hätte man partielle Beschreibungen, die in den einzelnen Branchen nicht übergreifend einsetzbar wären. Hinsichtlich einer internationalen Vergleichsstudie wie der hier diskutierten muss eine Ebene gefunden werden, die handhabbar und aussagefähig ist.

In der Literatur lässt sich eine Vielzahl von Ansätzen – sowohl auf der Grundlage von externen Tätigkeiten als auch von internen Bedingungen – in diesem Spannungsfeld finden. Der „Job-competency approach“ von Spencer und Spencer (1993) ist ein Beispiel für einen spezifischen, vor allem auf externen Tätigkeiten beruhenden Forschungsansatz aus den USA. Es werden Individuen am Arbeitsplatz analysiert, ohne dass vorher Annahmen über interne Bedingungen gemacht würden. Mit Hilfe sogenannter „open-ended behavioural event interviews“ wird bestimmt, welche Eigenschaften mit Erfolg am Arbeitsplatz assoziiert werden. Kompetenz wird hier mit denjenigen internen individuellen Eigenschaften assoziiert, die mit einer erfolgreichen Performanz (auf einer Skala von ausreichend bis überdurchschnittlich) am Arbeitsplatz verbunden sind. Ein ähnliches Beispiel ist die Ermittlung von Kompetenzen am Arbeitsplatz, wie z. B. beim Project Joint-Service Job Performance Measurement Standards (JMP), wo Performanz beim Militärpersonal in den USA bewertet wird (Wigdor & Green, 1991). Problematisch bei dieser Art von Ansätzen ist, dass sie für einen internationalen Vergleich zu speziell sind. Wenn national spezifische Annahmen über erfolgreiche Performanz gemacht und damit bestimmte Standards festgelegt werden, wird ein Vergleich mit anderen nationalen Standards erschwert.

Der andere Pol von berufsübergreifenden allgemeinen arbeitsbezogenen Kompetenzen wird in der Literatur am häufigsten mit dem Begriff „generic skills“ betitelt. Oft werden diese allgemeinen, berufsübergreifenden Kompetenzen auch als „enabling skills“, „core skills“, „key competencies“, „essential skills“, „necessary skills“ etc. bezeichnet (vgl. Nijhof, 1998; Murray, Clermont & Binkley, 2004). Generell zielen diese Ansätze auf die Entwicklung von transferfähigen

Kompetenzen, die für eine Vielzahl von Arbeitsaufgaben benötigt werden, sich aber von arbeitsplatzspezifischen Kompetenzen unterscheiden. Sie sind als eine Reaktion auf die ausufernde Entwicklung eng definierter und spezifischer Kompetenzen entstanden. In Deutschland hat sich diese Entwicklung in der Debatte um „Schlüsselqualifikationen“ (Mertens, 1974) niedergeschlagen, welche die Basis für eine allgemeine berufliche Handlungskompetenz abgeben, die dann in die Neuordnungsverfahren der Berufsbilder im dualen Ausbildungssystem in Deutschland Eingang gefunden hat.

In der Literatur findet sich eine Vielzahl generischer Kompetenzansätze, die wiederum hinsichtlich ihrer Abstraktionsebene variieren. Die Ausuferung des Schlüsselqualifikationskonzepts – so wurden über 650 verschiedene Vorschläge aufgelistet – ist mit einer der Gründe dafür, dass die Kompetenzdebatte in Deutschland so intensiv geführt wird.

Auf internationaler Ebene wurde eine große Anzahl von Projekten zur Ermittlung von beschäftigungsfähigen Kompetenzen („employability skills“) von bildungs- und arbeitsmarktbezogenen Organisationen durchgeführt. Die meisten dieser Untersuchungen gingen zunächst von breiten Definitionen aus und befragten oder beobachteten anschließend Arbeiter, Vorgesetzte oder Experten, um gemeinsame Kompetenzen für bestimmte Tätigkeiten herauszufinden. Ein Beispiel für ein solches Projekt ist die vom U.S. Department of Labour durchgeführte Studie SCANS („Secretary’s Commission on Achieving Necessary Skills“), in der zwei übergreifende Kategorien identifiziert wurden: Die Basiskompetenzen („foundation skills“) beinhalten grundlegende allgemeinbildende Kompetenzen (Lesen, Schreiben, Rechnen), Kompetenzen des logischen Denkens (Schlussfolgern, Problemlösen), Persönlichkeitseigenschaften (Verantwortung, Selbstschätzung) und arbeitsbezogene Kompetenzen, welche die Fähigkeiten bezeichnen, Ressourcen einzusetzen, zu interagieren, Informationen zu beschaffen und zu verarbeiten sowie mit Systemen und Technologien umzugehen (Secretary’s Commission on Achieving Necessary Skills, 1991).

Die allgemeinen berufsübergreifenden Ansätze sind von vielen Seiten und aus unterschiedlichen Gründen kritisch hinterfragt worden. Weinert (1999) gibt zu bedenken, dass manche Kompetenzen zwar allgemeingültig für eine Vielzahl von Lern- und Arbeitssituationen seien, aber keinen Ersatz für berufsbranchen-spezifische Kompetenzen darstellten – was ihnen freilich auch niemand unterstellt. Andere Autoren bestreiten, dass individuelle Kompetenzen kontext-unabhängig gemessen werden könnten (vgl. Sandberg, 2000). Der konstruktivistische Ansatz (vgl. z. B. Stoof, 2005) nimmt an, dass die Angemessenheit einer Kompetenzdefinition erst im Zusammenhang mit dem jeweiligen Kontext beurteilt werden könnte. In einer Studie von Stasz, Ramsey, Eden, Melamid & Kaganoff (1996) wurden Variationen und Gemeinsamkeiten von generischen Kompetenzen am Arbeitsplatz untersucht und dabei die Annahme, dass Kompetenzen uneingeschränkt berufsübergreifend transferfähig wären, widerlegt. Auch wenn berufsübergreifende Kompetenzen nicht „uneingeschränkt“ transferierbar sind, und in der Regel im

Zusammenhang beruflich spezifischer Ausbildungsprozesse auf der Basis allgemeiner Fähigkeiten vermittelt werden, ist aus analytischen Gründen an der Unterscheidung zwischen fachlich spezifischen und berufsübergreifenden Kompetenzen festzuhalten.

2.4. ARBEITSDEFINITION FÜR EIN KONZEPT BERUFLICHER KOMPETENZ

Die Vielfalt der Kompetenz-Dimensionen und -ebenen versuchen wir für das Erhebungskonzept eines Berufsbildungs-PISA unter Rückgriff auf den anthropologisch-pädagogischen Ansatz von Roth (1971) und Reetz (1999) in einem eigenen Ansatz zu ordnen. Nach Roth und Reetz lassen sich vier Kompetenzbereiche wie folgt differenzieren:

- Selbstkompetenz;
- Sachkompetenz;
- Methodenkompetenz;
- Sozialkompetenz.

Dieses Schema, dessen Komponenten und hierarchische Struktur weiter unten im Detail diskutiert werden, hat sich in der deutschen Diskussion weitgehend als Rahmen durchgesetzt. Seine Bereiche erlauben eine idealtypische Behandlung. Allerdings ist ein Problem gegeben: Die Bereiche lassen sich nicht eindeutig gegeneinander abgrenzen. Nimmt man beispielsweise die Sachkompetenz und dabei die Wissenskomponente, so ist klar, dass Wissen für die Bewältigung von berufsspezifischen Aufgaben erforderlich ist, aber auch für die angemessene Durchführung von Prozeduren, wie das für den Bereich der Methodenkompetenz zu erfassen wäre; Wissen ist ebenfalls notwendig, um sich sozial angemessen verhalten zu können. Der Erwerb und der Einsatz von Wissen wird über Motive und Einstellungen gesteuert; die Reflexion über den Einsatz des Wissens und dessen Erfolg ist als Metakognition (in einem weiten Sinn) erforderlich. Von daher scheint es uns zweckmäßig, die vier Kompetenzbereiche nach Roth und Reetz so anzuordnen, dass diese Querbezüge deutlich hervortreten – was entsprechend auch Konsequenzen für die Kompetenzmessung hat. Bei der Konstruktion des in Abbildung 2–6 wiedergegebenen Schemas haben wir Anregungen von John Stevenson (Australien) aufgegriffen.

Diese Darstellung veranschaulicht das Ineinandergreifen und das Wechselspiel der verschiedenen Kompetenzbereiche und der erforderlichen bzw. zu berücksichtigenden individuellen Fähigkeiten.

Abbildung 2–6: Matrix zum Zusammenhang von Selbst-, Sach-, Methoden- und Sozialkompetenz mit verschiedenen Performanzbereichen

Individuelle Fähigkeiten (in Anspruch genommen und interpretiert in verschiedenen Kontexten)	Kompetenzbereiche (Performanz in verschiedenen berufsspezifischen Kontexten)		
	Selbstkompetenz		
	Sachkompetenz	Methodenkompetenz	Sozialkompetenz
	(theoretische, analytische Anforderungen) „Umgang mit Konzepten“	(technische, funktionale Anforderungen) „Umgang mit technischer Ausstattung“	(interpersonale Anforderungen) „Umgang mit anderen“
Einstellungen Werte Wahrnehmungen			
Antriebe Motivation			
Metakognitive Strategien			
Deklaratives Wissen			
Prozedurales Wissen			
Strategisches Wissen			

Die verschiedenen Aspekte dieses umfassenden Konzepts sollen unter der Perspektive der beruflichen Aus- und Weiterbildung mit ihren möglichen und gewünschten Verknüpfungen und Handlungsformen unter Hinweis auf prinzipielle Forschungs- und Interventionszugriffe in einen Zusammenhang gebracht werden.

Mit diesem Ansatz werden die Auszubildenden bzw. die Arbeiter und Angestellten nicht nur als Träger und Ausführende betrieblicher Tätigkeiten gesehen, sondern auch und besonders in ihrer Persönlichkeit betrachtet. Von daher ist Selbstkompetenz über die jeweiligen Ausprägungen von Sach-, Methoden- und

Sozialkompetenz bestimmt; wesentliche Bestimmungsmerkmale sind aber auch Einstellungen, Motive, Metakognition sowie eine professionsethische Handlungsorientierung (vgl. hierzu auch Beck, 1995). Die Bedeutung dieser Aspekte zeigt sich insbesondere dann, wenn berufliche Tätigkeiten mit einem hohen Anteil komplexer und systemischer Anforderungen ausgeübt werden, wie sie sich beispielsweise dem Berufsfeld eines „Symbolanalytikers“ (vgl. hierzu Baethge, Buss & Lanfer, 2003) zuordnen lassen.

Als ein Beispiel kann für den kaufmännisch-verwaltenden Bereich die Einführung von SAP in größeren und mittleren Unternehmen herangezogen werden. Diese hat dazu geführt, dass Rechnungswesentätigkeiten in einem hohen Maße auf das Bearbeiten von Bildschirmmasken reduziert werden und zugleich das Rechnungswesen selbst für eine Aus- bzw. Weiterbildung im kaufmännisch-verwaltenden Bereich nicht mehr konkret-anschaulich gegeben ist. Damit wird es für die Auszubildenden, aber auch für den überwiegenden Teil der kaufmännischen Angestellten sehr schwierig, die Produktions- und Geschäftsprozesse nachzuvollziehen und zu verstehen. Unter einer *ökonomischen* Perspektive haben Mitarbeiterinnen und Mitarbeiter an Arbeitsplätzen dabei so gut wie keine Chance, Entwicklungen der Geschäftstätigkeit konstruktiv-kritisch zu begleiten. Unter einer *pädagogischen* Perspektive trägt diese Art der Tätigkeit kaum zur Herausbildung persönlicher Autonomie bei. Das gilt auch für die Entwicklung der Humanressourcen; denn die zugrunde liegenden Wissensarten (mit ihren deklarativen, prozeduralen, strategischen Komponenten) sowie Motivation, Interesse, Metakognition werden nicht angesprochen und können damit auch nicht zur Entwicklung von Kompetenzen beitragen, die ihrerseits Anknüpfungspunkte für weitere Lernprozesse abgäben. Damit ist auch eine Exklusion von Entwicklungsmöglichkeiten im Arbeits- wie im Privatleben gegeben.

Vor diesem Hintergrund sehen wir es als sinnvoll an, die berufliche Handlungskompetenz vor allem unter der Perspektive der Wissensentwicklung zu betrachten; denn Expertise in einer Wissensdomäne bzw. in einem Tätigkeitsbereich ist zugleich eine entscheidende Bedingung für den Erwerb von transferfähigen Fertigkeiten, aber auch für die Entwicklung von Leistungsmotivation, Interesse und Metakognition – wie diese ihrerseits wiederum den Wissenserwerb fördern. Diese Sichtweise entspricht auch modernen Managementinstrumenten, wie zum Beispiel dem „organisationalen Lernen“ oder dem „Wissensmanagement“, die gerade eine aktive, gestaltende Arbeitnehmerrolle voraussetzen (vgl. z. B. Tuomi, 1999). Wieweit man im Hinblick auf eine Managementsicht, über die „gerade in führenden Unternehmen die Bedeutung von Lernprozessen hervorgehoben“ wird, „die zu einem qualitativ hochstehenden und individuell mit dem Blick auf die Ganzheitlichkeit betrieblicher Leistungsprozesse ausgezeichneten Arbeitsverhalten führen“, einer Koinzidenz ökonomischer und pädagogischer Vernunft (Achtenhagen, 1990, S. VII), kann an dieser Stelle offen bleiben. Eine solche Annahme zur Grundlage von Forschungsprojekten zu machen (vgl. Heid & Harteis, 2005), erscheint indes sinnvoll.

Im Folgenden diskutieren wir die Kompetenzbereiche etwas mehr im Detail, um unsere Überlegungen zu dem vorgeschlagenen Messverfahren zu verdeutlichen.

Selbstkompetenz

Im Rahmen des Kompetenzkonzepts steht Selbstkompetenz für alles, was als Persönlichkeitsentwicklung in einem engeren Sinn gefasst wird. Dabei geht es nicht nur um Kognition, sondern auch um emotionale, motivationale und moralische Entwicklungsprozesse (Achtenhagen, 1996, S. 27). Nach Reetz (1999, S. 42) betrifft Selbstkompetenz „die Fähigkeit zu moralisch selbstbestimmtem humanen Handeln. Dazu gehört neben der Behauptung eines positiven Selbstkonzepts (Selbstbildes) vor allem die Entwicklung zu moralischer Urteilsfähigkeit“. Für ein Berufsbildungs-PISA muss es darum gehen, allgemeine Persönlichkeitsmerkmale zu erfassen, denen ein Einfluss auf die berufliche Leistung zugesprochen werden kann. Hierzu gehören vor allem Einstellungen, Werte, Wahrnehmungen, Antriebe, Motivation und metakognitive Strategien.

Die Probleme der Beschreibung der Komponenten von Selbstkompetenz sind beeinflusst von Annahmen über ihre Entwicklung in einzelnen Lebensabschnitten. Dabei bestehen große Forschungslücken. Dass ein Großteil der Literatur zum lebenslangen Lernen die Einflüsse von Motivation und Metakognition systematisch vernachlässigt (vgl. die entsprechenden Beiträge in Achtenhagen & Lempert, 2000), ist schwer nachvollziehbar. Leistungsmotivation mit dem dazugehörigen Verhalten entwickelt sich ca. in einem Alter von 11 bis 12 Jahren und Metakognition ca. mit 14 bis 15 Jahren. Die Bedeutung von motivationalen und metakognitiven Strukturen wird sehr unterschätzt, was mit erklären kann, dass viele der Ausgaben für Maßnahmen der Weiterbildung nicht zu den gewünschten Ergebnissen führen – weil eben die Motivationsstruktur der Beschäftigten nicht angemessen berücksichtigt wird.⁸

Untersuchungen zur Entwicklung von Motivation während der Berufsausbildung zeigen – zumindest für den kaufmännischen Bereich in Deutschland – ein Absinken der Werte für interessiertes Lernen sowohl für die berufliche Schule als auch für den Betrieb. Während die schulischen Werte niedrig bleiben, steigen sie für den Betrieb bis zum Abschluss der Ausbildung wieder an (Prenzel, Kramer & Drechsel, 2001).

Wir können daher annehmen, dass es lohnenswert ist, Informationen über die motivationalen und metakognitiven Strukturen von Auszubildenden zu gewinnen. Gerade wenn mehrere Messzeitpunkte oder auch eine Retrospektive angesprochen sind, vermögen Daten zur Motivationsstruktur wesentliche Aufschlüsse zur Kompetenzentwicklung zu geben. Für die Erfassung der Motivation benötigen wir

8 Wenn Firmenvertreter ehrlich über ihre Erfahrungen mit Ausgaben für Weiterbildungsmaßnahmen berichten, geben sie zu, dass ca. 50% ihrer Gelder umsonst ausgegeben würden.

Itemkombinationen aus gewohnheitsbasierten und situativen Komponenten. Daher sollten Items bezogen auf motivationale Präferenzen (vgl. Deci & Ryan, 1985), selbstbezogene Kognition und Anstrengung und Ausdauer herangezogen werden (Kuhl, 1984; 1998). Dabei kann versucht werden, neben der „Trait“-Komponente auch „State“-Komponenten zu erheben (vgl. hierzu Winther, 2005). Hinsichtlich der Metakognition gibt es Konzepte, die versuchen, Probleme relativ unabhängig von Inhalten zu beschreiben. Anerkannte Items zur Beschreibung und Evaluation von metakognitiven Strukturen existieren im Bereich von Bachelor- und Master-Studiengängen in Form eines Fragebogens für den Gebrauch von motivierten Lernstrategien (MSLQ; Pintrich, Smith, Garcia & McKeachie, 1991); aber auch für den Bereich der beruflichen Bildung lassen sich neuere Skalen-Entwicklungen finden (vgl. Bendorf, 2005). Reflexive berufliche Tätigkeiten und Anforderungen sind wichtig für Aussagen zu Lern- und Verarbeitungsstrategien und Erfolgzuschreibungen.

Erhebungen von Motivation und Metakognition sind relevant, wenn man Aufschlüsse über die integrierte Kompetenzentwicklung zusammen mit Kognitionen und Prozeduren gewinnen will. Die PISA-Studien für den allgemeinen Bereich haben gezeigt, dass es sinnvoll und notwendig ist, Informationen zu Einstellungen, Werten, Wahrnehmungen und Antrieben zu erheben. Die entsprechenden Instrumente sollten daher auch im Berufsbildungs-PISA eingesetzt werden. Sofern Kurzfassungen zu den „Big Five“ des Persönlichkeitsinventars angemessene Reliabilitäten aufweisen (vgl. Rammstedt, 2004), sollte überlegt werden, auch diese mit einzusetzen. Gerade wegen der zu erwartenden großen Heterogenität der Stichproben sind Maße wichtig, mit deren Hilfe valide Einordnungen möglich werden.

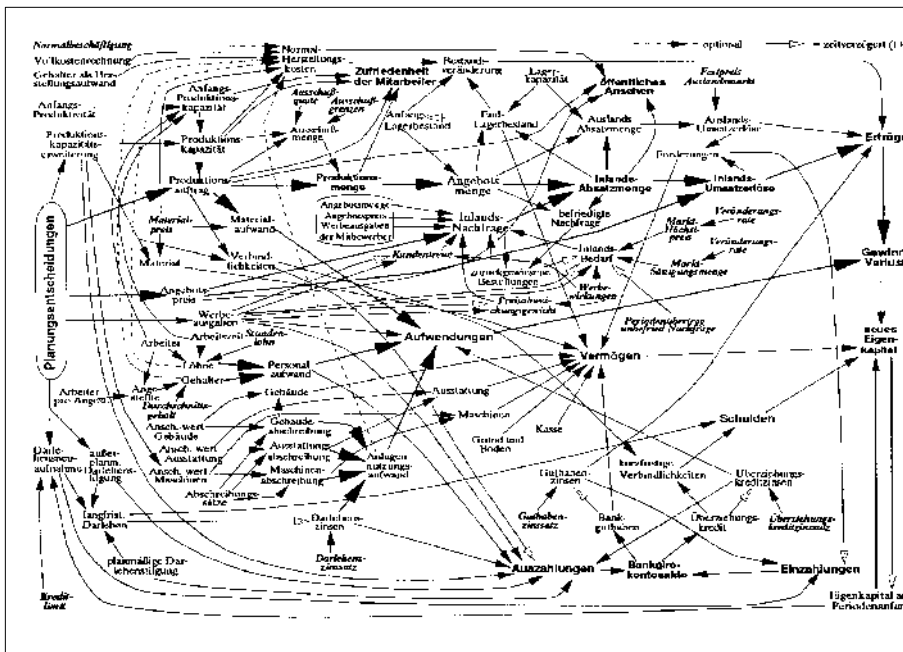
Sachkompetenz

Neuere empirische Studien belegen, dass die Berücksichtigung des Wissensbegriffs und dabei die Differenzierung nach Wissensarten: deklaratives, prozedurales, strategisches Wissen, plausibel und zugleich praktikabel sind (vgl. Bransford, Brown & Cocking, 2000)⁹. Für den Bereich des deklarativen Wissens, d. h. vor allem Faktenwissen, geht es um neue Muster, die in höherem Maße betrieblichen Anforderungen entsprechen. Angesichts der Komplexität und Vernetztheit betrieblicher Prozesse sollen deklarative Wissensbestände gerade nicht mehr lineari-

9 Shavelson, Ruiz-Primo und Wiley (2005) p. 415) schlagen als vierte Kategorie „Schematic knowledge“ (knowing `why`) vor. Eine genauere Analyse dieses für den naturwissenschaftlichen Unterricht entwickelten Konzepts zeigt aber große Überlappungen mit den anderen drei Wissensarten. Im Hinblick auf die Beherrschung der Ziele und Inhalte in der jeweiligen Wissensart unterscheiden sie „extent“ (how much?), „structure“ (how is it organized?) und „others“ (How efficient? How precise? How automatic?).

Abbildung 2–7 zeigt, wie komplexe, vernetzte betriebliche Strukturen wiedergegeben werden können. Es handelt sich um eine Aufbereitung für den Anfangsunterricht in Betriebswirtschaftslehre, die von der Idee her dem Ansatz des „vernetzten Denkens im Management“ von Gomez & Probst (1987) folgt. Für den kaufmännischen-verwaltenden Bereich wären für ein Berufsbildungs-PISA Testitems vorzusehen, die auf ein Verständnis derartiger vernetzter Strukturen bezogen sind: Bei Shavelson, Ruiz-Primo & Wiley (2005), finden sich Vorschläge für entsprechend formulierte Testitems und Verfahren; vgl. auch das Heft der *Unterrichtswissenschaft*, 1999/2: Alternative Evaluationsverfahren, (herausgegeben von Achtenhagen & Baumert); denn ein effektives Handeln im Betrieb sowohl im Hinblick auf die Konzepte der Betriebsorganisation als auch auf Prozeduren des Rechnungswesens setzt deren Kenntnis voraus. Hier bieten sich in besonderem Maße Simulationen an. Das hier abgebildete Wissensnetz diene als Grundlage für eine Planspielkonstruktion.

Abbildung 2-7: Wissensstruktur für den Anfangsunterricht in Betriebswirtschaftslehre



Quelle: Preiß, 1992, S. 62.

Für den Bereich des prozeduralen Wissens geht es dann um die Anwendung von Wissen, d. h. vor allem darum, wie mit solchen Wissensnetzen und ihren Elementen umzugehen ist.

Ein zentrales Ziel beruflicher Bildung stellt der Auf- und Ausbau strategischen Wissens dar. Hierbei sind Verhaltensbedingungen und Problemlösungen angesprochen, die sich auf ein reichhaltiges deklaratives und prozedurales Wissen stützen und dabei insbesondere die gewünschten Haupteffekte und die unerwünschten Nebeneffekte versuchen, gegeneinander abzuwägen. Entscheidungen sind dann in Kenntnis ihrer möglichen negativen Konsequenzen zu treffen.

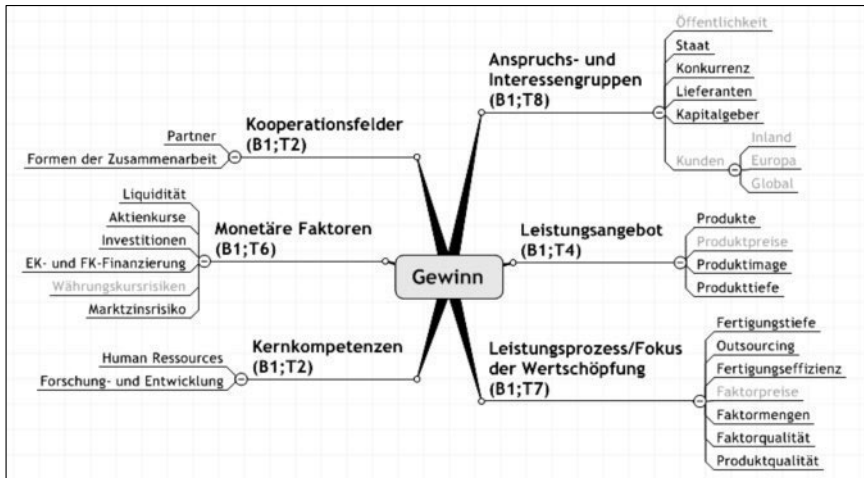
In der Neufassung der Bloomschen Taxonomie von Lernzielen durch Anderson & Krathwohl (2001), die sich dadurch auszeichnet, dass die Wissensdimension mit berücksichtigt wird, findet sich diese Aufteilung ebenfalls. Zugleich wird als zusätzliche Kategorie noch metakognitives Wissen eingeführt. Ein Vorteil dieser Neufassung des Wissensbegriffs liegt darin, dass sich deklaratives und prozedurales Wissen integrieren lassen, so dass beide Wissensarten in spezifischen Situationen zielgerichtet Anwendung finden können. Damit werden Prozesse eines einseitigen Lernens und eines vertieften Verständnisses sowie zugleich effektiver Handlungen in eine Balance gebracht. Hier haben Kompetenzerhebungen anzusetzen. Dabei wäre zu prüfen, inwieweit der Umgang mit Konzepten, mit einer technischen Ausrichtung oder mit Mitarbeiterinnen und Mitarbeitern kompetent erfolgt.

Bei dem hier verfolgten Konzept für ein Berufsbildungs-PISA spielt eine wichtige Rolle die Frage, wie sich Lernfortschritte für den deklarativen, prozeduralen und strategischen Bereich der Wissensanwendung abbilden lassen. Bezogen auf das Konzept „Gewinneinbruch bei einer Unternehmung“ kann beispielsweise ein Fall vorgegeben werden, der mit einem Kurzessay zu bearbeiten ist. Die Antworten lassen sich u. a. mit Hilfe einer Mindmap klassifizieren, wie das in Abbildung 2–8 dargestellt ist.

Das „B“ (Breite des Antwortverhaltens) und das „T“ (Tiefe des Antwortverhaltens) geben an, wie viele Argumente in welcher Form in der jeweiligen Lösung von den Probanden angesprochen werden. Die Differenzen bei zwei oder drei verschiedenen Testdurchführungen – sofern ein Panel realisiert wird – lassen sich dann im Sinne eines Lernfortschritts interpretieren. Die Frage, welches Messmodell zugrunde gelegt werden sollte, wird weiter unten diskutiert.

Zusammenfassend lässt sich festhalten, dass für ein Berufsbildungs-PISA domänenspezifisch Sach-, aber auch Methodenkompetenz eine ausschlaggebende Rolle spielen. Damit ergeben sich erhebliche Anforderungen sowohl an die fachdidaktische Abklärung der Ziel- und Inhaltsbereiche als auch an die Angemessenheit der Itementwicklung.

Abbildung 2–8: Mindmap als Klassifizierungsschema



Quelle: Erstellt im Rahmen des DFG-Projekts AC 35-24-1/2 (2005).

Methodenkompetenz

Um zu klären, was für die einzubeziehenden Ausbildungsgänge und die beruflichen Tätigkeitsprofile, auf die sie vorbereiten sollen, zentrale fachliche Kompetenzen sind, von denen man erwarten kann, dass sie in den unterschiedlichen Ausbildungssystemen zur Grundlage der Unterweisung gemacht werden, haben wir (in Anlehnung an die Methode der „Task-Analysis-Inventories“¹⁰) Ergebnisse aus am SOFI Göttingen durchgeführten Arbeitsplatzstudien (vgl. Baethge & Baethge-Kinsky, 2006 im Druck) mit Informationen über Ausbildungsprofile bedeutender Berufe/Tätigkeiten in Deutschland gekoppelt und auf dieser Basis Skill-Listen erstellt (vgl. Anhang). Die in diesen Listen aufgeführten Anforderungen wurden nach ihrer strategischen Relevanz für die Bewältigung der beruflichen Handlungs-

10 Bei der „Task-Analysis Inventories-Methode“ werden Anforderungen aus der Analyse von Funktionen, Arbeitshandlungen und den damit verbunden kognitiven Prozessen abgeleitet. Basis der Analyse sind Handlungsbeobachtungen von und Expertengespräche über abgegrenzte Tätigkeiten/Berufe und die jeweils auf sie entfallenden Aufgaben. Betrachtet werden das *Was* (Inhalte), das *Wann* (Dauer, Zeitpunkt und Häufigkeit von spezifischen Handlungssequenzen) und das *Wie* (Qualität, Niveau, Schwierigkeit, Fehlerkonsequenzen). Hierbei werden die Gebiete fachlichen Wissens wie auch die erforderlichen Fertigkeiten und Fähigkeiten nach ihrer strategischen Relevanz für die Bewältigung beruflicher Handlungssituationen (Kernkompetenzen) unterschieden. Auf diese Weise lassen sich Kerne fachlichen Wissens von ergänzenden Wissensbestandteilen ebenso abgrenzen wie auch in ihrer Verbindung zu Verhaltensanforderungen beschreiben.

situation ausgewählt und einzelnen Kompetenzdimensionen zugeordnet. Betrachtet wurden

- *industrielle Metall- und Elektro-Facharbeiter,*
- *Bankkaufleute,*
- *IT-Fachkräfte und*
- *Krankenpflegepersonal.*

Betrachtet man diese Anforderungsprofile im Quervergleich, so zeigt sich Folgendes:

1. In allen Tätigkeitsfeldern gibt es ein *domänenspezifisches, fachliches Wissen*. Dieses Wissen unterscheidet nicht allein und in erster Linie gewerblich-technische Tätigkeiten von kaufmännischen oder anderen Dienstleistungstätigkeiten, sondern markiert auf den Ebenen des deklarativen und prozeduralen Wissens Unterschiede auch innerhalb gewerblicher bzw. kaufmännischer Tätigkeiten: Bei den Metallfacharbeitern sind dies z. B. profunde Kenntnisse über Materialien und Verfahren der Metallbearbeitung; bei den Elektrikern/Elektronikern ist dies ein ausgeprägtes Wissen über den Aufbau und die Funktion von Leitungs- und Steuerungssystemen und den in ihnen verknüpften elektrischen, elektronischen, pneumatischen, hydraulischen und Software-Komponenten.
2. Für alle Tätigkeitsfelder gilt, dass die verschiedenen Wissensarten im Hinblick auf soziale Prozesse, auf Analyse- und Abstraktionsvermögen, auf Methoden- und Lernkompetenzen, wie auch im Hinblick auf organisatorische Strukturen und die eigene Einbindung in den Betrieb eine zentrale Rolle spielen. Dies hat etwas damit zu tun, dass eigenverantwortliches Arbeiten unter Zeitdruck, Wissensverschleiß, intensiver arbeitsbezogener Kontakt (mit Kollegen, Kunden oder Patienten), aber auch die Störanfälligkeit integrierter betrieblicher Prozesse ein Phänomen darstellen, das man in allen Branchen und dort auf allen Ebenen beobachten kann. Entscheidend ist, dass alle Wissenskompetenzen auch eine fachspezifische Einfärbung erhalten und insofern nicht allgemeine arbeitsbezogene Fähigkeiten darstellen. Das gilt auch für Verknüpfungen mit der Sozialkompetenz. So unterscheidet sich zum Beispiel die sozial-kommunikative Kompetenz im Rahmen industrieller Gruppenarbeit u. E. erheblich von derjenigen, die eine Krankenschwester im Umgang mit Patienten oder – noch einmal anders – ein Kundenberater in einer Bank beim Umgang mit Kunden aufbringen muss.
3. Bei unterschiedlichen beruflichen Funktionen (Produktion/Instandhaltung, Distribution, persönliche Pflege) spielen je spezifische Wahrnehmungsfähigkeiten und Sensibilitäten eine Rolle, die mit der Entwicklung von beruflicher Expertise zusammenhängen: Hierfür steht die technische Sensibilität der industriellen Berufe (durch Berühren, Hören und Sehen einer Anlage), die kommunikative Sensibilität des Bankkaufmanns (beim Aufspüren von Kundenbedürfnissen), auch die Sensibilität des Krankenpflegepersonals, mit der Aussehen, Schlaf und Appetit der Patienten registriert wird.

4. *Sensu-motorische Fähigkeiten* im Sinne manuellen Geschicks spielen zumindest bei den industriellen Fachkräften (beim Montieren oder Reparieren) und bei den Krankenpflegeberufen (beim Versorgen von Wunden, bei Injektionen und beim Umbetten von Patienten) eine wichtige Rolle (vgl. hierzu auch Achtenhagen & Grubb, 2001).

Die Frage der Arbeitsproben, d. h. der externalen Erhebung, stellt sich für den Bereich der Methodenkompetenz (wie auch folgend für den der Sozialkompetenz) in besonderem Maße. Hier werden Maßnahmen der Datenerhebung, die sich auf das Konzept der „vollständigen Handlung“ beziehen oder auf Simulationen zurückgreifen, eine besondere Rolle spielen.

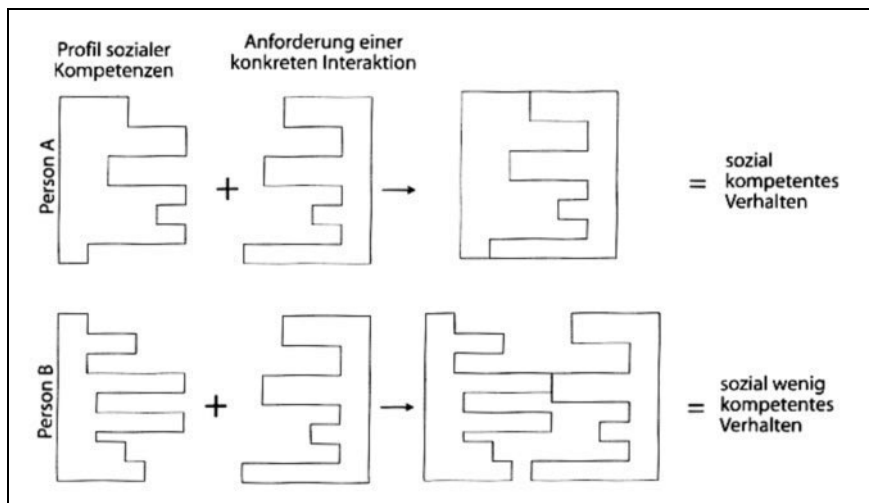
Sozialkompetenz

Sozialkompetenz umfasst die Gesamtheit der Fähigkeiten eines Individuums, die in kooperativen Situationen erforderlich sind, um im Umgang mit anderen sich bewegen und erfolgreich handeln zu können. Wie beim Kompetenzbegriff allgemein kann man mit Kanning (2003) bei Sozialkompetenz zwei Aspekte unterscheiden: Einerseits bezieht er sich auf das Potential eines Individuums, bestimmte Verhaltensweisen zeigen zu können (sozial kompetentes Verhalten), was die nicht-beobachtbaren individuellen Fähigkeiten bzw. internen Voraussetzungen einer Person umfasst; andererseits bezieht er sich auf das Verhalten von Individuen in konkreten, beobachtbaren Situationen (soziale Kompetenz). Soziale Kompetenz ist kontextabhängig und äußert sich unterschiedlich in Abhängigkeit von den Berufsfeldern.

Dementsprechend können *allgemeine soziale Kompetenzen* (ohne Spezifizierung auf bestimmte Situationen), wie beispielsweise Perspektivenübernahme und *spezifische soziale Kompetenzen*, über die nur Personen mit entsprechenden Lernerfahrungen (z. B. Altenpfleger, Verkäufer) verfügen, unterschieden werden. Dabei können sowohl spezifische Kompetenzen auch als eine auf Erfahrung basierende Ausdifferenzierung allgemeiner sozialer Kompetenzen verstanden werden, da diese sich nur selten völlig unabhängig voneinander entwickelt haben dürften. Sowohl die allgemeinen als auch die spezifischen sozialen Kompetenzen können entscheidend zur Generierung eines sozial kompetenten Verhaltens in einer konkreten Situation beitragen. Kanning (2003, S. 19–20) veranschaulicht diesen Sachverhalt wie folgt (Abbildung 2–9):

Das Profil sozialer Kompetenzen von Person A unterscheidet sich deutlich vom Profil von Person B. In einer bestimmten Interaktion stellen sich daher bei Person B im Gegensatz zu Person A deutliche Diskrepanzen heraus. In einer anderen Situation, die eher dem Profil sozialer Kompetenzen von Person B entspricht, würde dies anders aussehen.

Abbildung 2–9: Sozialkompetentes Verhalten.



Quelle: Kanning (2003).

Kanning (2003, S. 21) unterscheidet drei Gruppen sozialer Kompetenzen:

(a) *Perzeptiv-kognitiver Bereich*

Selbstaufmerksamkeit (direkt, indirekt); Personenwahrnehmung; Perspektivenübernahme; Kontrollüberzeugung (internal, external); Entscheidungsfreudigkeit; Wissen.

(b) *Motivational-emotionaler Bereich*

Emotionale Stabilität; Prosozialität; Wertepluralismus.

(c) *Behavioraler Bereich*

Extraversion; Durchsetzungsfähigkeit; Handlungsflexibilität; Kommunikationsstil (Unterstützung fordern und gewähren, Bewertung, Einflussnahme, Expressivität, Zuhören); Konfliktverhalten (verwirklichen eigener Interessen, Berücksichtigung der Interessen anderer); Selbststeuerung (Verhaltenskontrolle im sozialen Kontext, Selbstdarstellung).¹¹

Weber (2005) hat ein integriertes Instrument zur Messung einer „mindful identity negotiation“ auf der Grundlage der kulturell-historischen sozialen Lerntheorie und eines integrierenden Kommunikationsmodells entwickelt, das mit 17 hoch reliablen Items interkulturelle „communication skills“ abbildet (S. 251/252): „Presenting“, „Evoking“, „Facework“, „Supporting others“, „Personal identity needs“, „Self-reflection: Awareness“, „Self-reflection: Questioning“, „Personal Identity“, „Relational Identity“. Dieses Instrument wurde für die Evaluation interkulturellen Lernens entwickelt. Dabei betrifft „interkulturell“ nicht nur „verschiedene Nationen“, sondern in Betrieben auch unterschiedliche Hierarchien oder das Zusammen-

11 Kanning gibt dann noch eine Fülle von Hinweisen auf entsprechende Testinstrumente.

menwirken unterschiedlicher Berufsgruppen. An diese Vorarbeiten lässt sich anschließen.

2.5. VORSCHLAG FÜR DIE DATENERHEBUNG

Den beiden in Abschnitt 2.2. erörterten grundlegenden Ansätzen zur Erfassung von Kompetenzen – zum einen auf der Grundlage von externen Tätigkeiten, zum anderen auf der Grundlage von internen Bedingungen – entsprechen zwei Vorgehensweisen der Kompetenzmessung:

- (1) die Erfassung von Verhalten während der Arbeit (Performanz);
- (2) die Erfassung von Kompetenzen über die Durchführung von Tests (die Arbeitsproben einschließen können).

ad (1): Gegen die Erfassung von Verhaltensleistungen während der Arbeit sind drei gravierende Einwände im Hinblick auf die Durchführung eines Berufsbildungs-PISA zu erheben, die für die Messung berufsfachlicher Kompetenz ein hohes Maß an Bedeutung besitzen:

- Die Erbringung spezifischer Arbeitsaufgaben setzt eine Einigung über den Stellenwert der zu erbringenden Leistung für die untersuchten Berufe bzw. das untersuchte Aufgabenbündel voraus. Zugleich muss sichergestellt werden, dass die zu untersuchenden Aufgaben in den verschiedenen Ländern in vergleichbarer Weise als relevant angesehen werden.
Unsere Ausführungen haben deutlich gemacht, dass es sehr schwierig ist – und das ist auch die einhellige Meinung der Experten –, solche länderübergreifend gleichermaßen berufstypische Aufgaben zu definieren.
- Ein zweiter Einwand betrifft die zeitliche Erstreckung der Erhebung. Will man die Erbringung von Arbeitsproben valide, reliabel und objektiv für die Vergleichsstudie vorsehen, so muss mit einem erheblichen Zeitaufwand für die Durchführung gerechnet werden. Zieht man als Beispiel die sogenannten Berufsolympiaden heran, so sieht man, dass die dort zu erbringenden Arbeitsproben einen Zeitaufwand von bis zu 20 Stunden erfordern. Dieses Vorgehen lässt sich dort angesichts der geringen Teilnehmerzahl mit weitgehend identischen Arbeitsbedingungen realisieren.
- Der dritte Einwand betrifft die Begrenztheit des Ansatzes im Hinblick auf die Erhebung der von uns für die Berufsbildung für notwendig gehaltenen und hervorgehobenen Zielsetzungen: Sicher ließen sich so die arbeitsbezogenen Qualifizierungsaspekte abbilden, nicht aber die Aspekte der allgemeinen Persönlichkeitsentwicklung oder der gesellschaftlichen Teilhabe.

Von daher scheidet für unseren Vorschlag diese Variante aus theoretischen und methodischen Gründen aus, auch wenn sie für die Messung berufsfachlicher Kompetenzen eine hohe Validität haben könnte.

Unsere Einschätzung deckt sich mit den Nachteilen, die in vielen Einzelbeiträgen und resümierenden Stellungnahmen in den beiden zusammenfassenden, ge-

rade erschienenen Kompendien von Erpenbeck & von Rosenstiel (2003a) und Arbeitsgemeinschaft Betriebliche Weiterbildungsforschung (2005) geäußert werden, die zugleich auch auf Lehrbücher zur Arbeitspsychologie verweisen. Übereinstimmend wird herausgestellt, dass Leistungserhebungen in den Arbeitsprozessen selbst zu aufwändig und zeitraubend wären. Zudem wäre unter dem Aspekt der Repräsentativität das Arbeits- und Tätigkeitsbündel sehr umfassend zu ermitteln. Vor allem die Frage selten im Arbeitsprozess abgefragter, aber äußerst wichtiger Tätigkeiten spielt hier hinein wie auch der Sachverhalt, dass eine Versuchsperson zwar über die notwendigen Kompetenzen verfügt, diese sich aber in der Analyse nicht zeigen bzw. zeigen müssen. Von daher wird für die Kompetenzerfassung durchgängig von Messungen während der Arbeitsprozesse selbst abgeraten und auf Simulationen, komplexe Erhebungsverfahren sowie biographische Aufzeichnungen verwiesen.

Während der Workshops wurde unter Hinweis auf das Verfahren der „GWA“ (Generalized Work Activities) der Vorschlag gemacht, dieses Verfahren einzusetzen. Es böte sich vor allem deshalb an, weil es auf einer allgemeineren Ebene operierte, was eine Vergleichbarkeit sowohl zwischen verschiedenen Berufsfeldern als auch entsprechend zwischen Tätigkeiten in verschiedenen Ländern ermöglichte. Von der Durchführung her ließe sich eine GWA-Erhebung innerhalb eines akzeptablen Zeitraumes durchführen.

Schwierig zu beantworten ist die Frage, ob dieses Vorgehen berufsspezifisch genug ist. Zudem fehlt – jedenfalls nach unseren bisherigen Wahrnehmungen – ein Hinweis, dass von dieser Art der Performanzfeststellung her differenziert genug auf Kompetenzbereiche und ihr Wechselspiel rückgeschlossen werden kann. Es wäre zu überlegen, ob das GWA-Erhebungsverfahren zu Zwecken einer Triangulation eingesetzt werden sollte. Als einziges Erhebungsverfahren halten wir es zur Zeit noch für zu wenig für Fragen der Berufsbildung adaptiert. Toolsema und Nijhof haben ihre bisherigen Studien im Bereich „Higher Education“ und nicht direkt im Berufsbildungsbereich durchgeführt. Wir wollen nicht ausschließen, dass bei einem Transfer vom Hochschul- zum Ausbildungsbereich Schwierigkeiten auftreten könnten, die klare Aussagen überdeckten.

ad (2): Auch dieser Vorschlag weist Vor- wie Nachteile auf:

- Für die angemessene Erfassung von Kompetenzen sind – wie im Falle des authentischen Prüfens im Arbeitsprozess – die Tätigkeiten bzw. Situationen zu identifizieren, die für berufliche Tätigkeiten in einem Berufsfeld als ausschlaggebend anzusehen sind. Dabei muss auch hier die internationale Abstimmung gegeben sein. Die Beispiele, die über die verschiedenen National Vocational Qualifications-Modelle erzeugt werden, können bei diesem Vorgehen hilfreich sein. Es ist auch auf Bemühungen zu verweisen, die Abschlussprüfungen für anerkannte Ausbildungsberufe situationsspezifischer zu gestalten. Von daher scheint es möglich zu sein, über angemessene Tests die Performanz für Sach-, Methoden- und Sozialkompetenz zu erfassen, wobei die letztere am wenigsten leicht zugänglich ist.

- Ein Vorteil, sich der Kompetenzerfassung über entsprechende Tests zu nähern, kann darin gesehen werden, dass das von uns vorgeschlagene integrierte Modell in seinen Bereichen in einer umfassenden, integrierenden Weise zum Einsatz gelangen kann:
- So lassen sich mit entsprechend komplexen Aufgaben die Zellen in Abbildung 2–11 abdecken. Es kann festgestellt werden, ob angemessenes Wissen vorhanden ist und auch angewendet wird.
- Zugleich kann über die Erhebung z. B. von Motivation oder Metakognition erfasst werden, wie Selbstkompetenz mit Sach- oder Methodenkompetenz korreliert. Das erlaubt auch Rückschlüsse auf die inneren Bedingungen für erfolgreiches lebenslanges Lernen.

Der Umfang der entsprechenden Testaufgaben lässt sich der ins Auge gefassten Erhebungsdauer anpassen. Bei einer entsprechenden Konstruktion können bei mehrfacher Messung sowohl quantitative als auch qualitative Veränderungen erfasst werden. Die Testaufgaben können auch dann eingesetzt werden, wenn ein retrospektives Vorgehen zum Einsatz kommt.

Diese Überlegungen lassen sich mit Hilfe der Abbildung 2–10 spezifizieren:

Der besseren Darstellung halber haben wir den Kompetenzbereich „Selbstkompetenz“ über den Zeilen der Matrix abgebildet, um die entsprechenden Ausprägungen (Einstellungen; Werte, Wahrnehmung; Antrieb, Motivation; Metakognitive Strategien) sowie den Bezug zu Sach-, Methoden- und Sozialkompetenz zu verdeutlichen.

Für alle „Zellen“ gilt, dass alternativ oder ergänzend zu Testvorgaben auch biographische Berichte zum Einsatz kommen könnten. Hier ist insbesondere zu prüfen, wie auch übergreifende Zielsetzungen für die Berufsbildung mit erfasst werden. Die endgültige Auswahl und Zusammenstellung der Erhebungsinstrumente hat sich auch an der vorzugebenden Durchführungsdauer zu orientieren.

Abbildung 2–10: Plan für die Messung berufsübergreifender und berufsfachlicher Kompetenzen (mit Verweisen auf Testinstrumente)

Individuelle Fähigkeiten (in Anspruch genommen und interpretiert in verschiedenen Kontexten)	Kompetenzbereiche (Performanz in verschiedenen berufsspezifischen Kontexten)		
	Selbstkompetenz		
	Sachkompetenz	Methodenkompetenz	Sozialkompetenz
	(theoretische, analytische Anforderungen) „Umgang mit Konzepten“	(technische, funktionale Anforderungen) „Umgang mit technischer Ausstattung“	(interpersonale Anforderungen) „Umgang mit anderen“
Einstellungen Werte Wahrnehmungen	Selbstwirksamkeit (Selfefficacy) als Prädiktor für den Leistungsbereich (4 Items) (in der PISA-Studie eingesetzt: Kunter et al., 2002, S. 168)		
Antriebe Motivation	Handlungskontrolle in der Leistungssituation (4 Items) (in der PISA-Studie eingesetzt: Kunter et al., 2002, S. 171) Interesse und Motivation (18 Items) (Prenzel et al., 2001; verkürzte Version; im DFG-Schwerpunktprogramm eingesetzt) Selbstkonzept (verbal) (3 Items) (in der PISA-Studie eingesetzt: Kunter et al., 2002, S. 169)		
Metakognitive Strategien	LIST-Fragebogen zu Lernstrategien (Wild & Schiefele, 1994); verkürzte Version aus eigenem DFG-Projekt: 22 Items ($\alpha = 0,74-0,81$)		
Deklaratives Wissen	Komplexe Aufgaben zur simultanen Erfassung von Sach- und Methodenkompetenz		Fragebogen zu „kritischen sozialen Situationen“ („mindful identity negotiation“) (17 Items; $\alpha = 0,89$; Weber, 2005) (enthält Items zum deklarativen, prozeduralen und strategischen Wissen im Bereich der Sozialkompetenz)
Prozedurales Wissen			
Strategisches Wissen	Je nach Berufsfeld: komplexe Aufgabe zur Sachkompetenz komplexe Aufgabe zur Methodenkompetenz		

Fazit

Wir schlagen vor, interne Bedingungen beruflicher Performanz zu erheben. Entsprechende Testitems sind berufsfeldspezifisch national und international zu entwickeln. Es bietet sich an, zur Kontrolle und möglichen Triangulation eine „GWA“-Analyse durchzuführen. Anhand ausgewählter Beispiele soll gezeigt werden, wie Testitems konstruiert und eingesetzt werden könnten. Nach einer Darstellung des Konzepts der „vollständigen Handlung“ sollen Beispiele für die Erhebung von Fachkompetenz in den Bereichen „Car-Mechatronik“, „Nursing“ und „Bankkaufmann“ vorgestellt werden. Angerissen werden auch die Konzepte der „Lernaufgabe“, der „Handlungsfunktionen“, des „Problemlösens und der Computergestützten Simulation“ sowie möglicher „fachdidaktischer Systematisierungen“. Überlegungen zum vorgeschlagenen Messmodell schließen sich an.

Das Modell der vollständigen Handlung

Wenn Möglichkeiten ausscheiden, berufliche Tätigkeiten direkt im Arbeitsablauf zu beobachten, dann bietet es sich an, die Itementwicklung auf der Grundlage des

Abbildung 2–11: Elemente einer vollständigen Handlung

Arbeitsauftrag verstehen bzw. Ziel festlegen	- verstehen, um welche Art von Auftrag es sich handelt - erkennen, welche Produkte in welcher Qualität erwartet werden - ggf. selbständig ein Ziel setzen - übergeordnete Ziele (z. B. Wirtschaftlichkeit) erkennen und verstehen
Ausgangssituation analysieren	- auswählen, beschaffen und auswerten von <i>Informationen</i> - welche Informationen werden benötigt, welche sind vorhanden, welche sind überflüssig? - wo, wie und von wem können Informationen beschafft werden? - wie sind diese Informationen zu integrieren? - erkennen, welche weiteren <i>Personen</i> (mit welchen Rechten, Pflichten, Fähigkeiten) an der Handlung zu beteiligen sind - auswählen der <i>Arbeitsmittel</i> (im weitesten Sinn, d. h. auch EDV-Programme, Zahlungsarten) - erkennen, welche Rahmenbedingungen (z. B. zeitliche Einschränkungen, Rechtsvorschriften) zu berücksichtigen sind
weitere Arbeitsschritte bestimmen	- erkennen, welche Arbeitsschritte in welcher Abfolge durchgeführt werden müssen Lücken, Widersprüche und überflüssige Schritte sowie deren Folgen in einem Handlungsplan erkennen - evtl. alternative Handlungspläne vergleichend bewerten
weitere Arbeitsschritte ausführen	- Abarbeiten der einzelnen Teilschritte
Handlungsergebnisse kontrollieren	- die Qualität eines Ergebnisses prüfen - beurteilen, ob bzw. inwieweit ein Ziel erreicht ist - Fehler erkennen - Fehlerursache ausfindig machen - Fehlererfolg abschätzen können

Quelle: BIBB (1999).

„Modells der vollständigen Handlung“ (vgl. u. a. Aebli, 1987) vorzunehmen, da von hier aus Aufschlüsse gewonnen werden können, wie Tätigkeitsabläufe internalisiert und/oder beherrscht werden. Dieses Vorgehen wurde bereits bei der Prüfungsgestaltung erprobt (vgl. BIBB, 1999).

Das Modell des „Car-Mechatronic“

Ein Beispiel für die Zugrundelegung des Modells der vollständigen Handlung für die Generierung von Items sollte für den Bereich des Ankaufs und der Aufberei-

Abbildung 2–12: Ankauf und Aufbereitung von Unfallfahrzeugen für den CG-Fuhrpark

Auftrag	Lernsituation	Unterrichtssequenz	Zeit	Leitfragen
U N F A L L R E P A R A T U R	Fahrzeug-sichtung <i>didaktischer Fokus:</i> <i>Analyse des Systems</i>	Schadens-beurteilung	20 h	- Welche Aufgaben hat die Karosserie? - Wie ist eine Karosserie aufgebaut? - Welche Sicherheitsmerkmale weist eine moderne Karosserie auf? - Wie wird eine fachmännische Schadensbeurteilung an der Karosserie eines verunfallten Fahrzeuges durchgeführt?
	Reparatur schadhafter Karosserieteile	Karosserie vermessen Teilersatz-/Abschnittsreparatur	60 h	- Wie wird eine Karosserie vermessen? - Welches sind die Kriterien einer Teilersatz- und einer Abschnittsreparatur? - Was ist hinsichtlich des Karosserieschutzes und der Lackierung erneuerter Teile zu beachten? - Wie wird das Fahrwerk eines Fahrzeuges vermessen? - Was ist hinsichtlich Felgen und Reifen eines Fahrzeuges zu beachten? - Welche Achsen, Federungsarten und Dämpfungssysteme gibt es für heutige Kfz? - Welche Bedeutung haben Radstellungsgrößen? - Wie wird ein Kfz gelenkt? - Welche Möglichkeiten der aktiven Fahrwerksstabilisierung bietet die heutige Technik?
	Reparatur am Fahrwerk	Felgen, Reifen, Achsen Radstellungsgrößen Lenkung		

Quelle: Wolter & Petersen (2002).

tung von Unfallfahrzeugen in der Ausbildung von „Car-Mechatronics“ gegeben werden (Wolter & Petersen, 2002). Ausgangspunkt ist die Überlegung, dass in allen Ländern sich Autounfälle ereignen und dass es übereinstimmende Handlungen gibt, die durchgeführt werden müssen und an denen in komplexer Weise die fachlichen Kompetenzen von KfZ-Mechatronikern in ihren unterschiedlichen Wissens- und Methoden-Dimensionen testen lassen.

Das Modell gestattet es, eine oder mehrere komplexe Aufgabenstellungen zu formulieren, die für die Sachkompetenz oder die Methodenkompetenz jeweils deklaratives, prozedurales und strategisches Wissen ermitteln: Die Wissenskomponenten lassen sich im Hinblick auf den Umgang mit Konzepten bzw. mit der technischen Ausstattung erheben. Es bieten sich auch Computersimulationen an. Da die Abgabe eines beschädigten Autos in einer Werkstatt sehr oft eine kritische Situation darstellt, kann es sich auch anbieten, Items zur Sozialkompetenz im Umgang mit dem Kunden zu entwickeln.

Für den internationalen Vergleich könnte sich der „Car-Mechatronic“ in besonderem Maße eignen, da hier bereits ein europäisches Berufsbild mit der Auflistung der Arbeitsaufgaben existiert, auf das das deutsche Berufsbild bezogen werden kann.

Abbildung 2–13: Berufliche Arbeitsaufgaben des europäischen Berufsbildes „CarMechatronic“

1. Standardservice
2. Verschleißbehebung
3. Fahrzeugpflege
4. Administrative Dienstleistungen
5. Standarderweiterungen und Zusatzinstallationen
6. Große Inspektion, einschließlich Sommer-, Winter- und Urlaubscheck
7. Service-Dienstleistungen (AU, HU, SP)
8. Schadensbehebung (Fahrwerk, Karosserie, Lenkung, ...)
9. Fehlersuche und Reparatur
10. Reparatur von Aggregaten
11. Erweiterungs- und Zusatzinstallationen
12. Sonderdiagnose und -reparatur
13. Unfallschäden
14. Reklamationen
15. Prüfmaßnahmen bei Systemen und Alternativen

Quelle: Rauner & Spöttl, 2002, S. 144.

Das Modell des Nursing (Krankenpfleger)

Die international zu findenden Programme (hier ein Beispiel für Oregon) zeigen, dass alle Kompetenzbereiche angesprochen sind. Die in Abbildung 2–14 genannten Beispiele belegen den Bezug zu den Wissensarten für die Sach- und Methodenkompetenz, aber auch für die Sozialkompetenz (die Auseinandersetzung mit Eltern etc.). Insgesamt markieren die genannten Ziel- und Inhaltsbereiche die Basis für die Konstruktion von Testaufgaben und eine Itementwicklung. Im Hinblick auf den Umgang mit Patienten gibt es Simulationen für den Kompetenzerwerb.

Abbildung 2–14: Pflichten und Verantwortungsbereiche für die Krankenpfleger

Untersuchung von Patienten (Patient Examination)

Typical tasks: examines heart, lungs and respiratory system using stethoscope to identify heart murmurs and determine the nature and extent of other disorders; checks for masses and enlargements in major organs and glands, such as for evidence of enlarged liver, enlarged thyroid, ovarian and breast cysts and tumors; examines teeth, throat, ears, eyes, skin, and other parts of the body for signs of malformation or disease; conducts microscopic examination of samples to identify infections and blood cell counts; draws blood, takes smears, and orders laboratory studies; interprets results of laboratory studies.

Behandlung von Patienten (Patient Treatment):

Typical tasks: develops and carries out complete patient treatment plan as primary provider of health care services; treats infections, sprains, and simple fractures; sutures wounds not involving major blood vessels, nerves or tendons; suggests food supplements, non-prescription medicine, and prescribes antibiotics and other drugs to the extent authorized; refers cases to physicians and health specialists for conditions outside this employee's ability to diagnose or treat according to established protocol; confers directly with one or more physicians regarding diagnosis and care-management problems.

Ausbildung (Training):

Typical tasks: instructs nurses and health specialists in techniques to expand their diagnostic, counseling, and referral skills; confirms and extends assessments made by clinic registered nurses; authorizes and oversees treatment given under that employee's supervision; explains normal growth and development, abnormalities, diseases, diagnostic methods and treatment to patients and their relatives; follows up with parents, relatives, schools, and health resource organizations to obtain specialized help and or financial assistance, and to promote understanding.

Quelle: State of Oregon, Department of Administrative Services, Human Resource Service Division (2005).

Das Modell für den Bereich der Bankausbildung

Das folgende Beispiel geht einen Schritt weiter, indem es eine mögliche konkrete komplexe Aufgabe zeigt, für die verschiedene Probleme mit Hilfe von „Short Essays“ zu beantworten und zu lösen sind. Zugleich soll demonstriert werden, wie mit einer komplexen Aufgabe unterschiedliche Wissensarten angesprochen werden können.

Abbildung 2–15: Komplexe Aufgabe aus dem Bereich der Bankausbildung

Lesen Sie bitte nachfolgenden Artikel aus der Frankfurter Allgemeinen Zeitung vom 21.7.2004 aufmerksam durch:

VW korrigiert Gewinnziel nach unten

Mehr als 500 Millionen Euro unter Plan / Neues Oberklassemodell auf Eis gelegt

rit. HAMBURG, 20. Juli. Der Volkswagen-Konzern wird sein Ergebnisziel auch in diesem Jahr verfehlen. Nach Informationen dieser Zeitung rechnet Europas größter Automobilhersteller für 2004 nur noch mit einem operativen Ergebnis vor Sonder-einflüssen von 1,5 bis 2 Milliarden Euro. Ursprünglich wollte Volkswagen das Vorjahresergebnis von 2,5 Milliarden Euro übertreffen. Die revidierte Prognose soll am Freitag im Halbjahresbericht veröffentlicht werden, heißt es. Ein VW-Sprecher wollte diese Informationen nicht kommentieren. Den Informationen zufolge ist im ersten Halbjahr – trotz der stabilen Erträge aus der Finanzdienstleistungssparte – mit einem kräftigen Gewinneinbruch zu rechnen. Als Konsequenz aus der abermals verschlechterten Ertragslage hat der Vorstand die Pläne für den Bau des „C1“ vorerst auf Eis gelegt. Mit dem Modell, das 2007 in Serie gehen sollte, wollte VW die Lücke zwischen Passat und Phaeton schließen.

Trotz der Gewinnrevision wird VW am Freitag voraussichtlich keine Verschärfung des im Frühjahr eingeleiteten Restrukturierungskurses ankündigen. Damals wurde das Sparziel für 2004 und 2005 auf insgesamt 4 Milliarden Euro verdoppelt. Die jetzige Zurückhaltung hängt womöglich damit zusammen, daß man vor den im September startenden Tarifverhandlungen keine Drohkulisse errichten will. Gleichwohl steigt der Druck auf die Gewerkschaften, bei den Personalkosten kompromißbereiter zu sein. VW will die Personalkosten bis 2011 um 30 Prozent senken.

Schon in den ersten drei Monaten des Jahres war das operative Konzernergebnis um 46 Prozent auf 329 Millionen Euro eingebrochen. Seither hat sich das von Konjunktur- und Konsumschwäche geprägte Marktumfeld kaum verbessert. Der Verband der Automobilindustrie erwartet allenfalls noch einen stagnierenden Inlandsabsatz. Auf die europaweit zögerliche Nachfrage reagieren alle Hersteller mit zum Teil kräftigen Preisnachlässen; das hinterläßt deutliche Spuren in den Bilanzen von Ford und General Motors (mitsamt der Tochtergesellschaft Opel). Auch die französischen Hersteller Renault und PSA Peugeot Citroën sowie die italienische Fiat bekommen den Preisdruck zu spüren. Unter den Massenherstellern schwimmt Toyota indes gegen den Strom. Dank der hohen Produktivität eilt der größte japanische Autobauer bei Absatz und Ergebnis von Rekord zu Rekord.

VW hat die Kaufanreize im Verlauf dieses Jahres stetig erhöht. Das gilt vor allem für den neuen Golf: Nach der kostenlosen Zugabe einer Klimaanlage gibt es seit Juni zusätzlich eine Händlerprämie von 928 Euro für die Inzahlungnahme eines Gebrauchtwagens. Dies stabilisiert den Absatz, verengt aber die Gewinnmargen – und zwar nicht nur beim Golf. Denn um die Differenzierung der Konzernmodelle untereinander aufrechtzuerhalten, müssen auch die Preise der übrigen Modelle wie VW Polo, Skoda Octavia oder Seat Ibiza (direkt oder indirekt) gesenkt werden.

Zudem gerät VW auf wichtigen Auslandsmärkten unter Druck. In Amerika macht den Wolfsburgern die Dollar-Schwäche zu schaffen: Bei Durchschnittskursen zwischen 1,20 und 1,25 Dollar je Euro dürfte das operative Ergebnis im Gesamtjahr um rund 1,2 Milliarden Euro belastet werden. Auch toben in Amerika, wo der VW-Absatz unter dem Modellwechsel bei Jetta und Passat leidet, noch härtere Rabattschlachten als in Europa. All dies bescherte allein im ersten Quartal in Nordamerika einen Verlust von 235 Millionen Euro. Der chinesische Markt, auf dem VW in der Vergangenheit blendend verdiente, hat stark an Dynamik verloren. Dort versuchen die Konkurrenten – allen voran General Motors –, VW mit Kampfpreisen die Marktführerschaft zu entreißen. Die Wolfsburgern müssen wohl mit einem geringeren Ergebnisbeitrag aus China rechnen.

Volkswagen im Vergleich

Kurs 31. Dezember 2003=Index 100

Quelle: Thomson Financial Databroom (20.7.: 16 Uhr)

Toyota erhöht Absatzziel, Seite 16.
Ford profitiert von seiner Finanzsparte, Seite 16.

Quelle: Frankfurter Allgemeine Zeitung (2004).

Item für deklaratives Wissen (Reproduktion)/Sachkompetenz:

Welche Gründe für den Gewinneinbruch bei der Volkswagen AG erkennen Sie in dem Zeitungsartikel?

Sie sind Aktionär der Volkswagen AG. Der Wert Ihrer Volkswagen-Aktien betrug am 31.12.2003 17.500 Euro. Verwenden Sie für die folgenden Berechnungen bitte die Angaben aus der Grafik des Zeitungsartikels.



Quelle: Thomson Financial Datastream (2004).

Items zum prozeduralen Wissen (Reproduktion)/Sachkompetenz:

Wie hoch ist der aktuelle Wert Ihrer Volkswagen-Aktien am 20.07.2004?

Wie hoch wäre der Wert Ihrer Aktien am 20.07.2004 gewesen, hätten Sie die 17.500 Euro nicht in Volkswagen-Aktien, sondern in den „DAX 30“ investiert?

Im Folgenden berechnen Sie bitte Ihren Verlust im Vergleich zu einer Investition in den „DAX 30“.

Wie hoch ist Ihr Verlust in Euro?

Wie hoch ist Ihr Verlust in Prozent?

Item zum strategischen Wissen (Recognition)/Sachkompetenz:

Sie sind über Ihren Verlust empört. Welche Maßnahmen, von denen Sie sich prinzipiell eine Kurserhöhung versprechen, könnten Sie als Aktionär von der Volkswagen AG fordern?

Es bestünde auch die Möglichkeit, im Hinblick auf die Ziele und Inhalte im Bereich der Kundenberatung diese Vorgaben und Items in die Gestalt einer „vollständigen Handlung“ zu bringen – und dann entsprechend Testitems zu entwickeln. Dabei wäre es auch möglich, über den Einsatz von Computern Prozeduralisierungen abzufragen. Diesbezüglich ließen sich Aufgaben unter Bezug auf die Kundenberatung auch im Hinblick auf die Sozialkompetenz ausdehnen¹².

Problemlösen und computergestützte Simulationen

Im Hinblick auf die Notwendigkeit, im Rahmen eines Berufs-PISA die Performanz der Auszubildenden bzw. Ausgebildeten angemessen zu erheben, ist der Einsatz computergestützter Testverfahren in Betracht zu ziehen. In PISA-Deutschland ist ein entsprechender Einsatz erfolgt, der auf die Messung *fächerübergreifender* und *außerfachlicher Kompetenzen* zielte, wie Problemlösefähigkeit, verstanden als zielorientiertes Denken und Handeln in Situationen, für deren Bewältigung keine routinierten Vorgehensweisen verfügbar sind. Das umfasst einen weiten Bereich von Problemlöseprozessen:

- Die Problemstellungen können von der begrenzten Aufgabenbearbeitung bis zur Planung umfangreicher Projekte reichen.
- Der inhaltliche Kontext kann in unterschiedlichen Domänen angesiedelt sein und unterschiedliche Grade von Authentizität aufweisen (vgl. Achtenhagen & Weber, 2003).
- Das Problem kann unterschiedlich geschlossen oder offen sein, mit mehr oder weniger transparenten Zielen und Rahmenbedingungen, mit wenigen isolierten oder zahlreichen vernetzten Elementen.

12 Neben der Sortierung von Aufgaben nach ihrer Schwierigkeit (bzw. nach der Fähigkeit der Probanden) stellt sich die Frage, aufgrund welcher Anforderungsmerkmale eine Aufgabe schwieriger ist. Prinzipiell geht es dabei um die fachdidaktische Relevanz der Aufgaben; dabei ist Fachdidaktik als Begriff sehr weit gefasst, da er auch die am Arbeitsplatz bzw. im Betrieb zu vermittelnden Ziele und Inhalte mit einschließen soll. Für das Berufsbildungs-PISA heißt das, dass im Prinzip sechs verschiedene Fachdidaktikkonzeptionen – für jeden ausgewählten beruflichen Bereich jeweils eine – zu entwickeln bzw. anzuwenden wären. Hier taucht dann natürlich auch die Frage auf, ob es für diese fachdidaktischen Konzeptionen Elemente gibt, die allen gemeinsam oder zumindest einigen von ihnen gemeinsam wären. Hier ist beispielsweise auf Oates' dritte Kategorie „task-specific but independent of specific jobs“ zu verweisen, mit der die Grenze zur Berufsspezifität markiert wird. Hier ist auch der Punkt, an dem die „GWA“ ansetzen, indem sie über 42 Deskriptoren 35 „Berufe“ (gemäß US-amerikanischer Definition) vergleichen.

Diese Merkmale bestimmen insgesamt die Komplexität eines Problems (vgl. Abbildung 2–15 als Beispiel für die Ausbildung im kaufmännisch-verwaltenden Bereich). Als Mittel zur Untersuchung komplexer Aufgaben haben sich Computersimulationen bewährt, damit die vielfältigen dynamischen Abhängigkeiten der Variablen eines Systems dargestellt sowie der Umgang mit ihnen erhoben und berechnet werden können. In Interaktion mit dem Computer exploriert die Testperson das simulierte System, entwickelt und testet Hypothesen über Zusammenhänge, erwirbt Wissen und kann schließlich den Zustand des Systems steuern. Die in PISA-Deutschland eingesetzten Systeme sind in erster Linie „finite Automaten“, also solche technischen Systeme, die a) nur eine begrenzte (finite) Anzahl von Zuständen annehmen, b) ihre Zustände entweder durch Benutzereingabe oder durch einen autonomen Prozess erlangen und c) in Abhängigkeit von der Eingabe und erreichtem Zustand ein Ausgabesignal produzieren. Alle möglichen Zustands transformationen lassen sich somit formal angeben. Die Identifikation der Systemstruktur und die Steuerung des Systems stellen unterschiedliche Anforderungen dar. Das Ausmaß an Explorationstätigkeit wird als bedeutsamer Faktor für den Wissenserwerb angesehen (vgl. Kunter et al., 2002; vgl. auch die Literatur zum Einsatz von Computersimulation in der Berufs- und Wirtschaftspädagogik; vgl. Achtenhagen, Preiß & Weber, 2005).

Aus Sicht der Problemlöseforschung birgt der Einsatz von Computern die folgenden Vorteile (vgl. Wirth, 2003; Baumert et al., 2002; Wirth & Klieme, 2003):

- Die Anforderung an die Testperson wird erweitert um das aktive Suchen und kontinuierliche Verarbeiten von externer Information und ein Feedback zu den Wirkungen der eigenen Systemmanipulation;
- Problemsituationen können wesentlich authentischer gestaltet werden;
- Der Problemraum umfasst über die diskreten Zustände „falsch“ und „richtig“ hinaus differenzierte Zwischenzustände auf dem Weg vom Ausgangs- zum Endzustand. So können nicht nur Resultate, sondern auch Verlaufsmerkmale des Lösungsprozesses erfasst werden. Hieraus lassen sich Indikatoren über Strategien, z. B. die Systematik des Aufgabenlöseverhaltens, gewinnen.

In den Augen des deutschen PISA-Konsortiums ließen es diese Überlegungen als notwendig erscheinen, drei computergestützte Instrumente (Virtuelles Labor, Ökologisches Planspiel und Raumfahrtspiel) und drei schriftliche Problemlöseverfahren zur Konstruktvalidierung im Rahmen des PISA-Feldtests einzusetzen. Einige Verfahren konnten aus der Problemlöseforschung adaptiert werden, einige wurden neu entwickelt.

Das Fazit lautet, dass die Adaptation von Verfahren der Problemlöseforschung für ein large-scale assessment als gelungen bezeichnet werden kann. Die Reliabilitäten der computergestützten Verfahren sind insgesamt befriedigend bis gut. Ein relativer Nachteil der computergestützten Verfahren ist der signifikante Zusammenhang mit der Computererfahrung, der jedoch als nicht bedeutsam angesehen wurde. Akzeptabel war auch das Ausmaß der geschlechtsspezifischen Leistungsunterschiede. Diese Befunde besitzen für ein Berufsbildungs-PISA große

Bedeutung; denn computergestützte Lehr-Lern-Arrangements werden seit längerem systematisch in der Berufsbildung eingesetzt (vgl. Achtenhagen & John, 1992; Achtenhagen, 2004) – allerdings primär mikrostrukturell, d. h. auf die Lehr-Lernprozesse bezogen. Zudem sind sie in stärkerem Maße, als dies in der Problemlöseforschung üblich ist, auf beruflich-inhaltliche Problemstellungen bezogen. Die Aufbereitung dieser Arrangements für ein large-scale assessment sollte auf jeden Fall durchgeführt werden, damit ein entsprechender Einsatz vorbereitet und erprobt werden kann.

2.6. MESSFORMATE UND SKALIERUNG: ÜBERLEGUNGEN ZUM MESSMODELL¹³

Wir haben gezeigt, welch unterschiedliches Aussehen Testinstrumente prinzipiell annehmen könnten. Gleichzeitig wurde deutlich, welche Entwicklungsarbeit erforderlich würde, wollte man das ganze Arsenal an Messinstrumenten und -formaten – vor allem Arbeitsproben, komplexe Aufgaben, Simulationen – für die Kompetenzkonstrukte in den verschiedenen Berufsbereichen zur Anwendung bringen. Für eine Eingrenzung des möglichen Messinstrumentariums sprechen jedoch nicht in erster Linie der Entwicklungsaufwand, sondern vor allem die zeitlichen Restriktionen in der Durchführung von Messungen: Selbst wenn Arbeitsproben in vielen Bereichen validere Ergebnisse versprechen als Tests, ist eine beobachtungs-basierte Messung von Kompetenzen zum jetzigen Zeitpunkt fast auszuschließen; denn ein solches Vorgehen in den designierten Berufsfeldern würde national wie international erheblichen Aufwand nach sich ziehen und den Rahmen eines Berufsbildungs-PISA sprengen. Dies legt zumindest die folgende Übersicht nahe, die den Zusammenhang zwischen Bildungs-/Kompetenzaspekten, Messverfahren und Aufwand in verschiedenen Berufs-/Kompetenzfeldern ausweist (vgl. Abbildung 2–16). Danach lassen sich folgende Feststellungen treffen:

- Zwar werden heute der *Beobachtung zugängliche Arbeitsproben* in praktisch allen Berufsfeldern als Verfahren zur Messung von Kompetenzen eingesetzt. Doch erfordern solche Messverfahren einen erheblichen zeitlichen Aufwand; selbst im Fall der vom VDI vorgeschlagenen „Bildungsstandards für das Fach Technik an mittleren Schulen“ liegt der Messaufwand in einem einzigen Handlungsfeld (Arbeit und Produktion) bei 3 Zeitstunden, in den anderen Fällen (z. B. Krankenpflege; komplexe Arbeitsaufgabe in den Elektroberufen) liegt der Aufwand erheblich darüber.
- Wechselt man zu *Arbeitsproben mit schriftlichen Aufgabenstellungen*, bei denen ein Szenario vorgegeben und die Aufgabe in Form eines Kurzeassays oder aber eines an bestimmten Punkten detaillierten Lösungsentwurfs (z. B. Erstel-

13 Beim folgenden Text greifen wir auf Beratung durch Prof. Dr. J. Rost, Leibniz-Institut für die Pädagogik der Naturwissenschaften, Kiel zurück. Für seine Unterstützung in einer längeren Diskussion am 07.09.05 in Göttingen danken wir ihm.

lung eines Pflichtenheftes oder einer technischen Skizze) bearbeitet wird, so lässt sich der zeitliche Aufwand in Grenzen halten. Hier ergeben sich – wie das Beispiel der „schriftlichen Aufgaben“ in den industriellen Elektroberufen zeigt – Möglichkeiten der Prüfung komplexer Fach- und Methodenkompetenzen im vertretbaren zeitlichen Rahmen.

- Am geringsten erscheint der zeitliche Bearbeitungsaufwand für *multiple-choice-Formate*, wie sie beispielsweise in der theoretischen Facharbeiterprüfung in Deutschland bis heute üblich sind (PAL/KA). Ihre Eignung zur Prüfung deklarativen Wissens in eng eingegrenzten Berufsfeldern kann vorausgesetzt werden. Um jedoch auch prozedurale und strategische Wissensbestandteile abzufragen, ist dieses Antwortformat nicht ausreichend: Denkbar ist zum Beispiel ein Vorgehen, das einer Arbeitsprobe mit schriftlicher Aufgabenstellung entspricht: Vorgesaltet wird eine komplexe fachliche Problemskizze mit daran anknüpfenden geschlossenen (multiple-choice-Formate) und offenen Fragen.

Abbildung 2–16: Kompetenzaspekte, Messverfahren und Aufwand nach ausgewählten Berufsfeldern

Kompetenz-/Berufsfeld	Bildungs-/Kompetenzaspekte	Messverfahren	Aufwand
Technik für den mittleren Schulabschluss – Handlungsfeld „Arbeit und Produktion“ (VDI)	Methoden der Planung, Konstruktion, Fertigung, Bedienung und Wartung anwenden und bewerten können	Arbeitsprobe mit Beobachtungs- und schriftlichen Sequenzen	3 Stunden
	Werkstoffe und deren Haupteigenschaften kennen und deren Eignung für einen geplanten Gebrauchsgegenstand bestimmen können		
	Ordnungsprinzipien eines Arbeitsplatzes und einer Werkstatt berücksichtigen		
	Grundfertigkeiten im Umgang mit Werkstoffen, Werkzeugen, Maschinen und Geräten besitzen		
„Nursing Education“ (Krankenpflege) (Open University of the Netherlands)	Sorge für kranke, behinderte und sterbende Patienten, um ihr Befinden zu erleichtern	Arbeitsprobe mit Beobachtungs- (Patientenberatung) und schriftlichen (Erstellung von Diagnosen, Pflege- und Behandlungsplänen) Anteilen	5 Tage
	Einleiten präventiver Maßnahmen auf unterschiedlichen Levels zur Verringerung von Gesundheitsrisiken		
	Beratung von Patienten und Angehörigen zur Förderung eines gesunden Lebensstils		
Industrielle Elektroberufe – Teil 1 (Borch u. a.)	Unterlagen auswerten, Arbeitsabläufe planen, Material und Werkzeug disponieren	-- Arbeitsprobe (Komplexe Arbeitsaufgabe) mit	12 Stunden

	Teile/Aggregate montieren, verdrahten, verbinden und konfigurieren	schriftlichen Aufgabenbestandteilen multiple-choice Test (Fachkenntnisse)	
	Sicherheit von Anlagen und Betriebsmitteln beurteilen, elektrische Schutzmaßnahmen prüfen		
	Elektrische Systeme analysieren, Funktionen prüfen, Fehler suchen und beseitigen, Betriebswerte einstellen und messen		
	Produkte in Betrieb nehmen, übergeben und erläutern, Aufträge dokumentieren		
Industrielle Elektroberufe – Teil 2 (Borch u. a.)	Selbstorganisation der Arbeit im Rahmen der Geschäftsprozesse, des Qualitätsmanagements (betrieblicher Organisationszusammenhang)	Dokumentierte Arbeitsprobe	18 bis 30 Stunden
	Technische Problemanalyse, Entwicklung von Lösungskonzepten unter Berücksichtigung technischer, ökonomischer und betrieblicher Parameter, anwendungsgerechte Spezifikation, Komponenten und Verfahren auswählen	Schriftliche Aufgabe (Änderung eines Systems nach vorgegebenen Kundenwünschen)	2 Stunden
	Technische Unterlagen auswerten, Mess-, Prüf- und Diagnoseverfahren auswählen, funktionelle Zusammenhänge analysieren, Fehlerursachen bestimmen	Schriftliche Aufgabe (Funktions- und Systemanalyse)	2 Stunden
	Beurteilung wirtschaftlicher und gesellschaftlicher Zusammenhänge der Berufs- und Arbeitswelt	Test (Wirtschafts- und Sozialkunde)	1 Stunde
vor allem gewerblich-technischer Bereich (BIBB)	Ganzheitliche Arbeitsaufgabe verstehen und beurteilen (unterschiedliche Beispiele)	Lernaufgabe	2 Stunden bis 5 Tage

Quelle: Eigene Darstellung nach VDI (2004): Bildungsstandards im Fach Technik für den mittleren Schulabschluss; Borch u. a. (2004): Industrielle Elektroberufe; Smits u. a. (2005): Performance Assessment in Nursing Education; BIBB (2005).

Es spricht daher einiges dafür, für alle hier diskutierten Kompetenzen den Einsatz von Messformaten zu prüfen, die in der praktischen Anwendung möglichst wenig zeitlichen Aufwand benötigen und dennoch aussagekräftig genug sind, um die Ergebnisse von Arbeitsproben zu ersetzen oder gar simulieren zu können. Die zum jetzigen Zeitpunkt offenen Fragen bestehen darin,

- ob es gelingt, über ein singuläres Format alle Kompetenzdimensionen zu erfassen. Dies ist durchaus denkbar. So lassen sich zum Beispiel ein Essay in mehreren Dimensionen – z. B. nach Länge, Komplexität des Aufbaus, Anzahl

- der Argumente – auswerten und Antworten auf vorgegebene Items prinzipiell kombinieren (mehrdimensionale Items);
- oder ob die einzelnen Dimensionen jeweils für sich und über verschiedene Formate (Kombination aus Itembatterien und Essays) erfasst werden sollten. Hierfür spricht vor allem, dass ein reichhaltiges Frage- und Antwortformat Ermüdungserscheinungen vorbeugt.

Wie auch immer die Antwort auf diese Fragen letztlich ausfallen mag: Vorerfahrungen, die den Entwicklungsaufwand in Grenzen halten dürften, bestehen mit verschiedenen Formaten: Mit Essays als Antwortformat, die sich nach verschiedenen Aspekten vercoden lassen, mit frei formulierten Kurzantworten wie auch mit multiple-choice-Formaten (vgl. OECD, 2001; Altiner u. a., 2002; Scholz, 2003). Schließlich liegen auch erste Erfahrungen mit dem Einsatz von Computer-gestützten Tests vor, bei denen z. B. Fehlfunktionen einer technischen Anlage simuliert werden und die zu lösende Aufgabe in der Fehlersuche und -diagnose liegt.¹⁴

Ziel der Entwicklung der Messinstrumente muss es letztlich sein, zu einem Verfahren zu kommen, in dem für jede Dimension ein Einzel-Score sowie über alle Dimensionen hinweg ein Gesamt-Score ermittelt und ausgewiesen werden kann. Dies ist grundsätzlich möglich, unabhängig davon, welche der eben erwähnten Formate oder Formatkombinationen gewählt werden. Gleichwohl liegt in der Modellierung der Testinstrumente bzw. des Testmodells im Hinblick auf eine singuläre bzw. die gesamte Kompetenzen ein wesentliches Problem: Zunächst ist es schon schwierig, Testaufgaben mit einem angemessenen Schwierigkeitsgrad und hinreichender Trennschärfe zu finden. Schließlich sollen alle eingesetzten, unterschiedlichen Formate einheitlichen Anforderungen an die Skalierung entsprechen (vgl. Rost, 2004).

Von besonderem Interesse wird die Frage sein, wie Querschnittserhebungen in eine stringente Beziehung zu Veränderungsmessungen im Sinne der Erhebung eines kumulativen Kompetenzerwerbs gesetzt werden können. Mit den kurz vorgestellten Aufgabenformaten, die bereits existieren, scheinen entsprechende Veränderungsmessungen möglich. Auch die aktuell gegebenen Ansätze zur Erstellung angemessener Prüfungsaufgaben für die Kammerprüfungen am Schluss einer Ausbildung bieten viele Anregungen für ein zügiges Konstruieren valider und reliabler Aufgaben und der korrespondierenden Testitems.

Pretest: Konstruktion und Prüfung von Skalen

Die Entwicklung der Testinstrumente war in der vorliegenden Machbarkeitsstudie nicht zu leisten und bleibt der Pilotphase einer späteren Vergleichsstudie vorbe-

14 Noch keine Erfahrungen liegen mit dem Einsatz berufstypischer Lernaufgaben vor, deren Sammlung gerade vom BIBB vorgestellt wurde und bei der Teilaufgaben jeweils unterschiedlichen Anforderungsniveaus zugeordnet werden (vgl. „das Konzept der Lernaufgabe“ im Anhang).

halten. Gleichwohl lässt sich angesichts der angesprochenen Auswertungsperspektiven einerseits und der erörterten Modellierungsprobleme andererseits schon sagen, dass man um eine Testtheorie und ein Testmodell heuristischer Qualität, welches die Validierung der unterschiedlichen Tests im genannten Sinne erlaubt, nicht herum kommt. Nach eingehender Auseinandersetzung mit unterschiedlichen Testtheorien und -modellen spricht momentan alles für die Verwendung der Item-Response-Theorie (IRT) und des darauf basierenden „Mixed Rasch-Modells“ (vgl. dazu Rost, 2004).

- (1) Die IRT als wahrscheinlichkeitstheoretischbasierte Methode der Modellierung des Antwortverhaltens weist insbesondere dann Vorteile auf, wenn
 - es darum geht, einen Test neu zu entwickeln,
 - man mit größeren Mengen an ausfallenden Daten (missing-data) zu rechnen hat,
 - Veränderungsmessung (also die wiederholte Messung derselben Personen) ein Bestandteil des Designs ist und
 - Aussagen über Teilkompetenzen anhand von Untergruppen der Testaufgaben abgeleitet werden sollen.

Gerade wenn die geplante Untersuchung auf ein Panel hinauslaufen soll und damit auch *Messwiederholungen* (Absolvieren desselben Tests durch dieselben Personen) eingeplant werden müssen, steht bei der Auswertung derartiger Längsschnittdaten zunächst die Klärung an, inwieweit strukturelle oder quantitative Veränderungen auf der Ebene des Gesamttests den Kompetenzzuwachs bestimmen. Diese Fragen lassen sich ebenfalls am besten unter Rückgriff auf die Item-Response-Theorie klären.

- (2) Das „Mixed Rasch-Modell“ als besondere Anwendung der IRT erlaubt es u. a., anhand der Datenanalyse von Vortests
 - aus einer vorgegebenen Zahl an Items jene herauszufiltern, die den Test in Richtung Eindimensionalität optimieren, sowie Items so zu gruppieren, dass sie jeweils eine eigene Dimension erfassen (Trennschärfe!),
 - Personengruppen zu identifizieren, deren Kompetenzen nicht mit dem Rasch-Modell skaliert werden können (Unschärfebereiche), sowie
 - Personengruppen zu unterscheiden, in denen die gemessene Kompetenz mit unterschiedlichen Lösungsprofilen (Antwortstrukturen) definiert wird (Rost, 2004).

Am Ende des angesprochenen Prozesses der Testentwicklung sollen – so der derzeitige Bearbeitungs- und Diskussionsstand – sogenannte „Rasch-Skalen“ stehen, aus denen Scores sowohl für die einzelnen Kompetenzdimensionen als auch für die „Gesamt-Kompetenz“ ermittelt werden. Rasch-Skalen gibt es wiederum in unterschiedlichen Formaten. In der Praxis haben sich – wie das Beispiel der ersten PISA-Studie zeigt (vgl. OECD, 2001; Scholz, 2003) – vor allem dichotome Rasch-Skalen (Aufgabe gelöst/nicht gelöst) eingebürgert; in jüngerer Zeit werden auch polytome (ordinale) Rasch-Modelle verwendet bzw. erweisen sich ordinal-skalierte Tests, wie z. B. die Erhebung der „Big Five“ (5er-Graduierung), im

nachhinein als Rasch-Skalen (vgl. Rost u. a., 1999). Dichotome Skalen lassen sich grundsätzlich auch auf der Basis von Szenario-Formaten mit offenen Antworten erstellen, die dann nach unterschiedlichen Gesichtspunkten (Häufigkeit von Nennungen, Anzahl richtiger oder falscher Nennungen, Anzahl unterschiedlicher inhaltlicher Aspekte, Struktur genannter inhaltlicher Aspekte etc.) binär verkodet werden können. Hierbei – und dies macht die Schwierigkeit bei der Entwicklung entsprechender Formate aus – muss die lokale stochastische Unabhängigkeit des einzelnen Gesichtspunktes/Items gewahrt bleiben, d. h. von der Formulierung der Aufgabe her ist die Möglichkeit sicherzustellen, dass der Proband die (implizite) Einzelaufgabe auch dann erfolgreich löst, wenn er bei den anderen Teilaufgaben ein schlechtes Ergebnis erzielt.

Wie groß der Aufwand für die Optimierung der Tests sein wird, lässt sich nicht genau bestimmen. Für jede gewählte berufliche Branche sollte zunächst ein Itempool von 20 (dichotom angelegten) Items pro Kompetenzdimension erarbeitet werden, die einer Testerprobung zu unterwerfen wären. Diese Items sollten in komplexe Aufgaben eingebettet sein. Dabei wird davon ausgegangen, dass bei entsprechend sorgfältiger Itemkonstruktion ca. 12 Items pro Dimension den Ansprüchen des Messmodells genügen. Für diesen Prozess wären in jedem Land jeweils 100 Probanden für jede Branche vorzusehen, um die Messfehler hinreichend niedrig zu halten.¹⁵

15 Angesichts der Heterogenität von Personen, aber auch Aufgaben begnügen wir uns mit diesen Erfahrungswerten und verzichten darauf, Scheingenauigkeiten über Effektstärkensätzungen etc. zu erzeugen. Das ETS in Princeton schlägt für die Ermittlung von „Computer Literacy“ in den einzelnen Dimensionen vergleichbare Itemzahlen vor.

3. INSTITUTIONELLE UND INDIVIDUELLE EINFLUSSFAKTOREN FÜR DIE QUALITÄT BERUFLICHER BILDUNGSPROZESSE

Im Zentrum eines Berufsbildungs-PISA steht die Messung von Kompetenzen, die in der beruflichen Erstausbildung vermittelt werden. Allerdings kann sich eine solche Studie nicht in der Messung von Kompetenzen erschöpfen. In der internationalen Bildungsforschung ist man sich weitgehend darüber einig, dass die institutionellen und individuellen Bedingungen von Bildungsprozessen von entscheidender Bedeutung für die individuelle Kompetenzentwicklung sind (vgl. Europäische Union, 2001). Erklärungsmächtig im Hinblick auf den Informationsgehalt der Kompetenzmessung sind demnach neben den vorgängigen Lernerfahrungen und Lerndispositionen der Individuen auch die Ausbildungsorganisation und deren Einbettung in jeweils spezifische soziale, ökonomische und pädagogische Bedingungen (vgl. Baumert & Schümer, 2002). Über die Leistungsmessung hinaus sollte – wie bereits im Antrag ausgeführt – ein Berufsbildungs-PISA auch Informationen über andere Outcome-Effekte liefern, wie die Verwertung beruflicher Qualifikationen im Beruf, auf den Arbeitsmärkten und in der Biographie. Auch hier können institutionelle und individuelle Kontexte erklärungsrelevant werden. Dementsprechend stellt ein Untersuchungskonzept, das die für die Leistungsfähigkeit der beruflichen Erstausbildung wesentlichen institutionellen und individuellen Variablen systematisch in die Untersuchung einbezieht, einen wichtigen Bezugspunkt der Studie dar: Mit der internationalen Vergleichsstudie soll also auch ein Beitrag zur Aufklärung über diejenigen Kontextfaktoren erbracht werden, welche für die Vermittlung und Verwertung von Kompetenzen wichtig sind und von denen anzunehmen ist, dass sie nach nationalen Ausbildungs- und Arbeitsmarktstrukturen variieren.

Ausgehend von diesen Überlegungen stellt sich die Frage nach einem geeigneten Forschungsdesign, das es erlaubt, einen Beitrag zur systematischen Untersuchung von Unterschieden in beruflichen Bildungsprozessen und -ergebnissen zu leisten. Wir sind, im Einvernehmen mit den internationalen Bildungsexperten der beiden Workshops, zu dem Schluss gekommen, uns an einem Mehr-Ebenen-Modell zu orientieren: System-, Ausbildungsinstitutions- und Unterrichtsmerkmale sollen daraufhin geprüft werden, wie sie – in Interaktion mit individuellen und sozialen Einflussgrößen – die Vermittlung und Verwertung von Kompetenzen beeinflussen. Bezugsbasis ist das „Bildungsproduktions-Modell“ von Scheerens (1990), das auch die konzeptionell-theoretische Grundlage für die PISA-Studien lieferte. Wir greifen damit die theoretischen Vorarbeiten und Ergebnisse der Schuleffektivitäts-, Schulorganisations- und Lehr-Lernforschung auf und modifizieren sie für die Untersuchung von Berufsbildungsprozessen.

Die Operationalisierung der Bedingungsfaktoren lehnt sich ebenfalls an Scheerens' Modell an, das auch von der OECD aufgegriffen worden ist und das Input-, Prozess- und Outputvariablen auf verschiedenen Ebenen konzeptualisiert (vgl. auch Baumert et al.; 2001, OECD, 2003; Baethge et al. 2003). Dabei soll im Hinblick auf die institutionellen Bedingungen unterschieden werden zwischen systemischen Bedingungen, die durch die Bildungssysteme der einzelnen Länder vorgegeben sind, und den jeweils spezifischen Bedingungen einzelner Bildungsanbieter, unter denen die Auszubildenden in den verschiedenen Ländern lernen (vgl. Scheerens & Bosker, 1997). Welche Themen innerhalb dieser Kategorien bevorzugt untersucht werden und welche von Bedeutung für ein Berufsbildungs-PISA sein könnten, ist Thema der folgenden Abschnitte. Auf ihrer Grundlage wurde ein Raster von Indikatoren entwickelt, nach denen die Untersuchung institutioneller und individueller Einflussfaktoren für die Qualität von beruflichen Bildungsprozessen einzuordnen ist.

Da bereits einige internationale Studien zur Verknüpfung von Bildungserträgen und deren Bedingungen durchgeführt wurden, haben wir zunächst geprüft, mit welcher Absicht und auf welche Weise in diesen Studien die mikro- und makrostrukturellen Modalitäten von Bildungsprozessen beschrieben werden und ob die dort verwendeten Indikatoren einen hinreichenden Gebrauchswert für eine internationale Vergleichsstudie von Berufsbildungssystemen haben (Kapitel 3.1.). An den Ergebnissen dieser Prüfung lässt sich verdeutlichen, wie sich ein Berufsbildungs-PISA in die internationalen Vergleichsstudien einordnen lässt und wo seine spezifischen methodischen Probleme liegen. Daraufhin werden wir herausarbeiten, welche Indikatoren von hoher Bedeutung für die Abbildung institutioneller und individueller Faktoren für die Qualität von Berufsbildung sind (Kapitel 3.2.). In einem abschließenden Abschnitt werden wir der Frage nachgehen, welche Daten auf internationaler wie nationaler Ebene bereits zur Verfügung stehen und welche Daten im Rahmen eines Berufsbildungs-PISA erhoben werden müssten (Kapitel 3.3.).

3.1. KONTEXTBEDINGUNGEN IN INTERNATIONALEN VERGLEICHSTUDIEN ZU BILDUNGSSYSTEMEN UND -PROZESSEN

Nationale Bildungssysteme und Bildungspolitik sind zunehmend in den Zusammenhang internationaler Entwicklungen eingebunden. Nachdem die bildungspolitisch und analytisch relevanten Arbeiten mit Schwerpunkt auf der Ebene der Organisation für wirtschaftliche Zusammenarbeit (OECD) im Rahmen der Veröffentlichungen von „Bildung auf einen Blick“ (vgl. OECD, 2003) und der Messung von Schülerleistungen (PISA) vorangetrieben wurden, gilt dies spätestens nach der Jahrtausendwende auch für die Europäische Gemeinschaft. Angestoßen durch den Europäischen Rat in Lissabon im Jahre 2000 hat sie (über die Entschließen-

gen von Kopenhagen und Maastricht) ihre Anstrengungen erheblich verstärkt, auch die nationalen Ausbildungssysteme in den Prozess der europäischen Integration einzubeziehen (vgl. BLK, 2002, S. 5;). In der Folge werden bildungspolitisch relevante Aspekte im internationalen Kontext verglichen und auf dieser Grundlage Schlussfolgerungen für das nationale Bildungssystem gezogen. Besonderer Wert wird hierbei auf Transparenz gelegt, sowohl für ein besseres Verständnis der Systeme und ihrer Zielsetzungen als auch im Hinblick auf deren Ergebnisse (vgl. Corti, 2000).

Im Folgenden werden die Ziele und Methoden der gesichteten Arbeiten vorgestellt und danach bewertet, ob sie

- a) einen Beitrag zur Ermittlung wesentlicher Kontextfaktoren für den Erwerb und die Verwertung von Kompetenzen in der beruflichen Erstausbildung leisten können und
- b) geeignete Daten für den internationalen Vergleich von beruflichen Bildungssystemen und -prozessen zur Verfügung stellen.

– Internationale UNESCO – OECD – Eurostat (UOE) – Datenerhebung

Die UOE-Erhebung ist ein Gemeinschaftsprojekt der Organisation der Vereinten Nationen für Bildung, Wissenschaft, Kultur und Kommunikation (UNESCO), der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) und des Statistischen Amtes der Europäischen Union (EUROSTAT). Ziel des Projektes ist es, international vergleichbare Daten zu erheben, die Informationen zur Funktionsweise, Entwicklung und Wirkung von Bildung geben. Bei den erfassten Kennziffern nationaler Bildungssysteme handelt es sich im Kern um quantitative Daten zu den Bereichen Finanz- und Humanressourcen der Bildung, Bildungszugang, Bildungsbeteiligung und Bildungsergebnisse. Die UOE-Erhebung ist integraler Bestandteil für verschiedene Bildungsstatistikprojekte dieser Organisationen: World Education Indicators (WEI) der UNESCO, International Indicators of Education Systems (INES) der OECD sowie Schlüsselzahlen zum Bildungswesen in Europa (EUROSTAT) und anderer Kennziffernsysteme (vgl. BLK, 2002).

Das umfangreichste Spektrum an vergleichbaren Bildungsindikatoren bietet die OECD-Publikation *„Bildung auf einen Blick – OECD Indikatoren“* (OECD, 2003). Darin enthalten sind Informationen zu den in Bildung investierten Ressourcen, zur Funktionsweise von Bildungssystemen sowie zu den Erträgen aus Bildungsinvestitionen. Für den Niveauvergleich von Schul- und Ausbildungsabschlüssen wird im Wesentlichen auf der Ebene der ISCED-Klassifikation gearbeitet. Dabei stehen Daten für den Bereich der beruflichen Bildung selbst für eine grobe Typisierung nur vereinzelt zur Verfügung (vgl. Van den Berghe, 1997; NCVER, 2004). So wird eine signifikante Menge an Indikatoren lediglich auf der Ebene des Bildungsbereiches (Sekundarstufe II) angegeben und nur vereinzelt finden sich Informationen auf der Ebene der Ausrichtung (Allgemeinbildung oder Berufsbildung – VET Sek. II/bzw. ISCED 3 A–C), da die Trennung zwischen

schulischen und beruflichen Bildungsabschlüssen Gefahren der Doppelzählung eröffnet.

Die Aktivitäten der UNESCO sind neben einer Reihe anderer Bildungsaufgaben darauf gerichtet, Indikatoren und hierzu erforderliche Daten für diejenigen Länder bereitzustellen, die durch die EU und die OECD nicht erfasst werden. Die zentrale Publikation ist der „*Global Education Digest*“ (UNESCO, 2004). Auch für diese Reihe gilt, dass sämtliche Daten auf einem hochaggregiertem Niveau (Sekundarstufe II) ausgewiesen werden und somit keine Informationen zur Qualität von Berufsbildungsprozessen enthalten.

Einen eher allgemeinen Überblick über die Berufsbildung in den europäischen Mitgliedsstaaten enthält die Reihe „*Schlüsselzahlen zur Berufsbildung in der Europäischen Union*“. Vorgestellt werden die verschiedenen Etappen und Modalitäten der Berufsbildung sowie ihre Teilnehmer nach verschiedenen sozioökonomischen Merkmalen. Auf der Grundlage dieser Zahlen, die zum größten Teil auf Verwaltungsdaten basieren, und zusätzlichen Erhebungen, von denen der größte Teil aus den Arbeitskräfteerhebungen der Gemeinschaft entstammt, beschäftigen sich weitere Publikationen mit spezifischen Themen wie dem Übergang von der Schule ins Erwerbsleben (vgl. Kogan & Müller, 2003; Mainguet, 1999; Europäische Kommission, 1997).

Auch wenn die hier beschriebenen Bildungsstatistikberichte kaum geeignete Daten für den internationalen Vergleich von beruflichen Bildungssystemen und -prozessen zur Verfügung stellen, leisten sie doch einen wichtigen Beitrag zur Ermittlung wesentlicher Kontextfaktoren schulischer Lernbedingungen, die als Konsens der Fachwelt zur Bewertung von Bildung im internationalen Vergleich angesehen werden dürfen (OECD, 2003). Das Hauptaugenmerk der UOE-Datensammlung gilt dabei politischen Gestaltungsparametern.

Zu den von der UOE erfassten Faktoren gehören: *Bildungsbeteiligung, Bildungszugang, Bildungserwartung, Bildungsstand, Abschlussquoten, Erwerbsbeteiligung und Ertragsraten für Bildungsinvestitionen, Lernumfeld und Lernorganisation, Bildungsausgaben, Grundlegende statistische Bezugsdaten (Ökonomische Kennzahlen)*.

– Arbeitskräfteerhebung der Gemeinschaft (AKE)

Die Arbeitskräfteerhebung der Gemeinschaft ist eine Haushaltserhebung mit Schwerpunkt auf personenbezogenen Variablen, die in der Hauptsache der Betrachtung allgemeiner Merkmale des Arbeitsmarktes in Bezug auf Beschäftigung, Arbeitslosigkeit und Wirtschaftszweige dienen. Sie deckt aber auch die wichtigsten persönlichen Merkmale der Befragten ab (Alter, Geschlecht, Familienstand, Staatsangehörigkeit) sowie zusätzliche Fragenkomplexe, die in Zusammenarbeit mit EUROSTAT europaweit harmonisierte Daten zur kürzlich erfolgten und laufenden Teilnahme an Bildungsmaßnahmen liefern. Allerdings kann im

Hinblick auf die Konstruktion der Abgänger derzeit nur bezogen auf Altersgruppen, nicht jedoch auf Bildungsabschlüsse mit Genauigkeit gearbeitet werden. Somit werden auch im Rahmen der AKE keine geeigneten Daten für den Vergleich beruflicher Bildung zur Verfügung gestellt. Die hier verwendeten Variablen ermöglichen jedoch die Bestimmung individueller Indikatoren, die für die Analyse von Praktiken im Zusammenhang mit dem Übergang von der Schule ins Berufsleben und in den Arbeitsmarkt sowie den weiteren Bildungsweg geeignet sind.

Zu den von der AKE erfassten Faktoren gehören: *Sozialer Hintergrund der Befragten, Teilnahme an Bildungsmaßnahmen, Erste signifikante Tätigkeit nach Beendigung der Ausbildung, Beschäftigtenstatus, Stellung im Beruf.*

– Europäische Erhebung über betriebliche Weiterbildung (CVTS)

Im Rahmen der europäischen Erhebungen zur betrieblichen Weiterbildung (CVTS 1 im Jahre 1994 und CVTS 2 in 2000/2001) wurden vergleichbare Daten zur beruflichen Weiterbildung in Unternehmen erhoben (vgl. Grünewald et al., 2003). Das Hauptziel der CVTS 3 liegt in der Vervollständigung der Informationen zur betrieblichen Bildung: Im Jahr 2006 soll erstmals neben der betrieblichen Weiterbildung die Erstausbildung in das Erhebungskonzept integriert werden (vgl. BIBB, 2005a). Damit könnte die CVTS in Zukunft auch Daten über betriebliche Ausbildungsprozesse liefern. Bislang beschränkt sich ihr Informationsgehalt jedoch auf betriebliche Weiterbildungsaktivitäten.

Zu den im Rahmen der CVTS erfassten Faktoren gehören: *Bestehende Angebote, Formen und Anbieter von Weiterbildung, Teilnahmequoten an Lehrveranstaltungen, Teilnahmestunden, Kosten und Finanzierung der Weiterbildung, Unternehmensgröße, Professionalisierungsgrad der betrieblichen Weiterbildung.*

– CEDEFOP – Ländermonographien

Die Länderberichte des Europäischen Zentrums für die Förderung der Berufsbildung (CEDEFOP) bemüht sich um die Darstellung von Fakten zur allgemeinen und beruflichen Bildung in den EG-Mitgliedstaaten. Dabei beschränken sich die Autoren auf eine deskriptive Beschreibung allgemeiner Eckwerte, die auf je unterschiedlichen Bezugsdaten beruhen (vgl. CEDEFOP [Hrsg.] – Länderberichte, 1999–2002). Eine vergleichende Darstellung, die auf einem spezifischen theoretischen und systematisch indikatorengestützten Rahmen beruht, fehlt in den Ländermonographien noch.

Die in allen Monographien enthaltenen Elemente sind: *eine Beschreibung der systemischen Kontextindikatoren (der soziale, demographische und wirtschaftliche Kontext, unter denen das Berufsbildungssystem existiert), die Entwicklungsgeschichte nationaler Berufsbildungssysteme, die Darstellung des Erstausbildungs-*

systems, Hinweise auf die Steuerung des Systems, Informationen über die Finanzierungsmechanismen.

Den bisher vorgestellten Untersuchungen ist gemeinsam, dass sie lediglich eine empirische Basis für den Vergleich der Struktur von Bildungssystemen liefern, jedoch weder eine Darstellung der Qualität von Bildungsprozessen noch der in ihnen vermittelten Kompetenzen. Bildungsabschluss dient als Hilfsindikator für Lernerfolg und persönliche Fähigkeiten. Aus diesem Grund wurden weitere Vergleichsstudien zur Messung von Kompetenzen unter dem Aspekt gesichtet, welche institutionellen Merkmale, die etwas über die Qualität von Lernbedingungen aussagen, verwendet wurden.

– PISA – Studie

Die in Dreijahreszyklen durchgeführte OECD-Studie „Programme for International Student Assessment“ (PISA) stellt die bislang umfassendste internationale Erhebung von Schülerleistungen dar, wobei sich jede Einzelstudie auf eine andere Kompetenzdimension bezieht (2000 Lesekompetenz, 2003 Mathematik, 2006 Naturwissenschaften).

Im Rahmen der PISA-Studien wird nur ein Teil der Kontextfaktoren schulischen Lernens mit Hilfe eigener Untersuchungsinstrumente erfasst. So werden die Schulsysteme der einzelnen Länder überwiegend anhand der Datensammlung des OECD-Projekts INES beschrieben und mit Hilfe internationaler wie nationaler Dokumente und Datensammlungen sekundäranalytisch unterstützt (vgl. OECD, 2003; Baumert et al., 1999). In Deutschland wurden – zusätzlich zu den international verwendeten Untersuchungsinstrumenten – Fragebögen für Schulleitungen, Schüler und Eltern eingesetzt. Dazu gehörten im Schulleiterfragebogen vor allem Fragen zur Qualitätssicherung und zur regionalen Kooperation (hier unter dem Stichwort „Sozialkapital der Schule“ verhandelt), im Schülerfragebogen Fragen zu Kontextbedingungen, die neben Schule und Elternhaus erhebliches Gewicht für die Entwicklung der Schüler haben sollen, z. B. Fragen nach den Beziehungen der Schüler zu Gleichaltrigen und zur Medienkompetenz. Das Hauptaugenmerk der internationalen Fragebögen liegt indes auf den personellen und materiellen Ressourcen der Schulen (Baumert et al., 1999).

Zu den im Rahmen von PISA erfassten Kontextindikatoren gehören: *Persönelle und materielle Ressourcen der Schule*. Im Deutschland wurden zusätzlich erhoben: *Schulorganisation, Lernangebote und ihre Nutzung, Qualitätssicherung, Sozialkapital der Schule (regionale Kooperationen), Unterrichtsqualität, Schulklima, sozioökonomische Lage, Migration, Familienstruktur, Erziehungsklima und häusliche Unterstützung, Peer-group-Einbindung, Freizeitaktivitäten.*

- Internationale Vergleichsstudien der „generic skills“ von Erwachsenen: International Adult Literacy Survey (IALS)/Adult Literacy and Lifeskills (ALL)

Der „Adult Lifeskills and Literacy Survey“ (ALL) basiert methodisch auf dem bereits 1994 abgeschlossenen International Adult Literacy Survey (IALS). Im Rahmen von ALL sollen die „generic skills“ der 16 bis 65 Jahre alten Bevölkerung ermittelt werden. Neben verschiedenen Grundkompetenzen (Verständnis von Prosatexten, von schematischen Texten und von mathematischen Problemen) geht es bei diesem Projekt auch um die Messung von Teamfähigkeit und den Umgang mit modernen Informationstechnologien. IALS und ALL können demnach als Pendant zu der Leistungsmessung von Schülern (PISA) im Erwachsenenbereich angesehen werden. Dabei können eine Reihe von Informationen zum Zusammenhang zwischen dem sozioökonomischen Hintergrund der befragten Personen und anderen Variablen (z. B. Weiterbildungsverhalten) gewonnen werden. Allerdings werden, anders als bei PISA, die bildungspolitischen Einwirkungsmöglichkeiten als begrenzt angesehen (vgl. BLK, 2002). Dies liegt insbesondere an den fehlenden Informationen über die Rahmenbedingungen und Prozesse auf der Ebene der Bildungseinrichtungen, in denen die jeweiligen Kompetenzen im Lebensverlauf erworben wurden.

Zu den im Rahmen von IALS/ALL erfassten Kontextfaktoren gehören:

Beschäftigtenstatus, Stellung im Beruf, Einkommen, Durchschnittliche Arbeitszeit, Bildungsstand, Bildungsstand der Eltern, Herkunft, Teilnahme an Weiterbildungsmaßnahmen, Art der wahrgenommenen Maßnahmen, Teilnahme an gemeinnützigen Aktivitäten.

- Programme for the International Assessment for Adult Competencies (PIAAC)

Die Konzeption für eine international vergleichende OECD-Studie zu den Kompetenzen von Erwachsenen (PIAAC) hat sich weit reichende Ziele gesetzt. Zu ihnen gehören die Definition und Messung von Kompetenzen, die einen persönlichen und wirtschaftlichen Erfolg der Bevölkerung erwarten lassen; die Erfassung des Beitrages verschiedener Kompetenzen für den sozialen und ökonomischen Outcome auf verschiedenen Ebenen; die Bewertung des Einflusses von Bildungssystemen auf die Entwicklung von „generic skills“ und die Identifikation möglicher politischer Maßnahmen zur Sicherung und Steigerung von Kompetenzen in der erwachsenen Bevölkerung (vgl. Green, 2004; Dolton, 2004; OECD, 2004).

Um dieses Ziel zu erreichen, sollen Informationen zu den Bedingungen des Erwerbs und der Verwertung von Kompetenzen erhoben werden. Hierzu gehören: *Informationen über die nationalen Arbeitsmärkte, sozioökonomische und demographische Daten, Teilnahme der Bevölkerung an formellen und informellen, betrieblichen und außerbetrieblichen Bildungsprozessen, Charakteristika des Arbeitsplatzes und des Unternehmens, Informationen über das „Sozialkapital“ der Bevölkerung (z. B. zivilgesellschaftliche Partizipation, Gesundheit).*

Da sich PIAAC noch in der Planungsphase befindet, sind genauere Informationen zu Messkennziffern für Kontextbedingungen zurzeit nicht verfügbar. Mit ersten Resultaten ist erst im Jahre 2010 zu errechnen, die Erhebungen sollen im Jahre 2008 durchgeführt werden (vgl. Green et al., 2004). Zurzeit ist man noch mit der Entwicklung entsprechender Instrumente beschäftigt.

Die tabellarische Übersicht „Internationale Quellen zur Messung von Ausbildungsbedingungen“ stellt noch einmal dar, auf welche Weise bestehende internationale Vergleichstudien die mikro- und makrostrukturellen Einflussfaktoren von Bildungsprozessen beschreiben und ob die dort verwendeten Indikatoren und Ergebnisse einen hinreichenden Gebrauchswert für eine internationale Vergleichsstudie von Berufsbildungssystemen haben. Ein Vergleich der in den verschiedenen Studien erfassten institutionellen und individuellen Faktoren zeigt, dass die folgenden Input- und Output-Indikatoren besonders häufig auftauchen:

- Auf der individuellen Ebene sollen in erster Linie Daten zum sozioökonomischen Hintergrund der Befragten und ggf. zu deren derzeitigen Beschäftigtenstatus und Stellung im Beruf erhoben werden. Diese Informationen sollen dann in Beziehung gesetzt werden zu den in der Ausbildung und im weiteren Biographieverlauf erworbenen Qualifikationen bzw. Kompetenzen.
- Daten zu den eingesetzten finanziellen, materiellen und personellen Ressourcen der Bildungsanbieter, die in ein Verhältnis gesetzt werden sollen zu Qualifikationsniveaus und -abschlüssen.

Tabelle 3.1.: Internationale Quellen zur Messung von Ausbildungsbedingungen

Quelle	Inhalt/Bereich	Erfasste Faktoren	Statistische Einheit	Länder	Stärken und Schwächen
UNESCO- OECD- Eurostat (UOE)- Datenbasis/ INES	Erfassung demographischer, sozialer, finanzieller und organisatorischer Indikatoren, die nahezu den gesamten Teil des formellen Bildungssystems abdecken	Bildungsbeteiligung, Bildungszugang, Bildungserwartung, Bildungsstand, Abschlussquoten, Erwerbsbeteiligung und Ertragsraten für Bildungsinvestitionen, Lernumfeld und Lernorganisation, Bildungsausgaben, Grundlegende statistische Bezugsdaten (Ökonomische Kennzahlen).	Programme Haushalt Individuum	OECD- Staaten	International vergleichbare Daten zu Teilnahme, Abschluss, Ausgaben und Ressourcenaufwand für Bildung Informationen zur beruflichen Bildung nur auf hohem Aggregationsniveau vorhanden
Arbeitskräfte- erhebung der Gemeinschaft (AKE)	Teilnahme der Zielbevölkerung an unterschiedlichen Arten von Bildung und Weiterbildung	Sozialer Hintergrund der Eltern bzw. Auszubildenden, Teilnahme an schulischer und beruflicher Bildung, Abgeschlossene schulische und berufliche Bildung, Erste signifikante Tätigkeit nach Beendigung der Ausbildung, Beschäftigtenstatus, Stellung im Beruf	Haushalt Individuum	Mitglieds- staaten der EU	Versucht, die gesamte Wohnbevölkerung abzudecken, Verknüpfung mit Arbeitsmarkstatus (ISEI) und anderen Informationen (Übergängen in Beschäftigung) möglich. Bildungsabschluss als Hilfsindikator für Lernerfolg und persönliche Fähigkeiten; im Hinblick auf die Konstruktion der Abgänger kann man derzeit nur bezogen auf Altersgruppen arbeiten
Europäische Erhebung zur beruflichen Weiterbildung (CVTS F+H)	Befragungen zur beruflichen Weiterbildung in Unternehmen	Bestehende Angebote, Formen und Anbieter von Weiterbildung, Teilnahmequoten an Lehrveranstaltungen, Teilnahmestunden, Kosten und Finanzierung der Weiterbildung, Unter-nehmensgröße, Professionalisierungsgrad der betr. Weiterbildung (strukturelle Rahmenbedingungen)	Unter- nehmen	Mitglieds- staaten der EU	Daten stammen ausschließlich von Unternehmen; nur Aktivitäten zur Verbesserung beruflicher Tätigkeiten werden berücksichtigt In 2006 soll die Erstausbildung in das Konzept integriert werden.
CEDEFOP – Länderberichte	Darstellung von Daten und Fakten zur allgemeinen und beruflichen Bildung in den EG-Mitgliedstaaten	Soziale, demographische und wirtschaftliche Kontextbedingungen, unter denen das Berufsbildungssystem existiert, Struktur des Erstausbildungssystems, Steuerung und Finanzierung des Systems	System, Programme	Mitglieds- staaten der EG	Bereitstellung grundlegender Informationen über die Kontextbedingungen nationaler Berufsbildungssysteme
PISA	Untersuchung zur Effektivität nationaler Bildungssysteme (Schulen) in der Kompetenzvermittlung Berücksichtigt: Lesefähigkeit, mathematische und naturwissenschaftliche Kompetenzen	Ressourcen, Schulorganisation, Lernangebote und ihre Nutzung, Qualitätssicherung, Unterrichtsqualität, Schulklima, Sozioökonomische Stellung und kulturelles Kapital der Familie, Freizeitaktivitäten und peers	Individuum	32 Länder, von denen 28 der OECD angehören	International vergleichbare Daten bzgl. des Leistungsstandards zum Ende der Pflichtschulzeit Berücksichtigt nur die 15-Jährigen
IALS/ALL	IALS: Untersuchung zur Schreib- und Lesefähigkeit sowie den Rechenkenntnissen der Erwachsenen; ALL: Versuch, die Fähigkeit zur Partizipation am gesellschaftlichen Leben (in einem umfassenden Sinne) zu dokumentieren	Beschäftigtenstatus, Stellung im Beruf, Einkommen, Durchschnittliche Arbeitszeit, Bildungsstand, Bildungsstand der Eltern, Herkunft, Teilnahme an Weiterbildungsmaßnahmen, Art der wahrgenommenen Maßnahmen, Teilnahme an gemeinnützigen Aktivitäten	Individuum	Versch. OECD- Staaten	International vergleichbare Daten zur Schreib- und Lesefähigkeit (IALS), zu Rechenkenntnissen, Problemlösungsfähigkeiten, Teamarbeit, Medienkompetenz (ALL)
PIAAC	Konzept für eine international vergleichende Studie zur Messung von Kompetenzen von Erwachsenen und des Lebenslangen Lernens	Insbesondere Informationen über die Teilnahme der Bevölkerung an formellen und informellen/nicht-formellen Bildungsprozessen und das „Sozialkapital“ der Erwachsenenbevölkerung	Individuum	Versch. OECD- Staaten	International vergleichbare Daten zu den Kompetenzen und Lernbiographien der Erwachsenenbevölkerung Noch im Konzeptualisierungsstadium, deswegen noch keine Indikatoren

Vgl. BLK, 2002.

– Fazit: Forschungsdesiderate

Unter unserer Betrachtungsperspektive der Bereitstellung von Informationen zum Vergleich institutioneller Arrangements in der beruflichen Bildung und der in ihnen erworbenen Kompetenzen weisen die analysierten Studien spezifische Mängel auf, die Hinweise auf Forschungsdesiderate geben:

So fehlen etwa harmonisierte Daten zum Umfang einzelner Ausbildungsprogramme auf der Berufsfeldebene, zum Verhältnis theoretischer und praktischer Ausbildungsanteile, zum Angebot-Nachfrage-Verhältnis, zum Finanzierungsvolumen, zu Qualitätsstandards oder zum Übergang von beruflicher Ausbildung in den Beruf. Dies liegt im Wesentlichen daran, dass die bestehenden Daten auf einem so hohen Aggregationsniveau ausgewiesen werden, dass entsprechende Informationen in den meisten Fällen nur auf der Ebene von Bildungsbereichen (Primär-, Sekundär- und Tertiärbereich) erhältlich sind. Damit werden in kaum einer der vorgestellten Studien die Besonderheiten der beruflichen Erstausbildung adäquat berücksichtigt (vgl. NCVER 2004). Die im Rahmen der internationalen Berufsbildungs-PISA-Workshops eingeladenen Experten teilten einvernehmlich diese Einschätzung.

Des Weiteren konzentriert sich der überwiegende Teil der einschlägigen Quellen darauf, Strukturdaten zu erheben und zu präsentieren. Die dahinter stehenden qualitativen Prozesse sind so nicht zu erfassen und stellen die große Forschungslücke dar: In den auf Input-Output konzentrierten Untersuchungen werden Lernorte und Lernumgebungen als „black box“ begriffen (vgl. Scheerens 2004). So fehlen etwa Daten zur Unterrichtsqualität und den zum Einsatz kommenden Lehr-Lernmethoden. Eine Ausnahme bildet hier lediglich die Erhebung von Kontextdaten in der PISA-Studie, welche dabei helfen, die durch die hochaggregierte Strukturdaten gelieferten Informationen auf einem niedrigeren Aggregationsniveau zu erweitern und zu vertiefen.

Schließlich geben die vorliegenden Berichtssysteme wenig Auskunft darüber, mit welchen Kompetenzen die Jugendlichen aus dem Berufsbildungssystem austreten. Kompetenz wird allein in der Form des höchsten erreichten formalen Schulabschlusses erhoben. Nicht erhoben wird jedoch, was sich hinter diesen Zertifikaten verbirgt. Dort wo Kompetenzen in der Erwachsenenbevölkerung erhoben werden (z. B. ALL oder IALS), liegen wiederum praktisch keine Erkenntnisse über die Auswirkungen verschiedener Ausbildungsstrukturen auf deren Entfaltung vor.

Eine Vergleichsstudie zur Berufsbildung würde damit erstmals international vergleichbare Daten sowohl über tatsächlich vermittelte Kompetenzen in der Berufsbildung als auch über die kontextuellen Merkmale des Berufsbildungssystems und der Verwertungsseite erworbener Kompetenzen erheben und diese aufeinander beziehen können. Auf diese Weise sollte eine internationale Vergleichsstudie eine aussagekräftige Prüfung der „Produktivität“ unterschiedlicher institutioneller Arrangements ermöglichen und so einen hohen

bildungspolitischen Nutzen bieten. Pascaline Descy (CEDEFOP) hat diesen Tatbestand während des ersten internationalen Berufsbildungs-PISA-Workshops nachdrücklich hervorgehoben, jedoch auch darauf hingewiesen, dass aufgrund der vorhandenen statistischen Daten in den einzelnen europäischen Ländern ein Vergleich der einzelnen Systeme nahezu unmöglich sei. Dies ist vor allem auf die Tatsache zurückzuführen, dass die Berufsausbildungssysteme innerhalb der verschiedenen Staaten durch eine Vielfalt nationaler Besonderheiten gekennzeichnet sind, die eine Systematisierung nur schwer ermöglichen (vgl. CEDEFOP, 1998).

3.2. INDIKATOREN ZUR BESCHREIBUNG INSTITUTIONELLER UND INDIVIDUELLER BEDINGUNGEN VON BERUFSPROZESSEN – ZU EINEM MODELL VON AUSBILDUNGSPRODUKTIONS- UND -VERWERTUNGSVARIABLEN

Haben wir bis jetzt untersucht, was in internationalen Vergleichsstudien an Indikatoren zur Bestimmung von (Berufs-)Bildungsqualität zu finden ist, geht es im Folgenden darum, einzelne Indikatoren systematisch zu prüfen und in einem möglichen Modell zur Erfassung von Unterschieden in der Kompetenzproduktion und -verwertung zu verdichten.

Internationale Vergleichsstudien von Bildungssystemen und -prozessen gehen übereinstimmend davon aus, dass sowohl institutionelle als auch individuelle Kontextbedingungen die Qualität von Lernprozessen und deren Ergebnisse wesentlich beeinflussen. Lernen wird dabei in einem umfassenden Sinne als „persönliche Entwicklung in Interaktion mit gegenständlichen Umwelten und mit sozialen Milieus“ (Lempert, 1979, S. 88) betrachtet, so dass einseitig die Persönlichkeit des Lernenden bzw. die Umwelt betonende Aspekte abgelehnt werden. Institutionelle und individuelle Einflussfaktoren interagieren in den realen Ausbildungsprozessen. Erst wenn man beide Einfluss-Ebenen zueinander in Beziehung setzen kann, dürfte man auch das Gewicht der Faktoren näher bestimmen können.

Hierzu durchgeführte Untersuchungen haben als „Input-Output-Studien“ oder „Produktivitäts-Studien“ Eingang in die wissenschaftliche und politische Diskussion gefunden (vgl. Scheerens, 2004; OECD, 2003). Die ihnen zugrunde liegende Vorstellung von Bildungseinrichtungen als „Produktionsstätten“ von Qualifikationen findet ihren Ausdruck in der Modellierung von Variablen-Zusammenhängen als Input-Output-Beziehungen einer „Bildungsproduktionsfunktion“. Die nachstehende Gleichung stellt eine typische Spezifikation einer solchen Produktionsfunktion (vgl. Weiß & Preuschoff, 2004) dar. Danach ist das zu einem bestimmten Zeitpunkt erhobene Wissen eines Schülers (**Wit**) eine Funktion des kumulativen Einflusses von Faktoren seines familiären Hin-

tergrunds (**Fit**), der Peer-Group (**Pit**); schulischer Faktoren (**Sit**) und seiner Begabung (**Ii**):

$Wit = f(\text{Fit}, \text{Pit}, \text{Sit}, \text{Ii})$.

Die Analyse solcher Funktionen soll Informationen über die Richtung und den spezifischen Erklärungsbeitrag der einzelnen unabhängigen Variablen (Prädiktoren) über die Varianz beim Kriterium (Output-Variable, abhängige Variable) und die Signifikanz der geprüften Variablen-Beziehungen liefern.

Von besonderem politischen Interesse beim Produktionsfunktions-Ansatz sind solche Zusammenhänge, die Hinweise für einen effizienten Einsatz von Ressourcen liefern. Signifikante Zusammenhänge zwischen Bildungsausgaben und Schülerleistungen werden in verschiedenen Meta-Analysen (vgl. Hanushek, 1989; 1997) jedoch verworfen: „With over three decades of analysis, new studies have reinforced earlier conclusions: (...) there is no strong or consistent relationship between variations in school resources and student performance“ (Hanushek 1997, S. 141). Mit Blick auf amerikanische Untersuchungen weisen Weiß und Preuschoff (2004, S. 12) zwar auf den Effekt einer Veränderung der finanziellen Ressourcen hin, der aufzeige, dass eine Erhöhung der Ausgaben je Schüler mit einem Leistungszuwachs verbunden ist. Allerdings räumen sie ein, dass in diesen Studien durchgehend unklar bleibe, welche konkrete Mittelverwendung sich hinter dem Ausgabeneffekt verbirgt.

Die kritischen Feststellungen verweisen auf Theoriedefizite des Produktionsfunktions-Ansatzes. Vor allem der Sachverhalt, dass in den beschriebenen Wirkungsanalysen Merkmale der schulischen Ressourcenausstattung als *direkte Wirkfaktoren* konzeptionalisiert werden, erscheint problematisch. Von „distalen“, d. h. vom Unterrichtsgeschehen entfernten Bedingungs-Variablen ist jedoch theoretisch kein direkter Einfluss auf Schulleistung zu begründen, sondern nur ein indirekter, über „proximale“, d. h. prozessuale Bedingungsfaktoren, vermittelter Einfluss: das Angebot an Lerngelegenheiten und deren Nutzung durch die Schülerinnen und Schüler, die Qualität des Unterrichts etc. (vgl. Weiß & Preuschoff, 2004; Scheerens, 2004; Baumert et al., 2001). Neuere Wirkungsstudien, die den Einfluss von Bildungsausgaben zusammen mit differenten Verwendungsmöglichkeiten untersuchen, können dann auch zum Teil solche Zusammenhänge bestätigen (vgl. Weiß & Preuschoff, 2004).

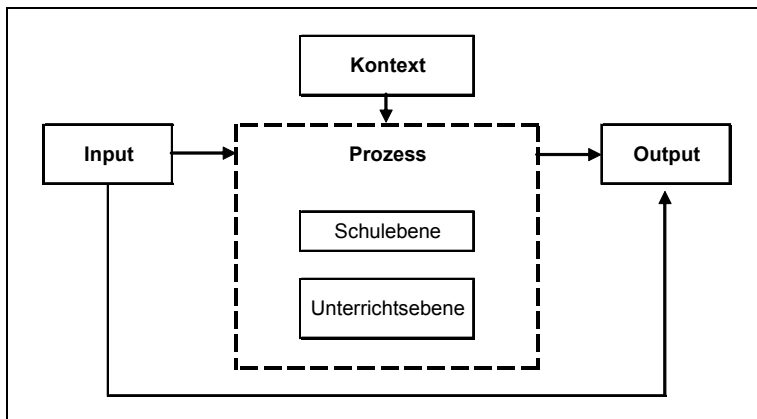
Auf einen Leistungseinfluss von Unterschieden in der Nutzungsintensität der in den Bildungseinrichtungen zur Verfügung stehenden Ressourcen verweisen die internationalen PISA-Befunde ebenso wie die früher im Rahmen der Lehr-Lern-Forschung durchgeführten Untersuchungen zur Leistungswirksamkeit unterschiedlicher Instruktionsqualität (vgl. u. a. Treiber, 1982).

Demnach sollte sich ein Berufsbildungs-PISA neben einer Darstellung wesentlicher Input- und Output-Variablen insbesondere darauf konzentrieren, die „black box“ der Lernorte, Lernumgebungen und Lernprozesse zu öffnen, und so herausfinden, was tatsächlich in beruflichen Bildungsprozessen passiert. Die Notwendigkeit einer solchen Konzeptualisierung wurde in den internationalen Workshops von allen Bildungsexperten hervorgehoben. Die Befunde zum Zusammenhang von Kontextbedingungen mit den Ergebnissen der Kompetenzmessung sollen dabei helfen, das Verständnis von Leistungen in beruflichen Bildungsprozessen zu vertiefen und zu bildungspolitischen Anregungen führen.

Des Weiteren konnte man Einigkeit darüber erzielen, dass die Input-, Prozess- und Outputvariablen möglichst auf vier Ebenen und deren möglichen Zusammenhängen zu konzeptualisieren sind:

- A) Die System- bzw. Makroebene der nationalen Institutionen und Organisationen,
- B) die Mesoebene der Bildungseinrichtungen (Schule und Betrieb),
- C) die Mikroebene der in den Bildungseinrichtungen stattfindenden Interaktionsprozesse (in der Schulklasse bzw. am Arbeitsplatz) und schließlich
- D) die Ebene der Individuen (Auszubildenden).

Tabelle 3.2: Bildungsproduktionsmodell nach Scheerens



Quelle: Scheerens, 1990.

Dabei bilden die Systemmerkmale keine Grundlage für ein Benchmark nationaler Berufsbildungssysteme, da die interne Heterogenität beruflicher Bildungssysteme in und zwischen den Ländern keine Vergleichbarkeit auf dieser Ebene zulässt (vgl. CEDEFOP, 1998). So hat sich beispielsweise das deutsche Berufsbildungssystem bis heute so stark ausdifferenziert, dass man nicht mehr von *einem* Berufsbildungssystem sprechen kann. Systemmerkmale (z. B. Steuerungs- und Finanzierungsmodalitäten, rechtliche und administrative Vor-

gaben etc.) sollen lediglich die systemweiten Informationen liefern, die zur Beschreibung und Erklärung spezifischer Merkmale der Ausbildungssektoren und der darunter liegenden institutionellen Ebenen (Schule/Betrieb bzw. Unterricht/Arbeitsplatz) notwendig sind.

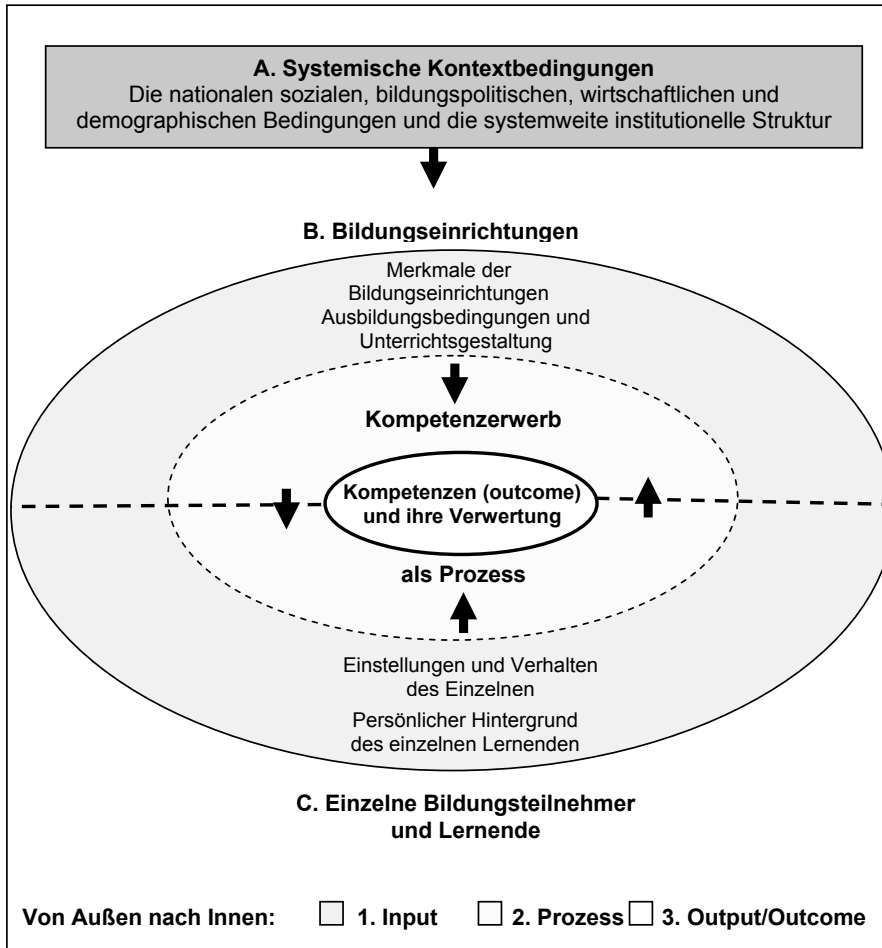
Bei der Beschreibung der individuellen und institutionellen Bedingungen der Vermittlung und Verwertung von Kompetenzen unterscheiden wir zwischen folgenden vier Aspekten:

- 1) Institutionelle Rahmenbedingungen und Vorgaben für das Bildungssystem als Ganzes (Kontext-Ebene; politische Steuerung);
- 2) Rahmenbedingungen und Vorgaben auf der Ebene einzelner Bildungseinrichtungen sowie der persönliche Hintergrund der Lernenden (Input-Ebene);
- 3) Prozesse institutionalisierter Bildung, das heißt die Ebene der Umsetzung materieller (technisches Equipment, Unterrichtsmaterialien), personeller (Qualifikation der Lehrkräfte, Wissen und Einstellungen der Auszubildenden) und symbolischer Ressourcen (Lehrpläne u. a.) in Unterrichts- bzw. Ausbildungsgestaltung (Prozess-Ebene);
- 4) Ergebnisse von Bildungsprozessen auf der Ebene der Zertifikate und der Verwertung von Kompetenzen in der Arbeit, auf den Arbeitsmärkten und in der Biographie (Outcome- bzw. Output-Ebene).

Die dem Berufsbildungs-PISA zugrunde liegende Konzeption der Zusammenhänge zwischen institutionellen und individuellen Bedingungen im Hinblick auf den Erwerb von Kompetenzen und deren Verwertung lässt sich schematisch wie folgt darstellen:

Für die Anlage und den Aufwand der Studie gilt es im Folgenden herauszuarbeiten, welche institutionellen und individuellen Faktoren auf den verschiedenen Input-, Prozess- und Output-Ebenen als zentral angesehen werden können, um beobachtbare Unterschiede in der Kompetenzentwicklung und -verwertung zu erklären, und über welche Indikatoren sie zu messen sind. Wir greifen dabei die wichtigsten Indikatoren der internationalen Vergleichsstudien (Kapitel 3.1.) auf und stellen sie in den Zusammenhang des Produktions- und Verwertungsmodells von Kompetenzen (s. Abbildung 3.1) und ergänzen sie um solche Aspekte, welche die Lücke zwischen fehlenden Informationen zur beruflichen Erstausbildung und den dort ablaufenden Prozessen schließen. Im Folgenden werden, ausgehend von bereits vorhandenen Untersuchungen, für jede der verschiedenen Ebenen Indikatoren zur Beschreibung verschiedener Input- und Prozess-Variablen vorgestellt.

Abbildung 3.1: Zusammenhänge institutioneller und individueller Ausbildungsbedingungen auf den Erwerb von Kompetenzen und deren Verwertung



3.2.1. Das Bildungs- und Beschäftigungssystem als Ganzes (Makroebene)

Die Notwendigkeit einer Erhebung von Informationen über das Bildungs- und Beschäftigungssystem als Ganzes beruht auf der Annahme, dass die im Rahmen eines Berufsbildungs-PISA zu vergleichenden Formen beruflicher Erstausbildung nur vor dem Hintergrund der jeweiligen sozialen, bildungspolitischen, wirtschaftlichen und demographischen Bedingungen eines Landes zu verstehen sind. Wie bereits weiter oben erwähnt, bilden solche Informationen

keine Grundlage für ein Benchmark nationaler Berufsbildungssysteme. Die hierunter rubrizierten Informationen und hochaggregierten Strukturdaten sollen letztlich dabei helfen, die Informationen auf einem niedrigeren Aggregationsniveau zu vertiefen und in ihrem makrostrukturellen Zusammenhang interpretieren zu können.

Auf der Makroebene des Bildungs- und Beschäftigungssystems als Ganzes unterscheiden wir zwischen den folgenden Kategorien:

A. Das Bildungs- und Beschäftigungssystem als Ganzes

- | |
|--|
| <ol style="list-style-type: none">1. Systemische Kontextbedingungen2. Systemweite institutionelle Struktur<ul style="list-style-type: none">– Steuerung und Koordination– Standards und Normen |
|--|

A1) Systemische Kontextbedingungen

Die Abbildung systemischer Kontextbedingungen verweist auf den Tatbestand, dass Berufsbildungsprozesse eingebettet sind in anderen gesellschaftlichen Subsystemen. Diesbezüglich sind verschiedene soziale, kulturelle, wirtschaftliche und politische Bedingungen zu identifizieren, die einen indirekten Einfluss auf das System der beruflichen Bildung ausüben. Beispielhaft zu nennen wären hier (möglicherweise nach Regionen zu differenzierende) nationale Daten zur aktuellen wirtschaftlichen Wachstumsrate, die Lage der öffentlichen Haushalte und die Steuerpolitik, demographische Entwicklungen oder kulturelle Faktoren wie die Einstellung zu Migranten (vgl. Lassnigg, 2003).

A 2) Systemweite institutionelle Struktur

Zur Beschreibung der Prozesse auf der Makroebene des Bildungs- und Beschäftigungssystems würden wir momentan drei Indikatoren vorsehen: Die Steuerung und Koordination des Berufsbildungssystems, Standards und Normen (inklusive Curricula) und die Finanzierung der Ausbildung.

Steuerung und Koordination: Eine entscheidende Aufgabe der Steuerung und Koordination von Berufsbildungssystemen besteht darin, ihre Anpassungsfähigkeit an den Bedarf des Einzelnen, des Arbeitsmarktes und der Gesellschaft insgesamt zu gewährleisten (vgl. Descy & Tessaring, 2002). Darauf bezogene institutionelle Faktoren sollen Aufschluss darüber geben, welche

Steuerungs- und Koordinationsformen diese Anpassungsfähigkeit am besten sicherstellen.

Ein zentraler Aspekt der Analyse ist die Beschreibung des Koordinationssystems und der in ihm ablaufenden Abstimmungsprozesse zwischen Anforderung und Nachfrage (vgl. Lassnigg, 2000). Dabei geht es uns primär um das Verhältnis zwischen den internen Akteuren des Bildungssystems und seinen wichtigsten externen „Stakeholders“, der Wirtschaft und den Vertretern der Ausbildungsadressaten. Eine mögliche Differenzierung im Hinblick auf die strukturelle Kopplung von Bildungs- und Beschäftigungssystem bieten Descy und Tessaring (2002) an, die zwischen engen, losen und variablen Kopplungen unterscheiden (vgl. Tabelle 3.3).

Tabelle 3.3: Beziehungen zwischen Bildungs- und Beschäftigungssystem

Art der Beziehung	Typische Länder	Merkmale
Enge Beziehung	Deutschland, Dänemark, Niederlande, Österreich, Tschechische Republik, Ungarn	Ausdifferenziertes Bildungssystem, mit der Arbeitswelt verknüpft (z. B. Praktika, Lehrlingsausbildung), relativ hohe Übereinstimmung von Qualifikationen und Berufen, Berufsausbildung als Voraussetzung für den Eintritt ins qualifizierte Erwerbsleben, hohe Stellung von Zeugnissen bei Arbeitgebern.
Lose Verknüpfung	Australien, Japan, Kanada, Vereinigte Staaten	Bildungssystem ist in der Regel schulbasiert mit breit angelegter allgemein beruflicher Bildung; die spezielle berufsfachliche Ausbildung findet vorrangig am Arbeitsplatz statt und zwar im Anschluss an den Bildungsgang, flexible Übereinstimmung von Qualifikationen und Berufen/ keine direkte Verbindung zwischen Ausbildungsfach und besetzter Stelle, niedrigerer Stellenwert von Berufsbildungszertifikaten im Vergleich zu allgemein bildenden Abschlüssen.
Variable Beziehung (vom einzelnen Bildungsgang abhängig)	England, Estland, Finnland, Frankreich, Griechenland, Norwegen, Portugal, Schottland, Schweden, Spanien	Gute Übereinstimmung von Qualifikationen und Berufen im Bereich der Lehrlingsausbildung bzw. einer spezifischen berufsfachlichen Ausbildung, weniger hingegen bei Vollzeitbildungsgängen, Probleme beim Übergang vom Bildungssystem zum Arbeitsmarkt aufgrund des stark ausdifferenzierten beruflichen Bildungswesens.

Quelle: Descy & Tessaring (2002).

Eine andere, auf die politische Steuerung zielende Differenzierungsmöglichkeit bietet Greinert (1995) an, der zwischen drei Steuerungsmodellen unterscheidet: dem liberal-marktwirtschaftlichen, dem etatistisch-bürokratischen und dem dual-korporatistischen. Die duale Berufsausbildung in Deutschland dient als Beispiel für ein korporatistisch gesteuertes Ausbildungssystem, in

welchem die Sozialparteien maßgeblich die qualitativen Standards für den betrieblichen Teil der Ausbildung bestimmen, die zugleich Standards für den schulischen Teil vorstrukturieren. Die Berufsausbildung in Frankreich gilt als ein staatlich gesteuertes schulbasiertes Ausbildungssystem, in denen die Standards durch das Bildungsministerium festgelegt werden. Die Sozialparteien haben lediglich eine konsultative Rolle bei der Strukturierung der Ausbildungsabschlüsse. Die Ausbildung für nationale berufliche Qualifikationen (NVQ) in England gilt dagegen als Fall eines weitgehend marktgesteuerten Ausbildungssystems. Um die Transparenz des Angebots auf dem Bildungsmarkt zu verbessern, erkennt der Staat weitgehend durch die Arbeitgeberseite entwickelte nationale Qualifikationsstandards an (vgl. Koch & Reuling, 1998).

Man kann erwarten, dass unterschiedliche Steuerungsformen die qualitativen und quantitativen Outcomes von Bildungssystemen mehr oder weniger stark beeinflussen. Als die angemessenste Form der Steuerung wird eine Verbindung von staatlicher Regulierung und marktwirtschaftlichen Prinzipien angesehen, die gleichzeitig ökonomischen Interessen und weiter gehenden sozialen und gesellschaftlichen Zielen Rechnung trägt (Descy & Tessaring, 2002). So ist aus internationalen Vergleichstudien ersichtlich, dass sich die leistungsfähigsten Bildungs- und Ausbildungssysteme in Europa durch ein hohes Maß an Verflechtung und Koordination zwischen Staat, Bildungs- und Ausbildungsanbietern und Sozialpartnern auszeichnen.

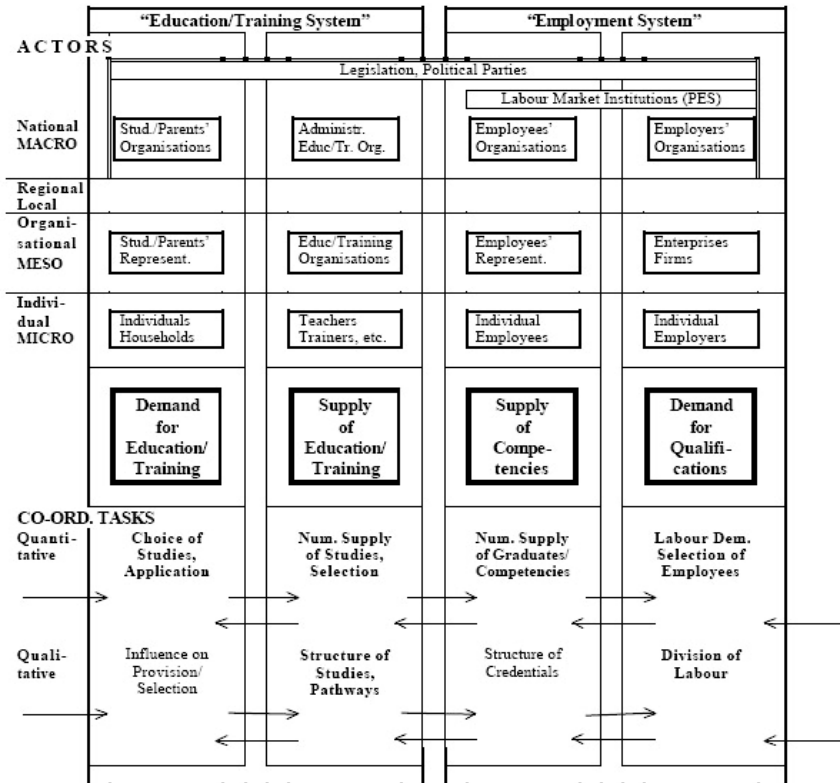
Die Differenzierung in Governance-Typen hat jedoch nur Modellcharakter: Alle vorhandenen Systeme stellen Varianten und/oder Kombinationen der drei Grundformen dar. Als eine Annäherung zur Beschreibung der stattfindenden Mechanismen bietet Lassnigg (2000) ein heuristisches Schema an, mit dessen Hilfe Koordinationsvorgänge und deren Interaktion beschrieben werden können und das für internationale Vergleichsstudien verwendet werden kann. Die folgende Darstellung zeigt das Beziehungsgeflecht zwischen den Akteuren, die auf unterschiedlichen Aggregationsniveaus jeweils nach Bildungs- und Beschäftigungssystem getrennt und nach Angebots- und Nachfrageseite differenziert werden können (s. Tabelle 3.4). Quantitativ geht es hierbei um Übergänge in Form von Selektions- und Kapazitätsentscheidungen, qualitativ vor allem um die Festlegung bzw. Veränderung der Ausbildungsstruktur und -profile.

Hinweise auf eine erfolgreiche Koordination und Steuerung des Berufsbildungssystems in quantitativer Hinsicht können makrostrukturelle Daten zu den institutionellen Ergebnissen (Outputs) liefern. Hierunter fallen insbesondere Abschluss- bzw. Abbrecherquoten¹⁶ in der beruflichen Erstausbildung und die

16 Abschlussquoten zeigen den gegenwärtigen Output der Ausbildungssysteme, d. h. den Prozentsatz derjenigen Schüler, die eine berufliche Erstausbildung erfolgreich abgeschlossen haben, bezogen auf die Gesamtpopulation im typischen Abschlussalter. Abschlussquoten sind in erster Linie ein Output-Indikator dafür, inwieweit es den Bildungssystemen gelingt, den Anforderungen des Arbeitsmarktes an qualifizierten Arbeitskräften

Gestaltung der Übergänge in den Beruf (z. B. Jugendarbeitslosigkeitsquoten). Governance-Typen können zur Auswahl oder makrostrukturellen Einordnung von in das Berufsbildungs-PISA einbezogenen Länder dienen.

Tabelle 3.4.: Akteure und Koordinationsmechanismen



Quelle: Lassnig (2000).

Standards und Normen: Allgemein verbindliche Standards für die Gestaltung der Ausbildungsqualität durch die einzelnen Bildungseinrichtungen können bestehen in Input-Standards für einzelne Qualitätsfaktoren, darunter fallen

zu entsprechen (vgl. OECD, 2003, S. 23). Abbrecherquoten sagen etwas über ein Missverhältnis zwischen institutionellen Angeboten und individuellen Bedürfnissen und Fähigkeiten aus, seien es solche zwischen beruflicher Angebotsstruktur und individuellen Wünschen, zwischen Ausbildungsformen und individuellen Wünschen, oder zwischen Ausbildungsformen und individuellen Ausbildungsbedingungen bzw. Ausbildungsveroraussetzungen. Sie signalisieren dabei Probleme des Übergangs von der Schule in die Ausbildung.

u. a. Standards bezüglich der Ausbildungsorganisation und Ausbildungsinhalte/Curricula (s. u.), Handlungsbedingungen des Lehrpersonals (z. B. erforderliche Abschlüsse, Weiterbildung, Tarifverträge über Arbeitsbedingungen, aber auch Verhaltensnormen), Verteilung von Ressourcen oder/und Output-Standards für Lernergebnisse und Prüfungsanforderungen. Die öffentliche Qualitätssteuerung auf der systemischen Ebene ist abzugrenzen von den hier nicht näher betrachteten trägerinternen Verfahren der Qualitätssicherung.

Eine zentrale Funktion von öffentlicher Qualitätssteuerung besteht darin, Vertrauen in die Qualität der Ausbildung zwischen allen Beteiligten herzustellen: „Die Jugendlichen sollen sicher sein können, dass die ihnen angebotene Ausbildung den öffentlich festgelegten Qualitätsnormen mindestens entspricht. Die Arbeitgeber sollen auf die Aussagekraft der Ausbildungsabschlüsse vertrauen können, was zugleich eine wesentliche Voraussetzung für die Funktionsfähigkeit des Erstbeschäftigungsmarktes ist. Die Einrichtungen des Bildungssystems sollen sich darauf verlassen können, dass den mit einem Abschluss verbundenen formalen Berechtigungen auch die entsprechenden Kompetenzen gegenüberstehen“ (Koch & Reuling, 1998, S. 7).

Hinsichtlich der Anordnung von Ausbildungsinhalten und deren Gestaltung lassen sich nach Lassnigg (2000) folgende Dimensionen unterscheiden:

- Eine erste Dimension betrifft die Art und Weise, wie Qualifikationselemente und Inhalte des Lernens konzeptualisiert werden. Die wichtigste Unterscheidung betrifft die traditionelle Aufteilung in schulische Gegenstände gegenüber spezifisch berufsbildenden Qualifikationselementen (Basic Job Oriented oder Generic Skills, vgl. Kapitel 2). Das Grundmuster der Aufteilung der allgemeinen und spezifischen Qualifikationselemente kann dabei eher „nach dem Modell der ‚Entkoppelung‘ von allgemeiner und spezifischer Bildung oder nach dem Modell der Verschränkung entsprechend der Lehrlingsausbildung“ (Lassnigg, 2000, S. 25) gestaltet werden. Eine weitere Differenzierungsmöglichkeit bietet die Gewichtung theoretischer und praktischer Ausbildungsanteile. Wie sich diese Gewichtung der allgemeinen Strukturmerkmale in den tatsächlichen Bildungsprozessen niederschlägt, ist auf den Ebenen der Institutionen zu ermitteln.
- Die zweite Dimension betrifft die Gestaltung von längeren und komplexen Ausbildungsgängen oder deren Zerlegung bzw. Sequenzierung in Form von flexiblen modularen Systemen (Modularisierung).
- Die dritte Dimension betrifft die Beziehung von Bildungsprozessen (Curricula) und Abschlüssen (Zertifikaten), die verbunden oder getrennt konzipiert sein können, sowie – in engem Zusammenhang damit – die Art der den Abschlüssen zugrunde liegenden Standards.
- Die vierte Dimension betrifft die vertikale und horizontale Verknüpfung von Bildungsgängen innerhalb des Berufsbildungssystems und im Verhältnis zur vorberuflichen Allgemein- zur und Hochschulbildung.

Besondere Bedeutung kommt der Normierung der Lehrertätigkeit zu. Hier geht es um die unterschiedlichen Qualifikationsanforderungen an das Lehrpersonal und die berufliche Weiterbildung. Es wird davon ausgegangen, dass sowohl die Erstausbildung als auch die berufliche Weiterbildung der einzelnen Lehrkräfte sich auf deren Fähigkeit auswirken, einen qualitativ guten Unterricht durchführen zu können (vgl. OECD, 2003).

Finanzierung: Bildungsausgaben können als Ausdruck sowohl für die Wertschätzung von Ausbildung durch die Anbieter als auch für die finanziellen Belastungen unterschiedlicher Akteursgruppen der Berufsbildung angesehen werden. Bildungsausgaben können im Allgemeinen über die Ausgaben pro Schüler für öffentliche wie private Bildungseinrichtungen gemessen wie auch als Anteil am BIP pro Kopf ermittelt werden. Unterschiede in den Bildungsausgaben pro Schüler dürfen dabei nicht nur auf Unterschiede bei den materiellen Ressourcen zurückgeführt werden, sondern auch auf Unterschiede in den relativen Gehaltsniveaus: Die Gehälter für Lehrkräfte machen den Hauptteil der Gesamtkosten aus (OECD, 2003). Die Verteilung der Finanzen kann einen indirekten Einfluss auf die Qualität des Lernumfeldes und der Lernbedingungen in der Ausbildung haben, wobei es wichtig wäre, finanzielle Inputs nach Ausgabearten differenzieren zu können. Eine weitere Differenzierungsdimension bildet der Finanzierungsmodus, der nach den verschiedenen Trägern der Netto-Ausbildungskosten fragt.

3.2.2. Bildungseinrichtungen

Die Indikatoren auf der Ebene der Einrichtungen beziehen sich auf die strukturellen und organisatorischen Bedingungen, unter denen Schulen und Betriebe ihre Aufgaben als Bildungseinrichtungen erfüllen. Formen des Lehrens und Lernens werden dadurch mehr oder weniger präformiert (vgl. Kunter et al., 2002).

B. Bildungseinrichtungen

Input

Merkmale und Organisation der Bildungseinrichtungen

- Ressourceneinsatz
- Lernortkooperation
- Qualitätskontrollen

Prozess

Ausbildungsbedingungen und Unterrichtsgestaltung

- Gestaltung der Ausbildungsprozesse
- Lernklima

B 1) Merkmale und Organisation der Bildungseinrichtungen (Input)

In diesem Komplex sollen auf der Ebene der Bildungseinrichtungen die Bedingungen des Unterrichts dargestellt werden. Die Informationen sind im Rahmen einer Befragung der Schulleiter bzw. der Verantwortlichen für die betriebliche Ausbildung zu gewinnen.

Ressourceneinsatz: Dieses Faktorenbündel soll Auskunft geben über die Verteilung und die Art des Einsatzes von Mitteln (bspw. Personal- und Sachaufwand), über die Schulen bzw. Betriebe bei der Ausbildung der Schüler verfügen. Hinsichtlich der Einsatzes *personeller Ressourcen* gilt das zahlenmäßige Schüler/Lehrerverhältnis als eine (wenn auch kontrovers diskutierte) indirekte Kennziffer zur Messung von Vermittlungsqualität (vgl. Baumert et al., 1999). Auch *materielle Ressourcen* spielen eine wichtige Rolle bei der Wissensvermittlung. Neben den klassischen Werkstätten zur Vermittlung praktischer Fertigkeiten gilt dies insbesondere für die neuen Informations- und Kommunikationstechnologien. Hierbei wird oft die durchschnittliche Anzahl von Computern pro Schüler als Indikator dafür verwendet, in welchem Maße den Schülern moderne Technologien zur Verfügung stehen¹⁷.

Lernortkooperation: Berufsbildung vollzieht sich in den meisten Ländern in unterschiedlichen Kooperationsformen zwischen Schulen, Betrieben und überbetrieblichen Einrichtungen. Ausmaß und Art der Lernortkooperation kann als Indikator für Vielfalt und Differenziertheit von Ausbildungsverfahren und für eine erfolgreich abgestimmte Umsetzung inhaltlicher Vorgaben zwischen ausbildenden Betrieben bzw. überbetrieblichen Einrichtungen und den zuständigen Schulen gelten (vgl. Kunter et al., 2002). Hierbei wird davon ausgegangen, dass Ausbildungsprogramme, die den Wechsel zwischen verschiedenen Lernorten vorsehen, Bildungseinrichtungen vor die Herausforderung einer aufeinander abgestimmten didaktisch-methodischen Umsetzung inhaltlicher Vorgaben stellen, die eine enge Zusammenarbeit der Organisationen erfordert. Kooperationen werden dabei durch teilweise gegensätzliche Bedingungen der beiden Lernorte erschwert, die durch unterschiedliche Zielausrichtungen und Entscheidungsspielräume bestimmt werden und ein unterschiedliches Selbstverständnis der Akteure in den jeweiligen Lernorten prägen. So zeigt die Geschichte der betrieblichen Berufsausbildung in Deutschland durchgängig das Problem des Spannungsfeldes von beruflicher Qualifizierung und betrieblichem Verwertungsinteresse (vgl. Euler, 2004; Weber et al., 2003).

Qualitätskontrollen: Innerhalb der vorgegebenen Verantwortungsbereiche steht den Schulen und ausbildenden Betrieben eine Reihe von Gestaltungsparametern zur gezielten Qualitätsentwicklung und Qualitätssicherung

17 Die erfolgreiche Integration erfordert jedoch mehr als die Verfügbarkeit entsprechender Ressourcen: Die Nutzung der neuen Technologien muss in den Lehrplänen verankert sein und das Lehrpersonal entsprechend geschult werden, um z. B. den Computer effektiv im Unterricht nutzen zu können.

zur Verfügung, die sich auch im Leistungsstand der Schulen manifestieren. Eine mögliche qualitative Differenzierung wäre die Unterscheidung zwischen formativer und summativer Evaluation und dabei insbesondere nach Art der Evaluation, z. B. Multiple-Choice-Tests vs. Gesellenstücke (vgl. Pellegrino et al., 2001).

B 2) Ausbildungsprozess und Unterrichtsgestaltung

Die auf der Makro- und Mesoebene beschriebenen Merkmale werden indirekt, vermittelt über proximale prozessuale Bedingungsfaktoren, im Ausbildungsprozess leistungswirksam (vgl. Kapitel 3.2.). Hiermit wird gewissermaßen die Umsetzung pädagogischer und betrieblicher Gestaltungsparameter ermittelt, die den Bildungseinrichtungen innerhalb des vorgegebenen (und weiter oben beschriebenen) Handlungsspielraums zur Verfügung stehen.

Gestaltung der Ausbildungsprozesse: Innerhalb der vorgegebenen Verantwortungsbereiche steht den Schulen und ausbildenden Betrieben eine Reihe von Gestaltungsformen zur Verfügung, die sich im Leistungsstand der Auszubildenden niederschlagen können. Dieser Indikator zielt auf die Darstellung der möglichen pädagogischen und organisatorischen Maßnahmen.

Schulische Lernsituation: Bezüglich der schulischen Ausbildung lassen sich die zum Einsatz kommenden pädagogischen und organisatorischen Gestaltungsformen (z. B. selbstorganisiertes Lernen, Projektlernen, Gruppenarbeit,) ermitteln und mit der Leistungsmessung in Beziehung setzen. Außerdem ist zu erheben, welche Einrichtungen (z. B. Bibliothek) und welche Aktivitäten (z. B. Förderunterricht) die schulischen Einrichtungen außerhalb des „normalen“ Unterrichts anbieten und wie diese von den Auszubildenden genutzt werden. Dabei wird vermutet, dass die intensive Nutzung solcher Lerngelegenheiten mit guten Leistungen verknüpft ist.

Betriebliche Ausbildungssituation: Bei der Beschreibung betrieblicher Lehr-Lernarrangements ist zu unterscheiden zwischen den Arbeitsinhalten, der Arbeitsumgebung und -organisation und den pädagogischen und didaktischen Arrangements, die das Arbeitshandeln begleiten (vgl. Onstenk, 2003). Im Hinblick auf die Lernrelevanz von Arbeitsplätzen in Ausbildungssituationen lassen sich nach Tramm (1992) die folgenden Qualitätsdimensionen unterscheiden:

a) Inhaltliche Erfahrungsmöglichkeiten

Welche inhaltlichen Erfahrungen (Begriffe, Theorien, Problemstellungen etc.) können Auszubildende in der Auseinandersetzung mit der Arbeitsaufgabe machen?

b) Formaler Anforderungsgehalt von Arbeitsaufgaben

Wie ist der Anforderungsgehalt von Arbeitsaufgaben einzuschätzen (bzgl. der Komplexität, Vollständigkeit und Problemhaltigkeit von Arbeitsaufgaben sowie des Handlungsspielraums, den die Arbeitstätigkeit bietet)?

c) Möglichkeiten einer aufgabenbezogenen, sozial-kommunikativen Auseinandersetzung

Welche Möglichkeiten einer aufgabenbezogenen, sozial-kommunikativen Auseinandersetzung mit Mitarbeitern bzw. Externen bieten die Arbeitsaufgaben?

Auf der Grundlage einer Analyse arbeitsplatzgebundener Lernprozesse in der beruflichen Bildung kommen Blockhuis (vgl. Nieuwenhuis, 2004) und Franke & Kleinschmitt (1987) zu einer Bildung von Variablen-Clustern, die einen starken Einfluss auf das Lernen von Auszubildenden haben. Franke und Kleinschmitt (1987) unterscheiden die folgenden Merkmale, die zum Aufbau beruflicher Kompetenzen beitragen:

- Problemhaltigkeit und Komplexität: bezieht sich auf die kognitive Herausforderung durch die Arbeit (Ausmaß an Denkopoperationen, Neuhaltigkeit der Arbeitsaufgabe, Komplexität, Anzahl der vorangehenden Planungsschritte).
- Variabilität: bezieht sich auf die Vielfalt der Arbeitsgegenstände, der Arbeitsverfahren und Arbeitsbedingungen als Voraussetzungen für die Transferierbarkeit vorhandenen Wissens.
- Handlungsspielraum: bezieht sich auf die Freiheitsgrade, die dem Auszubildenden bei der Bearbeitung einer Aufgabe zugebilligt werden.
- Handlungsfunktionale Integrität: bezieht sich auf die Einbeziehung des Auszubildenden in die Arbeitsprozesse.
- Soziale Unterstützung bei der Arbeit: bezieht sich auf das Feedback der Kollegen und Ausbildungsleiter bezüglich der Arbeitsleistung sowie der Möglichkeit, über auftretende Probleme zu sprechen und Erfahrungen austauschen.

Ähnlich unterscheiden Baethge und Baethge-Kinsky (2004) vier Dimensionen der Arbeitsorganisation, die einen lernförderlichen Einfluss haben:

1. betriebliche Entwicklungsmöglichkeiten;
2. Ganzheitlichkeit der Aufgabenstellung;
3. Kommunikations- und Kooperationsintensität (soziale Einbindung);
4. Partizipationschancen in der Arbeit.

Der größte Einfluss auf die Fähigkeit zu lernen kommt den betrieblichen Entwicklungsmöglichkeiten zu, aber auch die Ganzheitlichkeit der Aufgabenstellung und die Kommunikations- und Kooperationsintensität sind damit stark

gekoppelt. Den geringsten, aber dennoch markanten Zusammenhang zwischen der Lernförderlichkeit der Arbeit und der Lernfähigkeit zeigen die Partizipationschancen in der Arbeit. Insgesamt gilt jedoch, dass die Unterschiede zwischen den Dimensionen sehr gering sind, so dass prinzipiell alle vier Dimensionen zu berücksichtigen sind.

Lernklima: Dieser Komplex knüpft an die Forschung zu Schulklima und Unterrichtsqualität an (vgl. Kunter et al., 2002). Dabei soll erfasst werden, wie Auszubildende das allgemeine Schul- und Unterrichtsklima bzw. das Betriebs- und Arbeitsklima wahrnehmen. Mit der Erhebung dieser Klima- und Qualitätsvariablen verbindet sich die Hypothese, „dass die jeweils positiven Ausprägungen sich motivierend und lernstützend auswirken und deshalb Lernleistungen begünstigen“ (Kunter et al., 2002, S. 255). In Annäherung an PISA sollen hierbei die folgenden Dimensionen untersucht werden: Schul- bzw. Betriebszufriedenheit, die Zufriedenheit mit dem Unterricht in bestimmten Fächern bzw. am Arbeitsplatz, Lehrer-Schüler-Beziehungen und Schüler-Schüler-Beziehungen und die (generationenübergreifenden) Beziehungen zu den Kollegen in der Arbeit.

3.2.3. Einzelne Bildungsteilnehmer und Lernende

Wie im vorn präsentierten Gesamtmodell dargestellt (vgl. Abbildung 3.1), unterscheiden wir auf der Ebene einzelner Bildungsteilnehmer und Lernenden zwischen Informationen zum persönlichen Hintergrund des einzelnen Lernenden (Input/Kontext), den Einstellungen und Verhaltensweisen des Einzelnen (Prozess) und der Qualität individueller Bildungsergebnisse (Outcome). Zusätzlich zu den Untersuchungsinstrumenten zur Messung von Kompetenzen sollen die persönlichen Merkmale der Auszubildenden durch Fragebögen erhoben werden. Im Folgenden werden die einzelnen Kategorienklassen zu jeder Dimension vorgestellt.¹⁸

18 Als eine wesentliche Textgrundlage diene hierbei die Publikation „PISA-2000: Dokumentation der Erhebungsinstrumente“ des Max-Planck-Instituts für Bildungsforschung (Kunter et al., 2002).

C. Einzelne Bildungsteilnehmer und Lernende

Input

1. Persönlicher Hintergrund des einzelnen Lernenden
 - Sozioökonomische Stellung der Familie
 - Kulturelles Kapital der Familie
 - Lern- und Arbeitsbiographie

Prozess

2. Einstellungen und Verhalten des Einzelnen
 - Informationsverhalten und Lernzeit
 - Bildungsaspiration

Output/Outcome

3. Kompetenzerwerb und -verwertung
 - Qualifikationsniveau
 - Übergang in ausbildungsadäquate Tätigkeit und
 - Berufliche Mobilität

C 1) Persönlicher Hintergrund des einzelnen Lernenden (Input/Kontext)

In dieser Kategorie geht es im Wesentlichen um den Zusammenhang zwischen Aspekten des familiären Hintergrunds und der persönlichen Lerngeschichte auf der einen und den aktuellen Lernprozessen auf der anderen Seite. Hier sind vor allem der familiäre Kontext des Auszubildenden und sein soziales Umfeld zu thematisieren. Dass die Bildungslaufbahn in erheblichem Maße von familiären Sozialisationsprozessen – und damit von den sozioökonomischen und soziokulturellen Bedingungen des häuslichen Milieus – abhängig ist, gilt als eine seit langem gesicherte Erkenntnis der empirischen Bildungsforschung (Baumert et al., 2001). Mit Bezug auf das Kapitalkonzept von Coleman bzw. Bourdieu hat auch die PISA-Studie diese Faktoren in ihr Untersuchungsdesign mit einbezogen und dargelegt, dass der Einfluss des familiären Hintergrunds auf die Bildungslaufbahn nach wie vor besonders wirksam ist (vgl. Baumert & Schümer, 2002).

In der Regel wurde die soziale Herkunft mithilfe der sozioökonomischen Stellung der Eltern bestimmt. Heute gilt eine solche Bestimmung zwar immer noch als notwendig, allein aber als unzureichend: „Erstens haben die ökonomischen, politischen und sozialen Entwicklungen der vergangenen Jahrzehnte die Struktur und Funktion der Familie stark verändert, und zweitens ist es zu

einem relativ hohen Maß an Mobilität über Ländergrenzen hinweg gekommen, das zu kulturellen Differenzen innerhalb einzelner Länder geführt hat“ (Kunter et al., 2002, S. 225). Mit Blick auf die neueren Entwicklungen und ihre Bedeutung für den Schulerfolg von Jugendlichen sollen deshalb – zusätzlich zum sozioökonomischen Status der Eltern – Aspekte *des kulturellen und sozialen Kapitals* der Auszubildenden und deren Eltern erfasst werden. Die Eigenständigkeit des kulturellen und sozialen Kapitals der Auszubildenden ist besonders hervorzuheben. Die kontinuierlich steigende und bereits früh im Kindesalter einsetzende Bedeutung von peer-groups und Medien hat längst einen weitgehend eltern- und familienunabhängigen Erfahrungsraum der Jugendlichen konstituiert, der in hohem Maße bildungswirksam sein dürfte (s. auch weiter unten).

Sozioökonomische Stellung der Familie: Da hierbei die Berufsposition und die Bildungsabschlüsse der Eltern von zentraler Bedeutung sind (vgl. Baumert et al., 1999), sollte die soziale Herkunft von Auszubildenden mit Hilfe der sozioökonomischen Stellung ihrer Eltern bestimmt werden. Dieses Merkmal wird in der Regel auf der Basis des sozioökonomischen Index (International Socio-Economic Index, ISEI) über *Bildungsabschluss, Erwerbstätigkeitsstatus und berufliche Stellung der Eltern* erfasst. Diese Daten sind auch in PISA die wichtigsten Indikatoren für die sozioökonomische Stellung der Eltern. Der Internationale Sozioökonomische Index ist auf der Basis von Daten zu Bildung, Beruf und Einkommen von 74.000 Beschäftigten aus 16 Ländern erzeugt worden und hat sich in verschiedenen Vergleichsuntersuchungen bereits bewährt (vgl. Ganzeboom & Treiman, 2003).

Kulturelles Kapital der Familie: Jugendliche, die „bildungsfernen“ Schichten angehören und/oder nur wenig mit der herrschenden Kultur und Sprache ihres Landes vertraut sind, erbringen seltener hohe Leistungen als Jugendliche aus den gebildeten Schichten der Bevölkerung (vgl. Leschinsky & Mayer, 1999). Nach Bourdieu (1983) sind solche Leistungsunterschiede auf Unterschiede in der Verfügung über kulturelles Kapital zurückzuführen. Darunter sind alle kulturellen Ressourcen zu verstehen, die dazu beitragen, dass diejenigen Qualifikationen, Einstellungen, Verhaltensmuster und Wertorientierungen vermittelt werden, die zum Erfolg in einem spezifischen sozialen System führen. Zu den kulturellen Ressourcen gehören symbolische Machtmittel wie Kunstwerke oder Literatur, institutionalisierte Formen potentieller Macht wie Bildungszertifikate oder Titel und nicht zuletzt auch bestimmte Wahrnehmungs-, Denk- und Handlungsschemata, die im Sozialisationsprozess verinnerlicht werden (vgl. Kunter et al., 2002). Für ein Berufsbildungs-PISA sind in diesem Zusammenhang auch elterliche Berufstätigkeit und die Identifikationsmöglichkeiten, die sie Kindern bieten, im Auge zu behalten.

Des Weiteren ist im Rahmen einer Berufsbildungsstudie das *Weiterbildungsverhalten der Eltern* zu erheben, da davon auszugehen ist, dass es für die Wahrnehmungs- und Verhaltensweisen von Auszubildenden während und

nach der Ausbildung von Relevanz ist. Das kulturelle Kapital der Herkunftsfamilie eines Auszubildenden stellt somit einen wichtigen Indikator für die Herkunftsabhängigkeit von Bildungskarrieren und Kompetenzerwerb dar. Das Deutsche PISA-Konsortium weist darauf hin, dass Aspekte des kulturellen Kapitals bislang kaum quantitativ erfasst worden sind (Baumert et al., 2001).

Um festzustellen, über wieviel kulturelles Kapital die Auszubildenden verfügen, sollten u. a. ihre ethnische Herkunft, ihre Sprache und ihre Vertrautheit mit dem Aufenthaltsland sowie das Bildungsniveau ihrer Eltern erfasst werden (vgl. Kunter et al., 2002). Die *ethnische Herkunft* von Auszubildenden kann Unterschiede der Vertrautheit mit der in ihrem Aufenthaltsland vorherrschenden Kultur und Sprache anzeigen. Da die Staatsangehörigkeit längst kein hinreichendes Kriterium mehr ist, um zwischen verschiedenen Formen familiärer Sozialisation zu unterscheiden (vgl. Solga, 2005, S. 260 ff.), muss hier differenzierter nachgefragt werden: nach dem *Geburtsland der Auszubildenden und der Eltern, der erlernten Muttersprache und nach Schulbesuchszeiten im Ausland, dem Zeitpunkt bzw. Alter der Einwanderung* (vgl. Konsortium Bildungsberichterstattung 2006).

Das kulturelle Milieu der Herkunftsfamilie wird im internationalen PISA-Fragebogen mit Fragen nach dem *Besitz von Kulturgütern* (Musikinstrumenten, Kunstwerken, klassische Literatur) und *bildungsrelevanten Besitztümern* (wie z. B. Internetanschlüssen oder Taschenrechnern) erfasst. Allerdings ist hierbei zu beachten, dass die Aussagekraft dieses Indikators davon abhängig ist, ob die Auswahl der zu erfragenden Kulturgüter zu den entsprechenden Lernmilieus passt. So können sich Kulturgüter, die sich im Zusammenhang mit dem Schul-PISA als relevant erwiesen haben, bei Berufsbildungsprozessen als weniger aussagekräftig erweisen. Hier könnten technische Güter und praktische Momente der Alltagskultur (z. B. selbst getätigte Reparaturen) sowie das alltägliche Kommunikationsmilieu mehr Aussagekraft besitzen.¹⁹ Für sie sind Indikatoren zu entwickeln.

Lern- und Arbeitsbiographie: Im Hinblick auf unser Vorhaben, die Kompetenzen der Auszubildenden zu messen, ist davon auszugehen, dass die feststellbaren Leistungen nicht nur als das Resultat beruflicher Bildungsinstitutionen und -prozesse anzusehen sind, sondern zu einem großen Teil abhängig sind von der vorgängigen schulischen und außer-schulischen Lerngeschichte der Individuen (vgl. Lang & Pointinger, 2002, S. 101).

Zu den bildungsbiographischen Merkmalen zählen neben den erworbenen Schulabschlüssen Zertifikate in der nebenschulischen Weiterbildung und informelle Lernerfahrungen, wie sie in der Arbeit, im sozialen Umfeld (peers) oder auch in selbstorganisierten Lernaktivitäten (z. B. Umgang mit Medien) gewonnen werden können. Gerade der empirisch schwer fassbare Bereich des informellen Lernens hat in der bildungspolitischen Diskussion der letzten

19 An dieser Stelle wird sichtbar, dass Merkmale des sozialen und des kulturellen Kapitals ineinander übergehen.

Jahre an Aufmerksamkeit gewonnen (vgl. Straka, 2003; Baethge & Baethge-Kinsky, 2004)²⁰.

Bestätigt wird die Wichtigkeit informellen Lernens auch durch die Hinweise der internationalen Experten im Rahmen der Berufsbildungs-PISA-Workshops: Gerade im Hinblick auf die zahlreichen weniger oder kaum formalisierten Systeme der beruflichen Bildung in verschiedenen Ländern kommt informellen und nicht-formellen Lernformen ein sehr hoher Stellenwert für die Aneignung allgemeiner und beruflicher Kompetenzen zu.

Als ein weiterer Komplex kann das Freizeitverhalten der Auszubildenden und deren Einbindung in soziale Beziehungsgeflechte behandelt werden. Hintergrund für die Ermittlung von Informationen über das Freizeitverhalten der Auszubildenden sind die Erkenntnisse, die auf unterschiedliche Lebensstile von Jugendlichen verweisen, mit denen sich auch eine unterschiedliche Nähe zu ausbildungsspezifischen Anforderungen und intellektuellen Betätigungen verbindet (vgl. Kunter et al., 2002). Erhoben werden Freizeitaktivitäten einschließlich der Mediennutzung, wobei ein jugendkulturell vermittelter Zusammenhang zwischen einer eigenständig-kreativen Freizeitnutzung und guten schulischen Leistungen vermutet wird. Die sozialen Beziehungsmuster in den Gruppen gleichaltriger Jugendlicher geben ebenfalls Hinweise auf die Bedeutung, die die Art der Freizeitnutzung und der gegenseitigen Unterstützung für die schulischen Leistungen haben.

C 2) Einstellungen und Verhalten des Einzelnen (Prozess)

In dieser Kategorie geht es um die Erfassung des Auszubildenden als Lernenden. Die wesentlichen Merkmale (Motivation, Metakognition, Selbstwirksamkeit) wurden bereits in Kapitel 2 besprochen. Diese Merkmale sollen durch Angaben zum Informationsverhalten, zur durchschnittlichen Lernzeit und zur Bildungsaspiration der Auszubildenden ergänzt werden.

Informationsverhalten und Lernzeit: Neben den oben vorgestellten Lernansätzen von Auszubildenden sind zwei weitere Indikatoren bedeutsam für die Beschreibung individueller Lernprozesse: das Informationsverhalten des Einzelnen und seine aktiven Lernaktivitäten außerhalb des Unterrichts bzw. Arbeitsplatzes. Das Informationsverhalten soll zum einen Information über die Wissensbasis der Auszubildenden als Grundlage für die Berufswahl liefern, die beim Übergang von der Schule in die berufliche Ausbildung seine Entscheidungen steuert. Zum anderen soll es Informationen darüber liefern,

20 Als Impulsgeber einer international vergleichenden Untersuchung von Anteilen formeller und informeller/nicht-formeller Bildung sind vor allem internationale Organisationen wie die OECD mit ihren Studien zur Messung des Kompetenzzuwachses von Erwachsenen anzusehen (vgl. die Anmerkungen zu den internationalen Vergleichsstudien IALS, ALL und PIAAC in Kapitel 3.1.).

mit Hilfe welcher *Informationsressourcen* sich der einzelne Lerner sein Wissen aneignet. Die für schulische und außerschulische Lernaktivitäten verwendete Lernzeit gilt allgemein als ein wichtiger Schulleistungsprädiktor (vgl. Treiber, 1982). Er verweist auf die Motivation und Neigung des Einzelnen, seine Zeit in Lernaktivitäten zu investieren. *Lernzeit* lässt sich als *durchschnittliche Zeit für aktive Lernaktivitäten (gesondert nach Fach/Thema)* operationalisieren (vgl. Fisher & Berliner, 1985).

Bildungsaspiration: Angaben zur Bildungsaspiration sollen Auskunft geben über den angestrebten Ausbildungsabschluss und die Vorstellungen vom zukünftigen Beruf der Auszubildenden, die ein wichtiger motivierender Faktor zur Erzielung guter Ausbildungsleistungen sein können (vgl. Baumert et al., 2001). Die Berufswünsche der Schüler werden hierbei im Sinne persönlicher Projekte konzeptualisiert, verstanden als Handlungsprogramme zur Realisierung von Zielen. Dabei soll zum einen erfasst werden, inwieweit die Schüler an dieses Berufsziel gebunden bleiben, und zum anderen, wie sie die Erreichbarkeit dieses Ziels einschätzen (vgl. Kunter et al., 2002).

C 3) Kompetenzerwerb und -verwertung (Output/Outcome)

Über die Erhebung beruflicher Kompetenzen hinaus sollte ein Berufsbildungs-PISA auch Informationen über die Verwertung beruflicher Qualifikationen in der Arbeit, auf den Arbeitsmärkten und in der Biographie liefern. Solche individuellen Outputs sind nicht einseitig auf berufliche Bildungsprozesse und persönliche Merkmale zurückzuführen. Sie sind zu einem bedeutenden Teil auch durch Gelegenheitsstrukturen in Form von berufs-, branchen- und betriebsspezifischen Bedingungen bestimmt. Dies gilt sowohl im Hinblick auf die Möglichkeiten „an der zweiten Schwelle“, d. h. beim Übergang von der Ausbildung in den Beruf, als auch im Hinblick auf die Bewertung der in der beruflichen Bildung vermittelten Kompetenzen durch den Arbeitsmarkt. So ist bei der Erhebung beruflicher Bildungserträge die komplexe Dialektik zwischen Bildungssystem und Arbeitsmarkt zu berücksichtigen. Die folgenden Variablen sind demnach nur vor dem Hintergrund einer gründlichen Dokumentation systemischer Kontextbedingungen (s. o.) zu bewerten.

Qualifikationsniveau: Neben den tatsächlich vermittelten Kompetenzen charakterisiert das Qualifikationsniveau das formale *Niveau des in der Berufsausbildung erreichten Abschlusses* und spiegelt damit die Wertschätzung der Berufsausbildung innerhalb des Bildungssystems sowie bezogen auf die Anforderungen des Arbeitsmarktes wieder. Für eine differenzierte Betrachtung sollte es ergänzt werden durch Informationen über die mögliche horizontale und vertikale Verknüpfung mit anderen Bildungsgängen (Zugangsberechtigung).

Übergang in ausbildungsadäquate Tätigkeit: Dieser Indikator beschreibt die Qualität der beruflichen Einmündung junger Menschen nach einer Berufsausbildung und die Häufigkeit und Dauer beruflicher Mobilitätsereignisse.

Die Qualität der beruflichen Einmündung kann aus unterschiedlichen Perspektiven betrachtet werden: Der Begriff der *Ausbildungsadäquanz* beschreibt den Grad der Übereinstimmung zwischen der gegenwärtigen Beschäftigung von Schulabgängern und den vermittelten Kompetenzen in der Erstausbildung, wobei man nicht von zu engen Kopplungen ausgehen sollte. Es geht darum festzustellen, inwiefern das während der Ausbildung Gelernte bei der späteren beruflichen Tätigkeit auch genutzt werden kann (vgl. Büchel & Neubäumer, 2001).

Das Verhältnis von erlerntem und ausgeübtem Beruf verweist nicht zuletzt auf die oben beschriebene Verknüpfung von Ausbildungs- und Beschäftigungssystem, die Auskunft darüber gibt, wie weit die Struktur der Ausbildungsberufe auf die Struktur der Arbeitsplätze passt oder ob es zu systematischen Mismatches kommt. Daneben sind der Erwerbstätigkeitsstatus und die berufliche Stellung (vertikale Dimension und Einkommen) der Ausbildungsabgänger abzubilden. Bisher liegen Daten zu diesem Themenbereich von Eurostat zwar vor, jedoch nur mit Bezug auf das ISCED-Niveau des Bildungsabschlusses von Abgängern und ohne eine Differenzierung nach allgemeiner und beruflicher Bildung (vgl. Kapitel 4.1.).

Berufliche Mobilität: Seit längerem sind Arbeitsmärkte und Beschäftigungsstrukturen in den postindustriellen Gesellschaften von Diskontinuität, Unsicherheit und Volatilität gekennzeichnet. In der Kategorie der beruflichen Mobilität fallen zwei Prozesse zusammen: zum einen das erfahrene Ausmaß an Diskontinuität, das sich in Berufs-, Betriebs- und Arbeitsplatzwechseln äußern kann und eine Herausforderung für die berufliche Kompetenz des einzelnen im weitesten Sinne darstellt; zum anderen das Interesse an und der Umgang mit Mobilität. In beiden Dimensionen – Diskontinuitätserfahrung und Mobilitätsaktivitäten – sollte die berufliche Mobilität erfasst werden.

3.3. DATENVERFÜGBARKEIT

Im Hinblick auf die Datenverfügbarkeit gilt zunächst, dass derzeit auf internationaler wie nationaler Ebene in vielen Bereichen die empirische Basis für die Füllung der ausgewählten Aspekte und Indikatoren ausgesprochen lückenhaft ist und Daten selbst für eine grobe Typisierung auf der Ebene der Ausrichtung der Bildungsbereiche (VET Sek. II/bzw. ISCED 3 A–C) nicht zur Verfügung stehen. Selbst auf diesem Aggregationsniveau werden entsprechende Daten nicht für alle Länder ausgewiesen. Dies liegt in der Hauptsache daran, dass die Variabilität der beruflichen Bildung in Europa und auch im OECD-Raum viel

größer ist als in den anderen Bereichen des Bildungswesens und der Stand der vergleichenden Forschung viel weniger entwickelt ist. Schon die unterschiedlichen Definitionen beruflicher Bildung in den einzelnen Staaten führen zu grundlegenden Problemen länderübergreifender Vergleiche. Ein Tatbestand, der unter anderem auf die zum Teil stark abweichende Ausrichtung nationaler Statistiken und erhebliche terminologische Differenzen verweist (vgl. Kapitel 4.1.).

Besonders häufig fehlen Informationen zu angelsächsischen Ländern (Australien, Großbritannien, Vereinigte Staaten). Das Fehlen entsprechender Daten liegt häufig darin begründet, dass die modularisierten Ausbildungssysteme dieser Länder eine vergleichbare Unterscheidung zwischen allgemeinen und beruflichen Bildungsgängen und der sie vermittelnden Institutionen nicht zulassen. Weitere Schwierigkeiten erwachsen aus der stark variierenden Dauer post-sekundärer Bildungsgänge in diesen Ländern und der sich daraus ergebenden Problematik für die Ermittlung eines typischen Abschlussalters (vgl. NCVER, 2003a, S. 28).

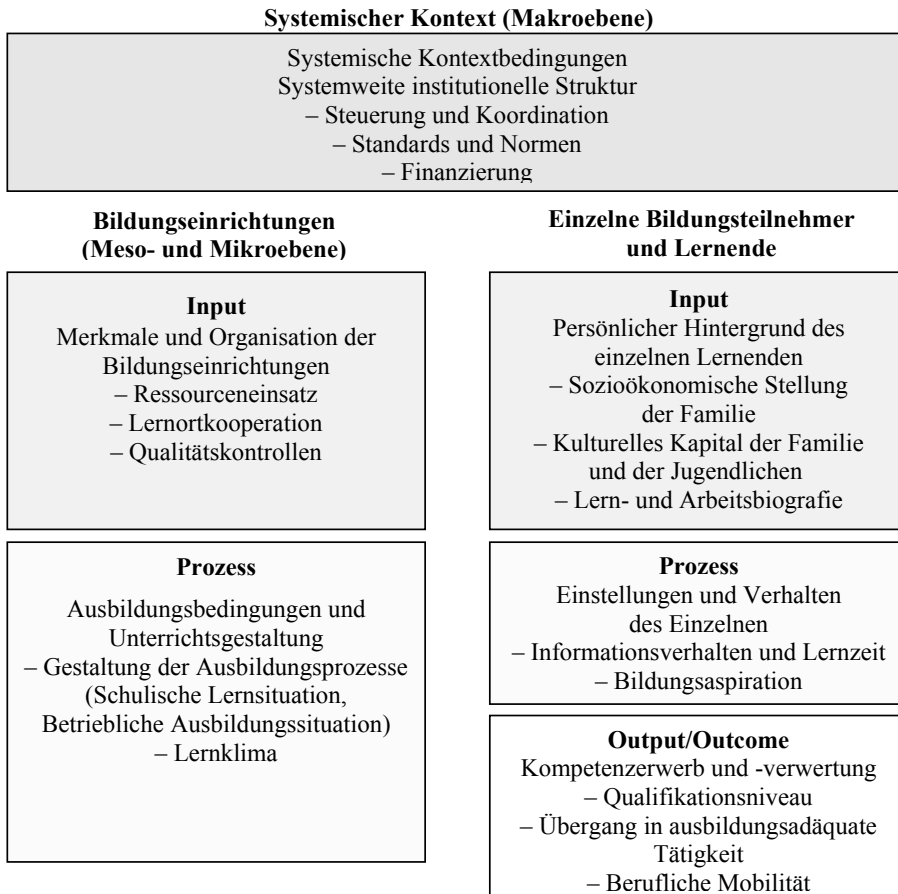
Aus diesem Grund schlagen wir vor, die hier vorgestellten Komplexe und Indikatoren selbst zu erheben. (Auf die damit einhergehenden Probleme und Lösungsstrategien einer geeigneten Sample-Bildung werden wir in Kapitel 4 eingehen.)

Die hier vorgestellten Indikatoren-Bündel (das Bildungs- und Beschäftigungssystem als Ganzes, Bildungseinrichtungen, einzelne Bildungsteilnehmer und Lernende) sind methodisch unterschiedlich zu erheben:

- Die Kontextindikatoren über die Analyse von Sekundärdaten, die gegebenenfalls zu ergänzen ist durch Interviews mit Bildungsexperten der einzelnen Länder.
- Die institutionellen Input-Bedingungen über eine Experten-Befragung von Schulleitern und den Verantwortlichen der betrieblichen Ausbildung.
- Die institutionellen Prozess-Bedingungen im Rahmen sowohl einer Institutionenanalyse als auch – vor allem beim zweiten Modell der Untersuchungsanlage (vgl. Kapitel 5.) – der Auszubildendenbefragung. Dies betrifft beispielsweise die Angaben der einzelnen Schüler zur Größe ihrer Klasse, dem Unterrichts- und Arbeitsklima oder den Lernangeboten. Dabei stellen die Angaben der einzelnen Auszubildenden subjektive Wahrnehmungen dar. Was die Gesamtheit der Schüler einer Klasse, einer Jahrgangsstufe oder einer Bildungseinrichtung zu den dort gegebenen Lernbedingungen sagt, ist aber erfahrungsgemäß verlässlich und dürfte das Lern- und Ausbildungsverhalten beeinflussen (vgl. Kunter et al., 2002). Die Daten aus den Auszubildendenfragebögen sind nicht ausschließlich als Individualdaten, sondern auch als Aggregate auf der Ebene der Einzelschule bzw. Ausbildungseinrichtung von Bedeutung.
- Die individuellen Bedingungen sind ebenfalls über eine Befragung der Auszubildenden im Zusammenhang der Kompetenzmessung zu erfassen.

Die folgende Abbildung liefert einen Überblick über die zu erhebenden institutionellen und individuellen Kontextbedingungen:

Abbildung 3.5.: Erhebungsplan institutionelle und individuelle Kontextbedingungen



4. DAS PROBLEM DER VERTIKALEN UND HORIZONTALEN VERGLEICHBARKEIT VON BERUFSBILDUNGSSTRUKTUREN: ZUR UNTERSUCHUNGSANLAGE UND SAMPLEKONSTRUKTION

Mit PISA und TIMSS existieren zwar internationale Vergleichsstudien des Typs „large scale assessment“. Dass diese für die Untersuchungsanlage eines Berufsbildungs-PISA nur einen begrenzten Vorbildcharakter haben, liegt an den Differenzen in der inhaltlichen und institutionellen Gestaltung zwischen allgemeinbildenden Schulen und Berufsausbildung. Haben erstere bei allen Unterschieden in der Mikroorganisation von Lernprozessen und personeller Zusammensetzung in verschiedenen Ländern doch einen gemeinsamen curricularen Grundkanon und eine im Durchschnitt ähnliche Altersstruktur der Klassen, so fehlen diese Voraussetzungen für ein Berufsbildungs-PISA. Berufsbildung ist inhaltlich sehr divergent entsprechend der Vielfalt beruflicher Domänen und institutionell sowohl in als auch zwischen Gesellschaften höchst heterogen.

Damit sind einer Machbarkeits-Studie eine Reihe von Problemen, deren tatsächliches Ausmaß noch zu klären ist, zur Lösung aufgegeben. Diese Probleme lassen sich in einer doppelten Vergleichsperspektive, der *vertikalen* und *horizontalen* Vergleichbarkeit, bündeln.

- *Vertikale Vergleichbarkeit* zielt auf das Niveau einer Berufsausbildung, das durch Institutionalisierungsform – z. B. auf Tertiär- (akademisch) oder auf Sekundärniveau – und Ausbildungszeiten (Zahl von Jahren) und/oder Altersstruktur der Teilnehmer definiert sein kann.
- *Horizontale Vergleichbarkeit* zielt auf den Sachverhalt, dass berufliche Curricula und Domänen in verschiedenen Ländern unterschiedlich definiert sind und es schwierig ist, inhaltlich vergleichbare Ausbildungsgänge oder Berufsfelder zu finden.

Die Sicherstellung beider Formen von Vergleichbarkeit ist für ein „large scale assessment“, das die Resultate der Kompetenzmessung mit den mikro- und makrostrukturellen Lernbedingungen verknüpfen will, unabdingbar. Muss man doch von der Prämisse ausgehen, dass sowohl die berufsinhaltliche Seite als auch die Frage des Niveaus und der Altersstruktur der Ausbildung Auswirkungen auf *alle* zur Messung anstehenden Kompetenz-Dimensionen hat. Wir wissen aus der Entwicklungspsychologie, dass allgemeine Kompetenzen wie z. B. Metakognition mit dem Lebensalter kovariieren (Hasselhorn, 2000).

Methodisch sicherzustellen, dass die jeweiligen nationalen Samples in etwa das gleiche Niveau der Berufsbildung als auch jeweils fachlich-funktionale

Äquivalente repräsentieren – also vertikal und horizontal vergleichbar sind –, ist schwierig. Im Falle der vertikalen Vergleichbarkeit liegt dies vor allem daran, dass die Berufsausbildung in den nationalen Bildungssystemen und ihren Traditionen verankert ist. Im Fall der horizontalen Vergleichbarkeit entstehen die Schwierigkeiten daraus, dass die Berufs- und Tätigkeitszuschnitte und die auf sie gerichteten Curricula der Berufsbildung in je national spezifischen Mustern der Arbeitsteilung und den Formen der Arbeitsorganisation begründet liegen (vgl. Heikkinen, 2000). Wir werden im Folgenden zunächst die Vergleichbarkeitsprobleme etwas genauer ausleuchten, um im Anschluss daran Konsequenzen für die Samplebildung zu ziehen und Lösungen zu erörtern.

4.1. VERTIKALE UND HORIZONTALE VERGLEICHBARKEIT UND DAS PROBLEM DER DEFINITION VON BERUFLICHER ERSTAUSBILDUNG

4.1.1 Vertikale Vergleichbarkeit

Die Aufgabenstellung der Machbarkeitsstudie sah die Festlegung des Untersuchungsgegenstands einer internationalen Vergleichsstudie auf „berufliche Ausbildung“ vor. Gemeint ist damit eine Konzentration des internationalen Vergleichs auf die Leistungsfähigkeit formalisierter Lernprozesse im Bereich der beruflichen Erstausbildung, d. h. jener institutionalisierten Bildungsprozesse,

- die in der Regel im Rahmen der Sekundarstufe II (ISCED 3a–c) absolviert werden,
- im Schwerpunkt die Vermittlung beschäftigungs- und arbeitsmarktrelevanter Qualifikationen bzw. Kompetenzen zur Aufgabe haben und
- mit deren Abschluss sich in der Regel der Übergang auf den Arbeitsmarkt verbindet.

Grundsätzlich ist berufliche Bildung auf der Systemebene weltweit, aber speziell in Europa als ein hochdynamisches Feld zu verstehen, das nur mit Mühe als ein Gebilde mit international weitgehend einheitlichen Konturen verstanden und verglichen werden kann. Dies gilt für die berufliche Bildung insgesamt wie auch für das, was als berufliche Erstausbildung bezeichnet wird und in einer institutionellen Definition häufig bestimmte Aspekte der Erstausbildung (z. B. innerhalb der Pflichtschule) ausblendet (Descy & Tessaring, 2002). Neben Fragen der Abgrenzung zum allgemeinbildenden Schulsystem und zu speziellen berufsvorbereitenden Programmen auf der einen Seite sind damit auch solche der Abgrenzung zum Weiterbildungssystem auf der anderen Seite angesprochen.

So werden beispielsweise in Italien berufsorientierende/-vorbereitende Maßnahmen explizit nicht zum Berufsbildungssystem gezählt; zudem weisen

insbesondere in Ländern mit einem wenig ausgebauten System der beruflichen Erstausbildung (z. B. Australien) Aus- und Weiterbildungsgänge eine geringe Trennschärfe auf und sind Formen der Ausbildung („Lehrlingsausbildung“), die hierzulande der Erstausbildung zugerechnet werden, in anderen Ländern (Frankreich, Finnland, Italien) als Programme anzusehen, die der „Chancenverbesserung“ von Jugendlichen dienen, die ansonsten aus dem Bildungssystem herauszufallen drohen. Die Entwicklung von ausgebauten „Übergangs“- oder „Chancenverbesserungssystemen“ (vgl. Baethge u. a. 2003) mit je eigenen Regeln und Verfahren kann als eine der wesentlichen Veränderungen im institutionellen Setting nationaler Berufsbildungssysteme in Europa gelten (vgl. Descy & Tessaring, 2002).

Einfach ausschließen aus der Betrachtung kann man solche Systeme aber schon deshalb nicht, weil teilweise erhebliche Anteile eines Altersjahrgangs diese durchlaufen, und sich solche Systeme – allen institutionellen Beharrungskräften zum Trotz (Leney & Deluca, 2000) – mit der Zeit durchaus zu einem eigenständigen Bestandteil beruflicher Erstausbildung im entsprechenden Land entwickeln können, d. h. auch jene „mittleren“ Kompetenzen vermitteln können, die in Deutschland in der Regel durch eine Ausbildung im „dualen System“ erworben werden.

Die Frage der Definition von beruflicher „Erstausbildung“ und der Zuordnung des in allen Ländern ausgeweiteten Sektors von Maßnahmen zur Berufsvorbereitung oder zum Übergang von allgemeinbildenden Schulen in Berufsausbildung ist mit der Berücksichtigung der „Benachteiligtenproblematik“ im Untersuchungsdesign verknüpft, d. h. jener Probleme im Zugang zu und in der Durchführung von Ausbildung, die als Aspekte sozialer Selektion (nach sozialer Herkunft, Geschlecht, ethnischer Zugehörigkeit) in der nationalen wie internationalen Berufsbildungsforschung verhandelt werden. Offen war vor allem, ob die entsprechenden Merkmale (z. B. Jugendliche/junge Erwachsene ohne qualifizierten Bildungsabschluss, Frauen, Jugendliche/junge Erwachsene mit Migrationshintergrund) als Schichtungskriterien einer Stichprobe dienen sollten:

Das Problem wurde auf dem ersten internationalen Workshop (18. – 20. Okt. 2004) verhandelt. In der ausführlichen Diskussion votierten die ausländischen Experten – unter Hinweis auf die in vielen Ländern prekäre Datenlage zu dieser Frage und die insgesamt schon hohe Komplexität der Vergleichsstudie – dafür, diese Problematik nicht zum expliziten Vergleichsgegenstand der Studie zu machen.

Angerissen ist damit das eigentliche Kernproblem der Samplekonstruktion: Aus der Vielzahl beruflicher Bildungsgänge und -programme aller in Frage kommenden Länder ist jeweils ein Sample zu definieren, das sowohl nach dem Niveau der in Ausbildungsgängen angezielten (bzw. in Tätigkeiten abgeforderten) Kompetenzen als auch nach dem fachlich-funktionalen Zugschnitt vergleichbar ist. Die Analyse der verfügbaren Informationen verweisen

auf mehrere grundlegende Probleme, die in der mangelhaften Verfügbarkeit aktualisierter und international einheitlich klassifizierter Daten ebenso wurzeln wie in unterschiedlichen, kulturell wie auch politisch bedingten nationalen Definitionen von Ausbildung und Tätigkeiten.

Eine Bestimmung von Programmen/Ausbildungsgängen auf gleichem Niveau kann sich international nur auf wenige – und häufig nicht besonders aktuelle – Datenquellen stützen. Informationen hierzu werden vor allem von der OECD, der UOE sowie von EUROSTAT/CEDEFOP bereitgestellt. Die brauchbarste Quelle zur Identifizierung quantitativ bedeutsamer qualifizierter Erstausbildungsgänge liefern vor allem die genannten europäischen Institutionen: In Zusammenarbeit von EUROSTAT und CEDEFOP wurden auf der Basis von Daten aus den Jahren 1995/1996 erstmals systematische Übersichten über die in den EU-Ländern existierenden Bildungsgänge (Programme) erstellt und für das Jahr 2000 aktualisiert. Diese Übersichten enthalten generell Angaben zu den einzelnen Ausbildungsgängen (Programmbezeichnung), dem Typ des Programms (berufsvorbereitend, allgemeinbildend oder berufsbildend), dem ISCED-Level des Ausbildungsziels und dem typischen Eintrittsalter. Darüber hinaus enthalten sie weitere Angaben, etwa zur Art der in diesem Bildungsgang erworbenen Arbeitsmarktqualifikation (allgemein oder fachspezifisch) wie auch zu den Teilnehmerzahlen; allerdings werden diese Angaben längst nicht für alle Länder ausgewiesen. Angaben über Ausbildungsanfänger sind in dieser Statistik überhaupt nicht enthalten.

Tatsächlich zeigt sich anhand von Datenanalysen wie auch an den von den internationalen Experten formulierten Statements, dass sich die ISCED-Klassifikation aufgrund ihrer systematischen Verbindung zu den Stufen des allgemeinen Bildungssystems nur begrenzt als Indikator für das Niveau der Ausbildung eignet – insbesondere dann, wenn es um die Identifizierung von Erstausbildungsprogrammen geht, die auf ein „mittleres Niveau“ der vermittelten Qualifikationen abzielen.

Die in Frage kommenden Programme der jeweiligen Staaten bewegen sich – je nach Ausbildungsfeld – nach der offiziellen Klassifikation auf den Levels 3 und 4 (und sind in Ausnahmefällen sogar auf Level 5 eingeordnet). Der ISCED-Level 3 bildet dabei die Untergrenze für Programme, die für die Sampleauswahl qualifizierter Ausbildungen herangezogen werden können. Insbesondere die in der Statistik für einzelne Länder auf dem ISCED-Level 3 ausgewiesenen berufsvorbereitenden Programme bzw. Kurzausbildungen²¹ erfüllen diese Kriterien jedoch nicht.

21 Hierzu zählen in Deutschland das Berufsgrundbildungsjahr wie auch die einjährigen Berufsfachschulen, in Dänemark die „basic vocational education (EGU)“, in der Schweiz die „Anlehre“, in Ungarn die „Secondary VET-School“, in den USA die „Vocational High Schools“ sowie im Vereinigten Königreich Ausbildungsaktivitäten, die zu einem niedrigen Grad der National Vocational Qualification (NVQ-Level 1 oder 2) führen (vgl. Tab. 1 im Anhang).

Mit Blick auf die für ausgewählte Berufs-/Tätigkeitsgruppen entwickelten und auf dem 2. Workshop vorgestellten Anforderungsprofile zeigt sich, dass die systematisch erhobenen Daten zu relevanten Ausbildungsprogrammen nur begrenzt die bestehenden Unterschiede innerhalb nationaler Bildungskonzeptionen, aber auch deren reale Veränderungen widerspiegeln. Ein Beispiel hierfür ist die Entwicklung in den *USA*: Galten dort die 2-jährigen Community Colleges sowie – mit Abstrichen – auch die Vocational High Schools als Institutionen der Vermittlung mittlerer beruflicher Kompetenzen (Münch, 1999), haben ihnen nach Auskunft unserer amerikanischen Experten zumindest in bestimmten Feldern (Banken, Informationstechnik) die 4-jährigen Colleges den Rang abgelaufen. Ähnliches gilt nach Aussage der französischen Expertin für *Frankreich*, wo die Ausbildungen zu Bankkaufleuten und in informationstechnischen Berufen im Rahmen von „BAC“-Programmen (Baccalaureat Professionnel bzw. Technologique) absolviert werden.

Ein weiteres Problem in diesem Zusammenhang bildet der Sachverhalt, dass fachliche *Ausbildungsfelder* bzw. fachliche Ausbildungsgänge keineswegs ein sicheres Definitionsprinzip zur vollständigen Erfassung qualitativ gleicher/ähnlicher Berufsausbildungsprozesse darstellen; vielmehr differieren die institutionellen settings, denen ein erheblicher Einfluss auf die Qualität der vermittelten Kompetenzen bzw. der Ausbildungsprozesse zuzuschreiben ist, auch innerhalb formal gleicher Ausbildungsfelder:

Zum einen können in einem Teil der Staaten sowohl die fachlichen Inhalte als auch die Formen ihrer Vermittlung (Kombinationen von theoretischen und praktischen, von schulischen und betrieblichen Lernprozessen) aufgrund „modularer“ Konstruktionen und unterschiedlicher Trägerstrukturen erheblich differieren, ohne dass das Gewicht solcher Differenzen und ihre Auswirkungen auf die fachliche Struktur und das Niveau der Kompetenz immer hinlänglich bekannt wäre. Dieses Problem existiert vor allem in Ländern, in denen eine eher lockere Beziehung zwischen Bildungssystem und Arbeitsmarkt (d. h. zwischen Qualifikationen und Tätigkeiten) besteht.

Besonders stark ausgeprägt finden wir dieses Problem in den *USA*, in *Australien*, aber auch in *Großbritannien*²², *Finnland*, *Frankreich*, *Griechland*, *Norwegen*, *Schweden*, *Spanien* u. a. Ländern – bei letzteren allerdings weniger im Bereich der Lehrlings- als vielmehr der vollzeitschulischen Ausbildung (Lasonen & Manning, 2001; Hannan, 1996). Selbst eine solche auf jüngeren Studien basierende Einschätzung kann gleichwohl schon durch Tagesaktualitäten überholt sein: Die *Niederlande* etwa, die nach den zitierten Studien für eine

22 Einen jüngeren Beleg hierfür lieferte eine jüngere Studie, bei der die Kompetenzen getestet wurden, die Auszubildende/Schüler des dem „Industriekaumann“ curricular vergleichbaren Ausbildungsgangs des „Advanced Business“ in England/Wales aufweisen. Die in dieser Fallstudie bei 151 den englischen Probanden gemessenen Kompetenzen lagen teilweise deutlich unterhalb des Niveaus, der nach dem Qualifikationslevel des Ausbildungsgangs (Level 3 GNVQ) zu erwarten war (vgl. Fulst-Biel & Ebner, 2005).

enge Beziehung von Qualifikationsvermittlung und Berufstätigkeit stehen, haben laut der Einschätzung des niederländischen Experten aufgrund der Modularisierung ihres Systems hier eher an Stabilität eingebüßt.

Zum anderen differiert das Einstiegsalter in berufliche Erstausbildung sowohl zwischen den Ländern als auch innerhalb der Länder erheblich und werden bestimmte fachliche Erstausbildungen in deutlich höherem Alter begonnen als früher: In vielen Ländern liegt zwar das Ende der Pflichtschulzeit (und d. h. meistens das Eintrittsalter in berufliche Erstausbildung) bei inzwischen 16 Jahren. Ausnahmen nach oben stellen Deutschland (18 Jahre) und die Niederlande (17 Jahre) dar. Aber gerade in osteuropäischen Ländern wie Polen, Tschechien oder in Ländern wie Portugal und Belgien liegt dies Alter deutlich darunter (vgl. Tabelle 4.1).

Tabelle 4.1: Typisches Alter am Ende der Pflichtschulzeit sowie nach Abschluss einer ISCED 3-Ausbildung nach Ländern
(Angaben in %)

	Typisches Alter	
	am Ende der Schulpflicht	bei ISCED 3-Abschluss
Australien	15	17
Tschechien	15	17–18
Dänemark	16	19–20
Deutschland	18	19
Großbritannien	16	*
Spanien	16	17
Frankreich	16	17–21
Niederlande	17	17–20
Norwegen	16	16–19
Polen	15	18–20
Portugal	15	18–20
Schweden	16	19
Schweiz	16	17–20
USA	16	k.A.

nicht ausgewiesen.

Quelle: OECD 2004.

Hinzu kommt, dass nach Aussagen der Experten auf dem 2. Workshop sich die Spanne des Alters, in dem eine solche Ausbildung aufgenommen wird, sukzessive zu höheren Altersstufen verschiebt – ein Phänomen, dem man auch in Deutschland begegnet: Nach der Statistik des BIBB betrug das Durchschnittsalter derjenigen, die im Jahre 2003 eine duale Ausbildung begonnen haben, über 19 Jahre und erstreckt sich die Spanne der quantitativ gewichtigen Altersjahrgänge (Anteil an Neuverträgen über 10 %) von den unter 16-Jähri-

gen bis hin zu den 20-Jährigen (vgl. <http://bibb.skygate.de/Z/B/30/9960-0000.pdf>). Ähnliches lässt sich laut dem dänischen Experten auch in seinem Heimatland beobachten.

Darüber hinaus werden bestimmte Ausbildungsgänge in einem höheren Mindest-Alter aufgenommen als andere: Ein spezifisches Beispiel hierfür ist die Krankenpflege-Ausbildung, deren Aufnahme in den meisten Ländern ein Alter von 18 Jahren voraussetzt.

Schließlich sehen immer mehr Berufsbildungssysteme – wie inzwischen auch das deutsche – sogenannte „Externenprüfungen“ von Personen vor, die eine weder zeitlich noch inhaltlich voll standardisierte Ausbildung durchlaufen und durch Ausbildungsstatistiken nur begrenzt erfasst werden.

In vielen europäischen und außereuropäischen Ländern findet sich eine hohe Quote von Berufsanfängern, die ihre Erwerbstätigkeit allein auf der Basis eines Abschlusses der Sekundarstufe 1 (ISCED-Level 2 und niedriger) aufgenommen haben, d. h. bei Aufnahme der Erwerbstätigkeit weder einen allgemeinbildenden noch einen berufsbildenden Abschluss der Sekundarstufe II aufweisen:

Im Durchschnitt aller damaligen EU-Länder betrug die Quote der Berufsanfänger mit einem Abschluss auf dem ISCED-Level 0-2 im Jahr 1997 32 % (vgl. Tabelle 4.2); die niedrigsten Quoten (mit Anteilen von 14 % bis 19 %) weisen hier Länder wie Finnland, Österreich, Deutschland, Frankreich und Schweden auf; die höchsten Quoten (mit Anteilen von 38 % bis 62 %) finden wir in Ländern wie Großbritannien, Spanien, Italien und Portugal. Ohne die Anteile für die USA genauer spezifizieren zu können, gaben unsere amerikanischen Experten zu erkennen, dass auch dort der direkte Übertritt von der Pflichtschule (High School) in die Erwerbstätigkeit ein massenhaftes Phänomen darstellt.

Bei allen Problemen der Aussagefähigkeit sowohl der ISCED-Klassifikation als auch von Daten, die 8 Jahre alt sind – die genannten Quoten dürften nur die Untergrenze des Ausmaßes bezeichnen, in dem Jugendliche im Anschluss an die allgemeinbildende Schulzeit und ohne eine wie immer geartete berufliche Erstausbildung direkt in den Arbeitsmarkt eintreten.²³ Da uns genauere, aktualisierte und vergleichbare Daten aus den amtlichen Statistiken nicht vorliegen, haben wir eine Schätzung* zum Institutionalisierungsgrad beruflicher Erstausbildung für die Länder der europäischen Union vorgenommen. Auch bei dieser Schätzung zeigt sich in etwa das gleiche Bild erheblicher Diskrepanzen zwischen einzelnen Ländern (vgl. Tabelle 4.3):

23 Es handelt sich hier um die Quote der Berufsanfänger, also nicht um die Quote derjenigen, die sich auf dem Arbeitsmarkt nach einer Stelle umsehen. Die Arbeitslosenquote der Personen ohne qualifizierten Bildungsabschluss lag in praktisch allen EU-Ländern (Ausnahme: Portugal) im Jahre 2000 deutlich über dem Durchschnitt (vgl. Fischer & Siebern-Thomas, 2004), also dürfte diese Personengruppe bei den Berufsanfängern tendenziell unterrepräsentiert sein.

Tabelle 4.2: Bildungsstruktur der Berufsanfänger (Juniors) in der EU im Jahr 1997 nach ISCED-LEVEL (Angaben in %)

	ISCED 0–2	ISCED 3	ISCED 5–7	insgesamt
Belgien	21	35	44	100
Dänemark	20	52	28	100
Deutschland	17	56	27	100
Griechenland	25	54	21	100
Spanien	45	15	39	100
Frankreich	18	44	38	100
Irland	25	38	37	100
Italien	50	42	7	100
Luxemburg	30	35	35	100
Niederlande	26	40	34	100
Österreich	15	76	9	100
Portugal	62	17	21	100
Finnland	14	55	31	100
Schweden	19	62	19	100
Großbritannien	38	36	26	100
EU-15 insgesamt	32	42	44	100

Quelle: Europäische Union, 2001, S. 49.

In Ländern wie Griechenland, Irland, Spanien und Portugal (und möglicherweise Italien – vgl. Fußnote Tabelle 4.3) spielt die im Anschluss an die allgemeinbildende Pflichtschule aufgenommene berufliche Erstausbildung eine eher marginale Rolle (Quoten um die 20 %), während sie in Ländern wie Belgien, Deutschland, Niederlande, Österreich, Tschechien und Ungarn (Quoten zwischen 60 % und 78 %) dominiert.

Das Interessante an diesen Befunden zum Übergangsverhalten ist jedoch ein anderer, damit verbundener Sachverhalt: Insbesondere in den anglo-amerikanischen Ländern (USA, Großbritannien, Australien) vollzieht sich ein erheblicher Teil dessen, was man in den deutschsprachigen Ländern als berufliche Erstausbildung fasst, entweder in post-sekundären (hoch-)schulischen Bildungsgängen oder aus bestehenden Beschäftigungsverhältnissen heraus. Letztere fungieren dann als Grundlage für Berufsfindungsprozesse und Bildungswahl, stellen die materielle Basis der Finanzierung späterer beruflicher Bildungsprozesse dar und bilden zugleich als informelles Lernarrangement einen Eckpfeiler beruflicher Qualifizierungsprozesse, die dann – gegebenenfalls – durch formalisierte Lernprozesse „off the job“ ergänzt werden.

Tabelle 4.3: Bildungsbeteiligung 17jähriger in ausgewählten Ländern nach institutionellen Formen – Angaben in % (2002) – Schätzung*

	Keine Bildungsbeteiligung	In Allgemeinbildung (Sek II)	In beruflicher Erstausbildung (Sek II)**	Gesamt
Belgien	0	30	70	100
Dänemark	17	39	44	100
Deutschland	7	34	59	100
Griechenland	30	42	28	100
Spanien	18	51	31	100
Frankreich	11	39	50	100
Irland	28	72 (52)	0 (20)	100
Italien	21	58 (82)***	21 (51)	100
Niederlande	17	26	57	100
Österreich	22	22	56	100
Portugal	30	50	20	100
Finnland	6	40	54	100
Schweden	4	48	48	100
Großbritannien	26	21	53****	100
Norwegen	7	39	54	100
Tschechien	2	19	79	100
Ungarn	15	74	11	100

* Der Schätzung wurden folgende Angaben zugrunde gelegt: die Verteilung der Schüler/Auszubildenden des Sekundarbereichs II auf institutionelle Formen sowie die allgemeine Bildungsbeteiligungsquote der 17jährigen Bevölkerung. Verknüpft wurden hierfür Daten der OECD und der Europäischen Union. Die Beteiligungsquote der 17jährigen haben wir deshalb herangezogen, weil in diesem Alter in den meisten Ländern die Pflichtschulzeit beendet und berufliche Ausbildung in diesem Alter in der Regel Erstausbildung (also) keine Weiterbildung ist.

** Ohne berufsvorbereitende Maßnahmen.

*** Für Italien gilt als strittig, ob die Fachoberschulen (istituti tecnici) dem allgemeinbildenden oder dem beruflichen Schulwesen zugerechnet werden.

**** Incl. Berufsvorbereitender Maßnahmen

Quelle: Eigene Berechnungen nach EU (2002), OECD (2003).

Für die *USA* wiesen die beiden amerikanischen Berufsbildungsexperten auf dem 2. Workshop darauf hin, dass es ganze Branchen (z. B. den Handel) gebe, in denen Beschäftigungsverhältnisse auf niedrigem Kompetenzniveau existierten, die als „Übergangsbeschäftigung“ (transition job) im Rahmen der Berufsfindungsprozesse 16-18jähriger Jugendlicher dienen. Illustriert wurde diese Abfolge von Pflichtschulzeit-Ende, Aufnahme von Erwerbstätigkeit, Berufswahl und Qualifizierung an einem Beispiel aus dem Bereich des Gesundheitswesens: Es wäre durchaus typisch, nach Schulende erst einmal eine Stelle als Putzfrau in einem Krankenhaus anzutreten, nach einiger Zeit arbeitsbeglei-

tend eine erste, vielleicht nicht besonders qualifizierte Ausbildung in der Krankenpflege zu beginnen, weiter zu arbeiten und später eine qualifizierte Ausbildung anzuschließen.

Ähnliches berichteten die Experten aus *Großbritannien*: Größere Teile von Beschäftigung im UK dienen einzig und allein dem Zweck, den Lebensunterhalt der ersten Jahre nach der Pflichtschule zu sichern, in denen man sich beruflich orientiert und weiterführende Bildungsgänge besucht (the function of the job is „to be a student“). Zudem sieht die Organisation des beruflichen Ausbildungswesens vor, dass die Bildungsabschlüsse grundsätzlich auch auf der Basis informell erworbener Kompetenzen, d. h. ohne Nachweis des Besuchs formalisierter Kurse/Programme erteilt werden können.

In die gleiche Richtung zielte der Beitrag von Stevenson (*Australien*): Danach erfolgt der größte Teil der Vorbereitung auf das Berufsleben *Australiens* in post-sekundären Institutionen wie etwa den TAFE-Colleges, wobei ein zunehmender Teil der Jugendlichen 18 Jahre und älter ist und die Ausbildung aus einem Beschäftigungsverhältnis heraus aufnimmt. Auch hier gilt die Regelung, dass informell erworbene Kompetenzen grundsätzlich den in formalisierten Bildungsgängen erworbenen gleichgestellt werden.

4.1.2. Horizontale Vergleichbarkeit

Die am Ausgangspunkt der Machbarkeitsstudie von uns vorgenommene Annäherung an die wichtigsten Tätigkeitsfelder, in denen die Studie angesiedelt werden sollte, erschien für die deutschen Verhältnisse auf den ersten Blick plausibel, reflektiert aber die Komplexität eines internationalen Vergleichs nur begrenzt.

Betrachtet man die Beschäftigungsstruktur im internationalen Vergleich nach Sektoren und bezieht die Entwicklungsdynamik in Europa mit ein, so lassen sich grob als allgemeine Trends die folgenden konstatieren (vgl. Tabelle 4.4): Der Dienstleistungssektor ist eindeutig der Bereich, der in der EU mit großer Dynamik wächst (+ 12 % an Arbeitsplätzen zwischen 1997 und 2002). Die Industrie (Verarbeitendes und Bau-Gewerbe) (+ 3 %) hinkt hinterher und der primäre Sektor verzeichnet einen Rückgang von 13 %. Trotz insbesondere in den Beitrittsländern ausgeprägter agrarischer Strukturen (Anteil in Rumänien: ca. 33 %, Polen 12 %) spricht diese Entwicklung dafür, den Vergleich auf gewerbliche und Dienstleistungsberufe/-bereiche zu beschränken.

Tabelle 4.4: Beschäftigungsstruktur nach Sektoren und Wirtschaftszweigen in ausgewählten Ländern 2002 (Angaben in %)

Land Sektor/Wirtschaftszweig	DK	D	E	F	NL	P	S	UK	CZ	RO	NO	CH	US	Sektorale Beschäftigungs- entwicklung in der EU 1997 bis 2002 (in %)
Land- und Forstwirtschaft, Fischerei	4	2	6	4	3	12	2	1	5	32	4	4	3	- 13
Industrie u. verarbeitendes Gewerbe	22	28	29	22	20	34	23	19	40	32	22	25	21	+ 3
Hochtechnologie-Bran- chen des verarbeitenden Gewerbes	6	12	5	7	5	4	7	7	9	6	*	*		*
Dienstleistungen	74	70	65	74	77	54	75	80	55	36	74	71	76	+ 12
Groß- und Einzelhandel, Instandhaltung und Reparatur von Kraftfahrzeugen	15	14	16	13	16	15	12	15	13	10	*	*		*
Gesundheits- und Sozialwesen	18	10	6	11	15	5	19	11	6	4	*	*		*
Hochtechnologie-Bran- chen des Dienstleistungs- sektors	5	3	3	4	4	2	5	5	3	2	*	*		*
Gesamt Sektoren	100	100	100	100	100	100	100	100	100	100	100	100	100	

* nicht ausgewiesen.

Quelle: Europäische Kommission, 2002.

Innerhalb der Industrie stellen die metallverarbeitenden Tätigkeiten von EU-weit 8 % einen größeren, per Saldo noch wachsenden Block dar. Ähnliches gilt – mit Abstrichen – für den Hochtechnologiebereich des Verarbeitenden Gewerbes. Beides spricht dafür, jeweils ein gewerblich-industrielles Be-

rufsfeld aus dem Baubereich sowie aus dem industriellen Sektor mit einzubeziehen. Innerhalb des Dienstleistungssektors stellen Groß- und Einzelhandel (einschließlich Instandhaltung und Reparatur von Kraftfahrzeugen) als traditionelle Branche mit europaweit ca. 15 % Beschäftigtenanteil einen ersten größeren Tätigkeitsbereich dar, dessen beschäftigungsstrukturelle Bedeutung international wenig variiert und zwischen 1997 und 2002 ein erhebliches Beschäftigungswachstum zu verzeichnen hat. Einen weiteren großen Bereich, auf den ca. 10 % der Beschäftigung entfallen (mit höherer Varianz des Anteil an der Gesamtbeschäftigung: 5 % in Portugal, 18 % in Dänemark) und der erhebliche Zuwächse zu verzeichnen hat, bilden das Gesundheits- und Sozialwesen. Schließlich ist hier noch der Hochtechnologiebereich der Dienstleistungen zu nennen, der mit ca. 4 % durchschnittlichem Beschäftigtenanteil ein vergleichsweise kleines, aber stark wachsendes Beschäftigungsfeld darstellt. Dies alles spricht dafür, aus den skizzierten Bereichen wichtige gewerblich-technische, kaufmännisch-verwaltende und personenbezogene Ausbildungsbereiche einzubeziehen.

Auf diese Weise sind in einer ersten Annäherung die Bereiche abgesteckt, auf die sich eine Auswahl der Ausbildungsfelder konzentrieren sollte:

- Gewerblich-technische Berufe aus Industrie und Handwerk
- kaufmännische bzw. kaufmännisch-verwaltende Berufe in Handel und sonstigen Dienstleistungen
- sozial- und krankenpflegerische Berufe aus dem Bereich der personenbezogenen Dienstleistungen sowie
- IuK-Berufe aus dem Bereich der informationstechnischen Dienstleistungen.

Damit ist freilich das Problem – wie die weiteren internationalen Diskussionen ergaben – noch keineswegs gelöst: Ein erster Grund dafür liegt darin, dass international kein einheitliches, an fachlich-funktionalen Domänen der Kompetenz orientiertes Konzept der Strukturierung und Differenzierung von Erstausbildungsgängen (oder Tätigkeitsfeldern) bzw. der Berichterstattung darüber verfolgt wird, auf das man sich stützen könnte (vgl. Lisbon-to-Copenhagen-to-Maastrich Consortium Partners, 2004). Dies heißt weniger, dass daran anschlussfähige Systematiken vollständig fehlten, sondern dass Basisdaten fehlen bzw. nicht entsprechend aufbereitet werden: Auf internationaler Ebene liegt neben der ISCO 88 (für den Bereich der Tätigkeitsfelder) für die Ausbildungsfelder („fields of training“) eine im Auftrag von CEDEFOP erstellte Systematik zur beruflichen Klassifizierung von diskreten Ausbildungsgängen vor (vgl. Andersson & Olsson, 1999). Beide Klassifikationen haben aber bislang keinen Eingang in die laufenden harmonisierten Berichterstattungen auf der EU-Ebene gefunden.

Eine dem deutschen Berufsbildungsbericht auch nur annähernd vergleichbare kontinuierliche und umfangreiche Berichterstattung ist – mit wenigen Ausnahmen (Australien) – nach den durchgeführten Recherchen in sonst kei-

nem Land verfügbar. Ebenso mangelt es – mit Ausnahme vor allem deutschsprachiger und einiger nordeuropäischer Länder – an entsprechenden, aktuellen statistischen Informationen (Registerdaten).

Ein weiteres Problem stellt sich damit, dass Ausbildungsprofile in den einzelnen Ländern entsprechend ihrer kulturellen Tradition unterschiedlich weit oder eng gefasst werden, eine unterschiedliche Dichte aufweisen und ihre Inhalte fachlich unbestimmt bleiben:

In den *USA* scheinen beispielsweise die 2-jährigen „Health care“-Ausbildungen, die nach deutschem Verständnis auf Krankenpflege ausgerichtet sind, sowohl die Ausbildung von Arzthelferinnen als auch von medizinisch-technischen Berufen einzuschließen (vgl. US Department of Education, 1996; National Center for Education Statistics, 2001; US Bureau of Labor Statistics, 2005).

Bei den „kaufmännisch-verwaltenden Berufe (Büroberufe)“ z. B. gibt es nicht nur Länder, die bislang keine fachlich einschlägigen Erstausbildungen in diesem Feld aufweisen (z. B. Norwegen); es zeigt sich darüber hinaus auch, dass die für die einzelnen Länder ausgewiesenen Ausbildungsgänge so gut wie keine oder doch nur wenig Überschneidungen aufweisen: Für Amerika beispielsweise können in diesem Bereich nur kommunale Verwaltungsberufe ausgewiesen werden, während diese Berufsausbildungen in anderen Ländern vollständig fehlen.

In Australien werden Ausbildungsprofile vielfach nach den betrieblichen Tätigkeitsbezeichnungen klassifiziert – jedoch nicht nach den über die absolvierten Module der „Training Packages“²⁴ definierten Inhalten.

Schließlich stellt sich das Problem, dass es zwar durchaus über viele Länder hinweg nach Niveau und fachlichem Zuschnitt vergleichbare Tätigkeiten/Ausbildungen gibt, diese aber – gemessen am gesamten Tätigkeitsfeld – nur einen kleinen Ausschnitt bzw. einzelne Tätigkeitsfelder gar nicht repräsentieren. Sowohl über die Analyse der internationalen Literatur als auch über die Diskussion auf beiden Workshops werden drei solcher Felder verstärkt ins Spiel gebracht: KFZ-Reparatur, Banking und qualifizierte Krankenpflege („registered Nurses“).

24 „Training Packages“ umfassen eine definierte Anzahl von frei kombinierbaren Modulen für häufig wirtschaftszweig-spezifisch definierte Qualifikationsbündel (vgl. National Training Information Service, 2005).

4.2. KONSEQUENZEN FÜR DIE SAMPLE-KONSTRUKTION

Die im internationalen Maßstab skizzierte

- Varianz des fachlichen Zuschnitts beruflicher Ausbildungsgänge,
- die (nach Niveau bzw. Alter differierende) ausgewiesene Vielfalt institutioneller Arrangements wie auch
- die erheblichen Unterschiede biografischer Arrangements im Übergang von der allgemeinbildenden Schule in eine spätere Berufstätigkeit (Kombination von formalisierten und informellen Ausbildungs- und Erwerbssequenzen)

in ihrer Gänze in dem Untersuchungsdesign der hier diskutierten Vergleichsstudie zu berücksichtigen, hieße diese zu überfordern. Dieses Grundproblem der Heterogenität beruflicher Bildungsprozesse war uns von Beginn an bewusst und führte uns zu der Einsicht, dass für die Vergleichsstudie kein Repräsentativ-Sample, das die gesamte Berufsbildung und alle Beschäftigungsfelder abdeckt, in Frage kommt. Mit dem im Rahmen der Machbarkeitsstudie entwickelten Mess- und Beobachtungskonzepten kommen weitere – sowohl forschungspragmatische als auch forschungsmethodische Gründe (Entwicklungsaufwand für fachliche Messkonzepte, geringe Besetzungszahlen von „Exotenberufen“ innerhalb der beruflichen Ausbildung) – ins Spiel, die eine Konstruktion geschichteter Stichproben nahe legen.

Damit ist zunächst eine grobe Richtung vorgegeben. Die weitere Suchstrategie zielt auf Sample-Konstruktionen, die aus dem Gesamt beruflicher Bildungs- und Einmündungsprozesse (in Erwerbstätigkeit) einen *wesentlichen* (wichtige Beschäftigungssegmente, wichtige institutionelle Formen) und *vergleichbaren Ausschnitt* (mittleres Niveau, ähnlicher fachlicher Zuschnitt, gleiches oder ähnliches Alter) *der beruflichen Erstausbildung* zu betrachten gestatten.

Die in diesem Zusammenhang auftauchenden – und im vorhergehenden Abschnitt ausführlich diskutierten – Probleme der Konstruktion international vergleichbarer Samples sind durchaus lösbar, freilich nicht alle in einem einzigen, homogenen Stichprobendesign und nicht ohne Kompromisse in der Gewichtung der Vergleichbarkeitskriterien. Die Gründe hierfür liegen vor allem in den schon erwähnten nationalen Unterschieden der institutionellen Verbindung von Ausbildungs- und Beschäftigungssystem einerseits und des Gewichts nicht-standardisierter Berufsbildungsbiografien andererseits. Was dies heißt, sei kurz erläutert:

- Die systematische Berücksichtigung nicht-standardisierter Berufsbildungsbiografien ist – bei einem vertretbaren Aufwand – nur über eine retrospektive Erhebung von Ausbildungserfahrungen innerhalb einer Querschnittsstudie möglich, bei der die Stichprobe aus der Beschäftigtenpopulation (Kriterien: Alter, Tätigkeit nach Niveau und funktionalem Zuschnitt) gezogen und die berufliche Bildungsbiografie rekonstruiert wird. Die Ein-

schränkung der Repräsentativität bei diesem Ansatz betrifft die institutionelle Seite der Ausbildung (Unschärfen nachträglicher Rekonstruktion von Niveau- und funktionaler Äquivalenz unterschiedlicher beruflicher Ausbildungsgänge).

- Wo – wie überwiegend in Deutschland – Ausbildungsfelder, -niveaus und -institutionen eng mit Tätigkeitsfeldern und -niveaus korreliert sind („Berufsprinzip“) und Ausbildungen weitgehend flächendeckend die mittleren Tätigkeiten abdecken, macht es dem ersten Anschein nach keinen Unterschied, ob man vertikal und horizontal vergleichbare Samples auf der Basis von Arbeitskräftestichproben (Tätigkeiten) oder aber von Registerdaten zur Ausbildung (Ausbildungsberufe) festlegt. Diese Art von regulierter Beziehung gilt jedoch selbst in Deutschland nicht für alle Beschäftigungssegmente; in anderen, insbesondere in den anglo-amerikanischen Ländern gilt dies nur für wenige Segmente bzw. Segmentausschnitte. Mit Blick auf eine Panel-Studie, deren Einstiegspunkt die Ausbildungsinstitutionen darstellen, ist daher im Prinzip der Weg eines *tätigkeitsstrukturbezogenen* Samplings angeraten. Gleichwohl bleibt – vor allem aus pragmatischen Gründen – der Weg eines rein *ausbildungstrukturbezogenen* Samplings zu erwägen. Die Unterschiede zwischen beiden Optionen liegen in dem noch zu leistenden Aufwand für die Vorbereitung einer Stichprobenziehung und in dem Gewicht, das die jeweilige Stichprobe in den Segmenten hat. Beide Wege verbinden sich mit einer Konstruktion beruflicher Teilsamples.

Erste Option: breite Samplekonstruktion (tätigkeitsstrukturbezogenes Sampling)

Diese Option wählt den Weg über die in einem bestimmten Alter (z. B. 24 Jahre) ausgeübte Tätigkeit und definiert das zu ziehende Ausbildungssample über Arbeitskräfteerhebungen.

Der Ansatzpunkt zur Definition von vergleichbaren Samples innerhalb der vordefinierten großen Beschäftigungssegmente (industriell-gewerbliche Berufe, kaufmännisch-verwaltende Berufe, Sozial-, Erziehungs- und Pflegeberufe, IuK-Berufe) ist die schon erwähnte ISCO88²⁵, die in den europäischen Arbeitskräfteerhebungen zur Anwendung kommt und sowohl das Niveau als auch das fachlich-funktionale Feld einer Tätigkeit ausweist. Auf diese Weise kann man – wie in dem hier aufgeführten Beispiel geschehen – Aggregate von vergleichsweise hoher Homogenität des fachlichen Zuschnitts wie auch des Niveaus definieren, die gleichwohl (zumindest für Deutschland) einen beträchtlichen Teil der Erwerbstätigen dieses Alters repräsentieren dürften (vgl. Tabelle 4.5):

Mit Ausnahme der Datenverarbeitungsfachkräfte (die mit einem ausgewiesenen Anteil von 1 % die IUK-Berufe nur unzureichend repräsentieren

25 International Standard Classification of Occupations, 1988.

dürften) bildet diese Auswahl jeweils gewichtige Felder der Erwerbstätigkeit jüngerer Beschäftigter und insgesamt immerhin 47 % der Erwerbstätigen in dieser Altersgruppe ab.

Die Aussagefähigkeit der ISCO-Daten in bezug auf Niveau und fachlichen Zuschnitt der Tätigkeiten kann insgesamt als gute Basis angesehen werden, auch wenn nicht zu übersehen ist, dass sie bei einzelnen Tätigkeitsgruppen an den Rändern Unschärfen aufweist und gerade – wie schon erwähnt – die IUK-Berufe nicht genau erfasst. Von daher ist – und dies ist der eigentlich aufwändige Schritt in der Vorbereitung einer solchen Sample-Konstruktion – ein Korrekturverfahren für die Bildung der Tätigkeitsaggregate vorzusehen, das schon mit Erfolg auf dem 2. Internationalen Workshop zur Ein- und Abgrenzung vergleichbarer Teilsamples eingesetzt wurde:

Tabelle 4.5: Ausgewählte Tätigkeitsfelder, ISCO-Code der feldspezifischen Tätigkeiten und Erwerbstätigenanteil des Feldes in der Altersgruppe der 15- bis unter 25jährigen nach Segmenten

Segment	Feld	ISCO-Code der feldspezifischen Tätigkeiten	Erwerbstätigenanteil des Feldes (15- bis unter 25jährige) (in %)
Gewerblich-technische Berufe (Industrie und Handwerk)	Metall- und Elektroberufe	72	15,0
Kaufmännisch-verwaltende Berufe	Handelsberufe Büroberufe	522, 523, 3416, [3419] 411, [343]	9,7 11,9
Sozial-, Erziehungs- und Pflegeberufe	Gesundheitsfachkräfte	322, 323	8,8
IuK-Berufe	Datenverarbeitungsfachkräfte	312	1,0
Alle Berufe			100,0

Das Korrekturverfahren basiert auf der vergleichenden Analyse von Arbeitsanforderungen qualifizierter Tätigkeiten, die nach der Hauptfunktion (also Produktion, Distribution, personenbezogene Dienstleistungen) wie auch den fachlichen Schwerpunkten unterschieden werden. Die auf dieser Basis entwickel-

ten Anforderungsprofile können dann in einem iterativen Verfahren in die Samplebildung einbezogen werden: Die Anforderungsprofile dienen einerseits als Referenzpunkt zur jeweils nationalen Ermittlung der Tätigkeiten, die unter dieses Profil zu subsumieren sind; auf diese Weise lassen sich auch Tätigkeiten zuordnen, die nach der ISCO 88 benachbarten Niveaus oder aber anderen Tätigkeitsbündeln zugeordnet werden. Umgekehrt lassen sich über die in diesem Verfahren aufgeworfenen Zuordnungsprobleme auch die Konturen der Anforderungsprofile unter einer internationalen Perspektive schärfen, so dass sie als Grundlage für die Auslegung der Messinstrumente in diesem Feld dienen können.

Die auf diese Weise entstehenden Aggregate können im Weiteren auf Basis nationaler Arbeitskräfteerhebungen mit dort vorhandenen Angaben zum Ausbildungshintergrund verknüpft und auf diese Weise die (quantitativ) wichtigsten Ausbildungsgänge in dem Tätigkeitsfeld und auf dem Niveau entschlüsselt werden. Diese Ausbildungsgänge bilden dann die Basis für die Ziehung einer institutionen-basierten Stichprobe. Das Vorgehen ist gleichwohl nicht ganz problemlos, weil – insbesondere für die anglo-amerikanischen Länder – nicht ausgeschlossen werden kann, dass solchermaßen breit definierte Stichproben mit einer Vielzahl institutioneller Arrangements einhergehen, die in der Stichprobenziehung zu berücksichtigen erheblichen Stratifizierungsaufwand mit sich bringen würde – sowohl was die Berücksichtigung unterschiedlicher Trägerstrukturen und regionaler Disparitäten als auch die Bildung altershomogener Gruppen anbelangt.

Zweite Option: Enge Sampledefinition (Auswahl verbreiteter und definierter Ausbildungsgänge)

Diese Option wählt den Weg über Ausbildungsgänge mit einem vergleichbaren curricularen Zuschnitt, die auf eine qualifizierte Tätigkeit in den großen Beschäftigungssegmenten vorbereiten und von Personen in einer bestimmten Altersspanne (zwischen 16 und 20) aufgenommen werden. Dieser Weg beinhaltet mehrere Schritte:

Der erste Schritt besteht in der Erstellung von zwei Übersichten: Die eine enthält die Vorauswahl der designierten Ausbildungsfelder (vgl. Tabelle 4.5), die andere die auf der Basis der erwähnten EUROSTAT/CEDEFOP-Daten erstellte und um Informationen für die außereuropäischen Ländern ergänzte Vorauswahl von Programmen²⁶, über die diese Ausbildungsfelder abgedeckt werden.

26 Programme meint institutionell unterscheidbare Bereiche der beruflichen Ausbildung, die durch die Zugehörigkeit zu Subsystemen des (Berufs-)Bildungssystems (z. B. duales System), durch die Charakteristik des erreichbaren Abschlusses oder des Curriculums definiert werden.

Der zweite Schritt bestand in einer Diskussion dieser Auswahl mit den ausländischen Experten, zu der wir des Weiteren die – schon erwähnten – Anforderungsprofile wichtiger Tätigkeiten heranzogen wie auch Informationen zur Verbreitung einzelner Ausbildungsgänge in anderen Ländern.

Auf dieser Grundlage – und unter Einbezug der von den Experten explizit gemachten Vorschläge – wurde der Vorschlag für ein Sample erarbeitet, das sowohl in seiner fachlich-funktionalen Struktur die großen Beschäftigungssegmente widerspiegelt als auch immer noch eine nach Niveau und Alterstruktur vergleichbare Samplekonstruktion gestattet (vgl. Tabelle 4.6).

Tabelle 4.6: Ausbildungsgang, Spektrum des ISCED-Level, Spanne des Eintrittsalters

Tätigkeitsfeld	Ausbildungsgang	Spektrum ISCED-Level (international)	Eintrittsalter (Spanne)
Metall- und Elektroberufe	KFZ-Mechatroniker	3	16–20
	Industriemechaniker	3	16–20
	Elektroniker	3	16–20
	(Industrie)*		
Kaufmännisch- verwaltende Berufe	Bankkaufleute	3–5	17–20
Gesundheitsdienst- berufe	Krankenpflege	4–5	18–20
IuK-Berufe	Fachinformatiker	3–5	17–20

* Fachrichtungen: Betriebstechnik, Automatisierungstechnik, Geräte und Systeme.

Der Bereich des Handwerks wird durch den *KFZ-Mechatroniker* repräsentiert. Hierbei handelt es sich um ein auf europäischer Ebene entwickeltes Berufsbild, zu dem wir jedoch Pendanten auch in den USA und in Australien finden. Der Bereich der industriellen Metall- und Elektroberufe wird hier durch die *Elektroniker für Betriebstechnik, Automatisierungstechnik sowie Geräte und Systeme*²⁷ sowie den *Industriemechaniker* abgedeckt; möglicherweise muss man in letzterem Fall noch einmal umdenken und anstatt dessen auf den *Werkzeugmacher* übergehen.²⁸ Die *Bankkaufleute*, *Fachinformatiker* wie auch

27 Die drei Berufs fassen wir im Folgenden – verkürzt – unter der Bezeichnung „Elektroniker Industrie“.

28 Die auf dem 2. Internationalen Workshop geführte Diskussion über das Anforderungsprofil des Industriemechanikers hat deutlich gemacht, wie schwierig dessen curriculare Äquivalente in anderen Ländern bestimmbar sind. Umgekehrt scheint der deutlich enger gefasste Werkzeugmacher (oder besser: Konstruktionsmechaniker) mit dem „Tool

die *Krankenpflegeberufe* stehen für Ausbildungen, die teilweise in anderen Ländern formal auf höherem Niveau angesiedelt sind als in Deutschland, in ihrem Curriculum aber offenbar große Ähnlichkeiten aufweisen.

Auch bei dieser Sample-Konstruktion sind noch nicht alle Probleme gelöst, wie die Diskussion auf dem 2. Workshop zeigte. Doch die auch bei unseren Experten in Einzelfällen vorhandenen Unsicherheiten, welches ihrer im Niveau differierten Programme den Samplevorgaben am besten entspricht (z. B. im Bereich der Krankenpflege), lassen sich vergleichsweise schnell lösen. In solchen Zweifelsfällen ist eine Rückkopplung der nationalen Forschergruppen mit ihren jeweiligen statistischen Ämtern vorzusehen und ein entsprechender Vorschlag in das Koordinationsgremium einer solchen Studie zurückzuspielen.

Bei aller Unsicherheit über die aus Einzelangaben ermittelten Spannen des Einstiegsalters deuten sich bei dieser Option Möglichkeiten einer altershomogenen Konstruktion des Gesamtsamples an. Die Schnittmenge hierfür wäre das Alter von 18/19 Jahren. Dies hieße letztlich, in einer am Beginn der Ausbildung angesiedelten Erst-Befragung mit diesen Jahrgängen zu starten (z. B. 18-jährige Ausbildungsanfänger).

Aufgrund der spezifischen Ausgangspunkte der beiden Konstruktionen (breit definierte Tätigkeiten vs. Curricula einzelner Ausbildungsgänge für – im internationalen Maßstab – vergleichsweise stark regulierte Beschäftigungen) ist bei der letztgenannten Option die Wahrscheinlichkeit gering, dass neben den auf eine designierte qualifizierte Tätigkeit zugeschnittenen und hier ausgewiesenen Ausbildungen weitere bedeutsame Ausbildungsgänge existieren, mit der diese Tätigkeit bewältigt wird – auszuschließen ist es freilich nicht. Von daher ist die bei der zweiten Option vorgenommene Auswahl über eine tätigkeitsspezifische Analyse des Ausbildungshintergrundes zu kontrollieren und gegebenenfalls zu ergänzen.

Zur Durchführung der Untersuchung – Ziehung der Stichprobe, Messung und Beobachtung

Die oben entwickelten beiden Optionen vergleichbarer Samplekonstruktion sind prinzipiell offen für unterschiedliche Vorgehensweisen in der Organisation von Stichprobe und Studie, legen aber zusammen mit den Grundprinzipien der Konstruktion von Mess- und Analysekonzepten bestimmte Grenzen des Designs fest, innerhalb derer Variationen möglich sind. Diese Grenzen sind bestimmt durch die Verbindung von mikro- und makrostruktureller Analyse, von Messung beruflicher Kompetenzen und Institutionenanalyse sowie durch eine prozessbezogene Untersuchungsperspektive (in der unterschiedliche zeitliche Ereignisse/Sachverhalte und Erfahrungen im Sinne einer Ursa-

and die maker“ – nicht nur, aber eben auch in den anglo-amerikanischen Ländern – ein vergleichbares Pendant zu haben.

chen-Wirkungsanalyse in einen kausalen Zusammenhang mit Ergebnissen der Kompetenzmessung gebracht werden sollen).

Folgt man der jüngeren forschungsmethodischen Literatur wie auch der Untersuchungsanlage international vergleichender Studien zur Kompetenzmessung, so spitzt sich unter diesen Prämissen die Frage des Vorgehens auf die Wahl zwischen einer *haushalts-basierten* und einer *institutionen-basierten Konstruktion* einer *repräsentativen Stichprobe* sowie zwischen einer *Querschnittsuntersuchung* (mit retrospektiven Elementen) und einer „echten“ Längsschnittuntersuchung (Panel) zu (vgl. Expertise Kühnel; Schnell, Hill & Esser, 1989; National Center for Education Statistics, 1998; COM/DELSA/EDU, 2004; Baumert u. a., 2001). Die Unterschiede im Vorgehen lassen sich wie folgt skizzieren:

- *Ziehung der Stichprobe:* Bei *Haushaltstichproben* wird die Zielpopulation in einem zweistufigen Verfahren ausgewählt: Zunächst werden die in Frage kommenden Personen (in einem bestimmten Ausbildungsgang/einer bestimmten Tätigkeit) ausgewählt und diese dann nach ihren eigenen Angaben Institutionen (Ausbildungseinrichtungen, Betrieben) zugeordnet. Ebenfalls zweistufig wird bei einer *institutionen-basierten Stichprobe* vorgegangen – nur umgekehrt: Auf der ersten Stufe werden Bildungseinrichtungen bzw. Betriebe erhoben und auf der zweiten Ebene Teilnehmer an Ausbildungsgängen bzw. Betriebsangehörige in den entsprechenden beruflichen Feldern. Wenn die Einrichtungen/Betriebe zufällig ausgewählt werden, dann handelt es sich auch in diesem Fall um eine repräsentative Stichprobe der Zielpopulation (vgl. Expertise Kühnel, im Anhang).
- *Organisation der Studie:* In *Querschnittsuntersuchungen* würde man bei den jeweils gezogenen Samples die Kompetenzen zu einem bestimmten Zeitpunkt (Ende der Ausbildung, nach dem Übergang in Beschäftigung) messen und die aktuellen Umgebungsbedingungen erheben. Um mögliche Ursachen und Wirkungen der Messergebnisse zu identifizieren, muss dieser Untersuchungstyp mit Verfahren retrospektiver Analyse kombiniert werden: Auf der *Individual-Ebene* würde dies durch die Aufnahme sozio-kultureller und -ökonomischer Hintergrundvariablen und biografischer Ereignisse im Interview geschehen. Auf der *institutionellen Ebene* könnte dies durch Querschnittsuntersuchungen zu verschiedenen Zeitpunkten (Trendstudien mit mehreren Wellen – Alterskohortendesign) bzw. dadurch realisiert werden, dass zum gleichen Zeitpunkt mehrere Stichproben untersucht werden, deren Angehörige sich in verschiedenen Phasen des Durchlaufs durch die Ausbildung bzw. des Übergangs in Beschäftigung befinden („Quasi-Panel“ – Institutionelles Kohortendesign). Demgegenüber würde man im Rahmen einer *echten Längsschnittuntersuchung* (Panel im engeren Sinn) bei den ausgewählten Teil-Samples die Kompetenzen am Anfang, am Ende und nach Übergang auf den Arbeitsmarkt messen und zugleich die Umgebungsbedingungen über persönliche Interviews und Re-

cherchen in den Institutionen (Ausbildungseinrichtung, Betriebe) erheben (vgl. Expertise Kühnel).

Die Entscheidung für jeweils eine der beiden Varianten bzw. eine bestimmte Kombination fällt aus mehreren Gründen nicht leicht: Ein wesentlicher Grund liegt darin, dass bestimmte Ansätze in internationalen Studien zur Kompetenzmessung besonders verbreitet sind, für andere auf internationaler Ebene (fast) gar keine Erfahrungen vorliegen. Betrachtet man sowohl die im schulischen Bereich angesiedelte PISA-Studie als auch die im Kontext der Diskussion zum lebenslangen Lernen entstandenen Studien zu Kompetenzen der Bevölkerung im erwerbsfähigen Alter („International Adult Literacy Survey/IALS“, „Adult Literacy and Life Skills Survey/ALL“, „Programme for the International Assessment of Adult Competencies/PIAAC“), so zeigt sich ein Übergewicht haushalts-basierter Stichprobenkonstruktionen (Ausnahme: PISA), die mit einem retrospektiven Querschnittsdesign kombiniert wurden (National Center for Education Statistics, 1998; COM/DELSA/EDU, 2004; Baumert u. a., 2001). Dies macht in dem Rahmen der jeweiligen Studien durchaus Sinn, bieten beide Ansätze doch eine Reihe von Vorteilen gegenüber anderen Ansätzen: Über Haushaltsstichproben kann prinzipiell der Zugang auch zu den Teilen der Grundgesamtheit von Personen und Individuen gewährleistet werden, die eher untypische Bildungsverläufe aufweisen (und damit in einer institutionen-basierten Stichprobe unterrepräsentiert wären). Querschnittsdesigns sind grundsätzlich weniger aufwändig als Wiederholungsbefragungen (vgl. Expertise Kühnel im Anhang).

Eine genauere Analyse der Studien zeigt freilich, dass diese Studien unter methodischen Gesichtspunkten nur begrenzten Vorbild-Charakter für die hier diskutierte Vergleichsstudie haben können (vgl. National Center for Education Statistics, 1998; Baumert u. a., 1999, COM/DELSA/EDU, 2004); denn die Organisation der Stichprobenziehung und der Studie scheint erheblich von der Art der erhobenen Kompetenzen (und ihrer Messverfahren) sowie der Bedeutung abzuhängen, welche die intensive (prozessbezogene) Ausleuchtung institutioneller Kontextbedingungen im Rahmen der Untersuchung hat:

Nimmt man diese Logik auf, dann spricht die spezifische Mess- und Beobachtungskonzeption der hier diskutierten Vergleichsstudie für eine institutionen-basierte Ziehung berufsfachlicher Stichproben, die als Panel organisiert werden. *Gegen* eine haushaltsbasierte Ziehung der Stichprobe spricht zusätzlich der hohe Aufwand: Auf diesem Weg wird eine Vielzahl von Personen kontaktiert, die nicht zur Zielpopulation (berufliche Teilsamples) zählen (was selbst aufwändig recherchiert werden muss). Für eine institutionen-basierte Samplekonstruktion, die als Panel organisiert wird, sprechen sowohl pragmatische als auch methodische Gesichtspunkte: Der Zugang über Institutionen erleichtert die Kontaktaufnahme mit der Untersuchungspopulation, die zeitkritische Aufnahme von Kompetenzen und relevanten Ereignissen in der Umwelt an verschiedenen Punkten des Übergangsprozesses von der Schule auf den Ar-

beitsmarkt und lässt kurz- und langfristige Ursachen wie Wirkungen verlässlich analysieren (vgl. Expertise Kühnel, Schnell, Hill & Esser, 1989).

Dennoch ist bei einer echten Panel-Konstruktion folgendes zu beachten: Wiederholungsbefragungen derselben Population oder Institution sind – in Abhängigkeit von der Zahl der Befragungszeitpunkte und dem Abstand zwischen ihnen – grundsätzlich aufwändiger als Querschnittsuntersuchungen und bergen spezifische Risiken: Hierzu zählt zum einen ganz allgemein die Panel-Mortalität, d. h. die Ausfälle, die dadurch entstehen, dass die Erstbefragten zu einem späteren Zeitpunkt nicht erreichbar sind oder die Auskunft verweigern. Solche Ausfälle sind durch systematische Panelpflege minimierbar und durch Erhöhung der Ausgangsstichprobe korrigierbar. Dies gilt weniger für die in diesem Untersuchungsfeld erwartbare spezifische Struktur der Ausfälle (Ausfall besonders mobiler Jugendlicher, Ausfälle durch Wechsel in von der Untersuchung nicht erfasste Berufsfelder²⁹), die Messergebnisse systematisch verzerren können, was nur über zusätzliche Kontrollgruppen aufgefangen werden kann (Expertise Kühnel).

Es bleibt zu klären, wie die beruflichen Teilsamples als repräsentative Stichproben gezogen werden können und welchen Einfluss dies auf den Stichprobenumfang hat.

Die jeweiligen Stichproben der Teilsamples werden über eine *2-stufige Ziehung* rekrutiert. In dieser werden auf der ersten Stufe die bei den Teilsamples in Frage kommenden Bildungseinrichtungen (eventuell bei betrieblich gebundenen Ausbildungen: Betriebe) erhoben und auf der zweiten Ebene Teilnehmer an den entsprechenden Ausbildungsgängen ausgewählt. Eine Zufallsauswahl der Einrichtungen (Betriebe) sichert die Repräsentativität der jeweiligen Teil-Samples ab. Allerdings birgt die Bildung solcher Cluster-Stichproben, in denen ganze Schulklassen bzw. alle die Auswahlkriterien erfüllenden Personen eines Betriebes befragt werden, die Gefahr von Verzerrungen in der Auswahl (Intraklassenkorrelation), die durch eine Erhöhung der Fallzahlen nach einem bestimmten Korrekturfaktor ausgeglichen wird (vgl. Expertise Kühnel).

Wie groß der Umfang einer Stichprobe im Rahmen des vorgeschlagenen Stichprobenplans (Teilstichproben mehrerer Ausbildungsgänge/-felder, institutionen-basierte Ziehung, Panel-Konstruktion) in den einzelnen Ländern ausfallen muss, lässt sich zurzeit nicht detailliert ermitteln. Allerdings lässt sich eine Modellrechnung für Deutschland vornehmen, in die folgende Annahmen eingehen:

29 Angespielt wird damit auf das Problem, dass wir bei einem Panel, das die Entwicklung domänenspezifischer Kompetenzen misst, in einem nicht unerheblichen Maße mit Personen zu tun bekommen können, die das Berufsfeld nach Beginn der ersten Ausbildung wechseln. Unabhängig davon, ob man diese Personen als „Ausfälle“ betrachtet oder aber ihre Kompetenzen mit einem anderen Messkonzept erfasst, verkleinert sich dadurch tendenziell das Sample.

- die geschätzte Größe der zum dritten Messzeitpunkt benötigten Stichprobe (mit einer Genauigkeit von 5 % vorgenommene einfache Zufallsauswahl),
- eine Panel-Mortalität von 50 % (durchschnittliche Größe bei einem Beobachtungszeitraum von 6 Jahren) sowie
- ein Korrekturfaktor für die Verwendung von Clusterstichproben (der für Schulen mit 2.9, für Betriebe mit 1.08 angesetzt wurde).

Auf dieser Basis ergibt sich folgender *Vorschlag für die Samplekonstruktion* (vgl. Tabelle 4.7):

Tabelle 4.7: Vorschlag für Samplekonstruktion (Modellrechnung)

	Größe rechnerische Basisstichprobe (einfache Zufallsauswahl, Fehlermarge 5 %)	Erhöhung des Umfangs durch geschätzte Panel-Mortalität (50 %)	Erhöhung des Umfangs durch Korrekturfaktor für Clusterung (1.08 bzw. 2.9)	Stichprobenumfang für Panelstudie
KFZ-Mechatroniker	400	+ 400	+ 64	= 864
Industriemechaniker	400	+ 400	+ 64	= 864
Elektroniker (Industrie)*	400	+ 400	+ 64	= 864
Bankkaufmann	400	+ 400	+ 64	= 864
Krankenpfleger	400	+ 400	+ 1520	= 2320
Fachinformatiker	400	+ 400	+ 64	= 864
Alle Ausbildungsgänge				= 6640

* Fachrichtungen: Betriebstechnik, Automatisierungstechnik, Geräte und Systeme.

Für die insgesamt 6 Ausbildungsgänge werden zum dritten Messzeitpunkt jeweils Stichproben von etwa 400 Personen benötigt. Mit Blick auf die avisierte Laufzeit des Panels und die erwartbare Mortalitäts-Rate muss dieser Ansatz auf jeweils 800 erhöht werden. Da die erwartbare Intraklassenkorrelation in Schulen deutlich höher ist als in Betrieben, ergibt sich bei dem Ausbildungsgang „*Krankenpflege*“ ein höherer Korrekturfaktor und führt bei dieser Gruppe zu einer Ausgangsstichprobe für das Panel, die 2320 Personen beträgt. Bei den anderen Ausbildungsgängen, dem *KFZ-Mechatroniker*, den *Industriemechaniker*

kern für Betriebstechnik, Automatisierungstechnik, Geräte und Systeme, dem Elektroniker, dem Bankkaufmann wie auch dem Fachinformatiker wird hingegen eine deutlich kleinere Ausgangsstichprobe benötigt, die jeweils 864 Personen umfasst. Insgesamt, d. h. über alle hier aufgeführten Ausbildungsgänge hinweg, läuft dieser Vorschlag auf ein Sample von 6640 Personen hinaus (vgl. Tabelle 4.7), die in der ersten Welle eines Panels befragt werden.

5. UNTERSUCHUNGSANLAGE FÜR EIN BERUFSBILDUNGS-PISA – ALTERNATIVE MODELLE

5.1. ZUR AUSGANGSLAGE

Rekapitulieren wir zunächst noch einmal die Aufgabe der Machbarkeits-Studie: Entwicklung von Perspektiven für eine internationale Vergleichsstudie in der beruflichen Bildung, die als „large scale assessment“ angelegt sein soll, und Prüfung der Bedingungen und Probleme ihrer Realisierung. Ziel eines Berufsbildungs-PISA sollte die Messung von Kompetenzen sein, die in der beruflichen Bildung vermittelt werden, sowie die Analyse der mikro- und makrostrukturellen Bedingungen, unter denen die Berufsbildung steht und die den Vermittlungserfolg beeinflussen. Über die vergleichende Messung von Kompetenzen soll erfasst werden, was in beruflichen Bildungsprozessen tatsächlich gelernt wird und was in Abschlusszeugnissen nicht hinreichend abgebildet werden kann, um hierüber Stärken und Schwächen unterschiedlicher Berufsbildungssysteme systematisch zu beleuchten. Die Kompetenzmessung sollte dynamisch angelegt sein, d. h. Kompetenzzuwächse erfassen und eine Verknüpfung der Messergebnisse mit den mikro- und makrostrukturellen Bedingungen ermöglichen, unter denen Kompetenzen erworben und genutzt bzw. verwertet werden.

Anders als bei anderen, auf Schülerpopulationen bezogenen „large scale assessments“ (wie PISA und TIMSS) kann ein Berufsbildungs-PISA sich nicht mit der Messung allgemeiner Grundkompetenzen bescheiden. Wer Berufsbildungs-PISA sagt, geht die Verpflichtung zur Messung funktional-fachlicher arbeitsbezogener Kompetenzen ein. Das macht die besondere Schwierigkeit eines Berufsbildungs-PISA aus und stellt die eigentliche Herausforderung für die Machbarkeits-Studie dar.

Die vorgelegte Studie hat mit Analyse des Forschungsstands und mit Hilfe internationaler Expertise (zwei mehrtägige Workshops) drei zentrale Prüfschritte für die Durchführbarkeit eines Berufsbildungs-PISA vollzogen:

- Prüfung der Konzepte zur Kompetenzmessung und Entwicklung von Vorschlägen für den Bereich der Berufsbildung (vgl. Kapitel 2);
- Klärung der makro- und mikro-strukturellen Bedingungen, welche die Qualität von Ausbildungsprozessen, mithin die Kompetenzentwicklung, beeinflussen und als Bezugspunkt für die Ergebnisse der Kompetenzmessung ins Erhebungsprogramm aufgenommen werden müssen (vgl. Kapitel 3);

- Ausleuchtung der Probleme internationaler Komparatistik in der Berufsbildung mit dem Ziel, die systemischen Koordinaten für die Untersuchungsanlage und die Sample-Konstruktion herauszuarbeiten und in einen Vorschlag umzusetzen (vgl. Kapitel 4).

Auf dem Hintergrund dieser Prüfschritte kommen wir zu dem Schluss, dass ein internationaler Vergleich der Leistungsfähigkeit von Berufsbildungssystemen über ein „large scale assessment“ zwar schwierig, aber aus wissenschaftlicher Perspektive möglich ist. Die von uns durchgeführten internationalen Workshops haben neben einer durchgängig positiven Beurteilung der Wichtigkeit eines Berufsbildungs-PISA durch alle Beteiligten auch die Einschätzung der zu überwindenden Schwierigkeiten kritisch bestätigt und klar gemacht, dass noch ein beträchtliches Ausmaß internationaler Entwicklungsarbeit zur Erstellung eines Untersuchungs-Designs erforderlich ist. Diese Arbeit betrifft nicht allein die wissenschaftliche Seite der Erarbeitung von Messmethoden für Kompetenzen und von Analyseschemata für die Erfassung institutioneller und individueller Kontextbedingungen, welche die Ausbildung, das Ausbildungsverhalten und die Ausbildungsverwertung beeinflussen. Methodenentwicklung in diesem Feld hat politische Implikationen hinsichtlich der Definition von Ausbildungszielen und -formen, die eine Konsensbildung über politische, ökonomische und kulturelle Grundlagen von Berufsbildung nötig machen. (Spätestens bei Klärung der Bereitschaft der politischen Administrationen in den Teilnehmerländern zur Finanzierung eines Berufsbildungs-PISA, die von einer Reihe der Experten auf den internationalen Workshops durchaus positiv eingeschätzt wurde, dürfte das Konsensbildungs-Problem virulent werden.)

Im Folgenden verdichten wir die Resultate aus der Auseinandersetzung mit dem Forschungsstand und aus den internationalen Experten-Workshops zu Vorschlägen für die Untersuchungsanlage eines Berufsbildungs-PISA. Der Plural „Vorschläge“ erklärt sich daraus, dass es unter forschungspragmatischen Gesichtspunkten nicht nur *eine* Lösung gibt, sondern abzuwägen ist, wie sich das Verhältnis von Aufwand und Ertrag einer Vergleichs-Studie darstellt und wie die bestmögliche Zielerreichung mit einem vertretbaren Ressourceneinsatz und in einem kalkulierbaren Zeitraum gewährleistet werden kann.

Die Anforderungen an ein wie immer geartetes Untersuchungs-Konzept sind nicht allein durch den Charakter als large scale assessment und die Verbindung der Kompetenzmessung mit den mikro- und makrostrukturellen Bedingungen des Kompetenzerwerbs bestimmt. Die Messung soll in quantitativ relevanten Berufsfeldern erfolgen und die wichtigsten Wege des Kompetenzerwerbs in den Ländern erfassen. Unter diesen Voraussetzungen stellen wir zwei Grundmodelle eines internationalen Berufsbildungs-PISA vor, die in sich weiter modifizierbar sind:

- erstens eine echte Panel-Längsschnittstudie mit mehreren Messzeitpunkten;
- zweitens eine Querschnittsstudie.

Beide Konzepte sind unter dem Gesichtspunkt der Ziele und der Forschungspragmatik einer Vergleichsstudie zu erörtern. (Auch hierbei können wir auf die Diskussionen in den internationalen Workshops zurückgreifen.)

5.2. DAS FORSCHUNGS-DESIGN EINER LÄNGSSCHNITTSSTUDIE

Unter messtheoretischen Gesichtspunkten und in der Perspektive eines internationalen Vergleichs der Leistungsfähigkeit von Berufsbildungsinstitutionen kann man eine Längsschnittstudie, die wenigstens drei Panel-Wellen umfasst, als den „Königsweg“ eines vergleichenden „large scale assessment“ bezeichnen (vgl. Abbildung 5.1).

Abbildung 5.1: Panel-Konstruktion: Systematik der einzelnen Wellen

	1. Welle	2. Welle	3. Welle
Stichprobenziehung	Institutionenbasierte Stichprobe	Panel	Panel
Kompetenzmessung	Eingangskompetenzen, Motivation	In der Ausbildung erworbene Kompetenzen (Kompetenzzuwachs)	Erweiterung der Kompetenzen nach Ausbildungsabschluss, Verwertung der Kompetenzen in der Arbeit, auf dem Arbeitsmarkt und in der Biografie
Erhebung institutioneller Daten	Systemische Kontextbedingungen, institutionelle Struktur der Ausbildungseinrichtungen	Merkmale und Organisation der Ausbildungseinrichtungen, Ausbildungsbedingungen und Unterrichtsgestaltung, Lernklima	Branchen- und berufsbezogene Erwerbsbedingungen, Regulationsformen des Übergangs in Beschäftigung oder auf dem Arbeitsmarkt
Erhebung von Individualdaten	Lebens- und Lernbedingungen, Informations- und Entscheidungsverhalten bei der Berufswahl, Bildungsaspiration, Lern- und Arbeitsbiografie	Besuchte Programme, Lerninhalte und -orte, Art der Arbeitserfahrungen, Wahrnehmung der Ausbildungsbedingungen und -situationen, Lernzeit, biografische Planungsperspektiven	Abschluss (zertifiziert), Erwerbsstatus, Stellung im Beruf, Einkommen, ausgeübte Tätigkeit, Weiterbildungsverhalten, berufliche Mobilität

Die Untersuchungsanlage

Die erste Panel-Welle soll kurz nach Beginn der Ausbildung stattfinden, um das Kompetenz-Eingangsniveau zu ermitteln.

Die zweite Panel-Welle sollte – unter der Voraussetzung einer dreijährigen Ausbildung – kurz vor Abschluss der Ausbildung nach gut zweieinhalb Jahren durchgeführt werden, um den während der Ausbildung erzielten Zue-
winn an Kompetenz zu messen.

Die dritte Welle stünde drei Jahre nach Abschluss der Ausbildung an, wenn die Auszubildenden den Übergang in den Arbeitsmarkt geschafft haben und einer Beschäftigung nachgehen oder sich in weiterführenden Bildungseinrichtungen befinden – oder aber auch arbeitslos sind. Ziel dieser Welle ist die Erhebung der Nachhaltigkeit von Kompetenzen, ihrer Erweiterung und ihrer Verwertung im Berufsleben sowie in der persönlichen Biografie.

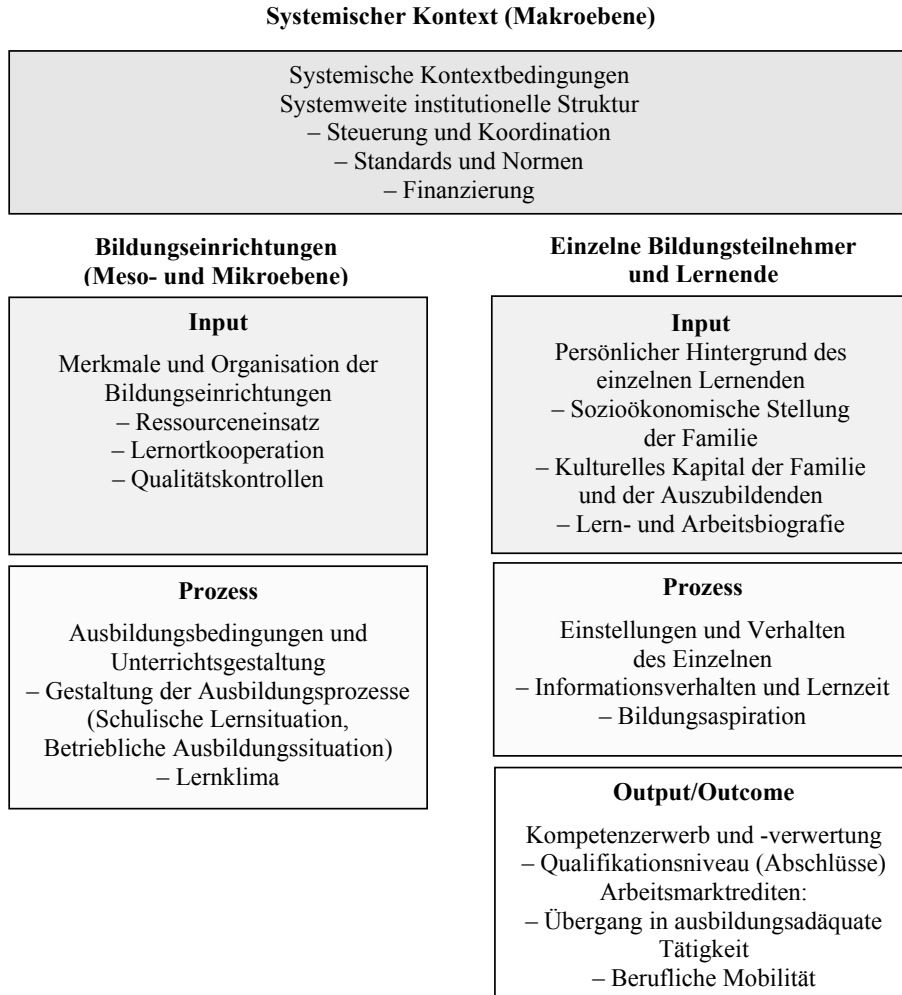
Das Erhebungsprogramm

- Kompetenzmessung in den allgemeinen, berufsübergreifenden und berufsfachlichen Kompetenzbereichen (vgl. Abbildung 5.2)
- Institutionelle und individuelle Lernkontexte (vgl. Abbildung 5.3)

Abbildung 5.2: Erhebungsplan Kompetenzmessung

Berufsübergreifende und berufsfachliche Kompetenzbereiche (Performanz in verschiedenen berufsspezifischen Kontexten)			
Individuelle Fähigkeiten (an Anspruch ge- nommen und in- terpretiert in ver- schiedenen Kontexten)	Selbstkompetenz		
	Sachkompetenz	Methodenkompetenz	Sozialkompetenz
	<i>(theoretische, analytische Anforderungen)</i>	<i>(technische, funktionale Anforderungen)</i>	<i>(interpersonale Anforderungen)</i>
	„Umgang mit Konzepten“	„Umgang mit technischer Ausstattung“	„Umgang mit anderen“
	Selbstwirksamkeit (Selfefficacy)		
Einstellungen, Werte, Wahrnehmungen	Handlungskontrolle in der Leistungssituation/ Interesse und Motivation/Selbstkonzept		
Antriebe	Lernstrategien		
Motivation			
Metakognitive Strategien			
Deklaratives Wissen	Komplexe Aufgaben zur simultanen Erfassung von Sach- und Methodenkompetenz		Fragebogen zu „kritischen sozialen Situationen“ (enthält Items zum deklarativen, pro- zeduralen und strategischen Wissen im Bereich der Sozialkompetenz)
Prozedurales Wissen			
Strategisches Wissen	Je nach Berufsfeld:		
	Komplexe Aufgabe zur Sachkompetenz	Komplexe Aufgabe zur Methodenkompetenz	
Allgemeine kognitive Grundkompetenzen (analog PISA)			

Abbildung 5.3: Erhebungsplan institutionelle und individuelle Kontextbedingungen



Die Stichprobe

- Berufsfelder: (vgl. Tabelle 4.6)
- Stichprobenplan für Deutschland: (vgl. Tabelle 5.1)

Tabelle 5.1: Stichprobenplan Deutschland

	Größe rechne- rische Basis- stichprobe (einfache Zufallsaus- wahl, Fehler- marge 5%)	Erhöhung des Umfangs durch ge- schätzte Panel- Mortalität (50%)	Erhöhung des Umfangs durch Korrekturfaktor für Clusterung (1.08 bzw. 2.9)	Stichprobenum- fang für Panel- studie
KFZ- Mechatroniker	400	+ 400	+ 64	= 864
Industrie- mechaniker	400	+ 400	+ 64	= 864
Elektroniker (Industrie)*	400	+ 400	+ 64	= 864
Bankkaufmann	400	+ 400	+ 64	= 864
Krankenpfleger	400	+ 400	+ 1520	= 2320
Fachinformatiker	400	+ 400	+ 64	= 864
Alle Ausbil- dungsgänge				= 6640

* Fachrichtungen: Betriebstechnik, Automatisierungstechnik, Geräte und Systeme.

Institutionen-Analyse

Als Ergänzung zur Individualbefragung schlagen wir eine Institutionen-Befragung vor. Sie soll durch Expertengespräche und Beobachtungen die institutionellen Kontextbedingungen der Berufsbildungsprozesse erfassen. Diese Ergänzung erscheint uns bei einer institutionenbasierten Stichprobe nötig, um im Vergleich der Ausbildung in Schulen und Betrieben Unterschiede in den gemessenen Kompetenzen in Beziehung zu schulischen und betrieblichen Arrangements setzen zu können. Das Erhebungsprogramm folgt den Aspekten und Indikatoren der Hauptbefragung. Die Stichprobe wird durch die in die Hauptuntersuchung einbezogenen Schulen, Betriebe und überbetrieblichen Einrichtungen gebildet.

Varianten des Untersuchungs-Designs

Zu dem aufgeführten Hauptvorschlag eines Panel-Designs mit drei Wellen lassen sich zwei Varianten denken, die schneller zu Ergebnissen führen, aber mit Einbußen bezogen auf die Reichweiten der Erkenntnisse verbunden sind:

- Die erste Variante besteht in einem Panel mit nur zwei Messzeitpunkten, deren erster am Beginn und deren zweiter am Ende der Ausbildung liegt. Im Zentrum dieser Variante steht die Messung des Kompetenzzuwachses während der Ausbildung mit Rückbezügen auf die institutionellen und individuellen Kontextbedingungen. Fehlen würde hier die Verwertungsseite der erworbenen Kompetenzen auf dem Arbeitsmarkt, in der Arbeit und in den realisierten Bildungsperspektiven nach Ausbildungsabschluss.
- Die zweite Variante sieht auch zwei Messzeitpunkte vor, jetzt aber am Ende der Ausbildung und drei Jahre nach Abschluss. Im Zentrum dieser Variante steht die Messung des Outcomes der Ausbildung in Arbeit, Arbeitsmarkt und biografischen Perspektiven, das Verhältnis von erworbenen Kompetenzen und Arbeitsmarktoutcomes. Fehlen würde eine Messung des ausbildungsbedingten Kompetenzzuwachses; aber eine Zuordnung der Kompetenzbestände am Ausbildungsende zu institutionellen Ausbildungsformen bliebe möglich. Zudem sind bei dieser Variante umfassende Ergebnisse nach relativ kurzer Zeit präsentierbar.

Stärken und Schwächen des Modells

Mit Blick auf die am Anfang dieses Kapitels noch einmal skizzierten Ziele eines Berufsbildungs-PISA liegen die Stärken eines Paneldesigns mit drei Messzeitpunkten in folgenden Punkten:

- Es ermöglicht vor allem – anders als in den schulbezogenen Vergleichsstudien (PISA, TIMSS) – die Messung von *Kompetenzzuwächsen* in der Zeit und ihre Zuordnung zu institutionellen Rahmenbedingungen und berufspädagogischen Verfahren.
- Durch den Einbezug des dritten Messzeitpunktes etwa drei Jahre nach Abschluss der Ausbildung werden detaillierte Aussagen sowohl zur Nachhaltigkeit von erworbenen Kompetenzen als auch zu ihrer Verwertbarkeit in Arbeit und Arbeitsmarkt und in der persönlichen Biografie (Bildungskarriere, Weiterbildung) möglich.
- Schließlich ist auch ein internationales Benchmarking auf der Ebene von Ausbildungseinrichtungen und Berufsfeldern wegen der Zuschreibbarkeit von Kompetenzzuwächsen zu Ausbildungsformen besser als bei jedem anderen Untersuchungsdesign möglich.

Diesen unbestreitbaren Stärken stehen nicht zu unterschätzende forschungspragmatische Probleme gegenüber:

- Die Durchführung einer Panelstudie ist allein schon wegen der Dauer und der erforderlichen Panelpflege in Bezug auf Kosten und Zeit sehr aufwändig. Allerdings sind auch bei dieser Variante schon nach der ersten Welle wichtige Ergebnisse zu präsentieren, insbesondere bezogen auf die Eingangsvoraussetzungen der Auszubildenden in den unterschiedlichen Kompetenz-Dimensionen.
- Da die Vorteile einer Paneluntersuchung nur realisierbar sind, wenn man sie institutionenbasiert anlegt, tritt das schwer zu lösende Problem der „vertikalen Vergleichbarkeit“, d. h. der Definition von Ausbildungsanfang und -ende und der Festlegung einer Alterskohorte auf. (Dieses Problem hatte das Schul-PISA, das von einer einheitlichen Alterskohorte, „den 15-jährigen“, ausging, nicht.)
- Wegen der Institutionenbasierung des Panels würden Jugendliche, die keine formalisierte Ausbildung absolvieren, aus der Untersuchungsgruppe herausfallen, und es wäre auch schwierig, Ausbildungsabbrecher oder -wechsler weiter zu verfolgen und Kompetenzmessergebnisse bei ihnen institutionell zuzuordnen.

Beide Varianten des Panel-Designs verringern den Erhebungsaufwand um drei Jahre. Je nachdem, ob das Hauptinteresse bei der Messung des Kompetenzzuwachses während der Ausbildung oder bei der Verwertungsseite liegt, fiel die Entscheidung für Variante 1 oder Variante 2.

5.3. DAS FORSCHUNGS-DESIGN EINER QUERSCHNITTSSTUDIE

Unter forschungspragmatischen Gesichtspunkten bildet eine Querschnittsstudie mit (zunächst) einem einmaligen Messzeitpunkt bei einer altershomogenen Gruppe eine ernst zu nehmende Alternative zu dem Längsschnittstudien-Design.

Die Untersuchungsanlage

Der Grundgedanke des Querschnitt-Designs besteht darin, die für ein Berufsbildungs-PISA formulierten Ziele (vgl. Kapitel 1.) in einem kürzeren Zeitraum und mit einem geringeren Ressourcen-Einsatz so weit wie möglich zu erreichen. Ein Querschnitts-Design stellt also einen pragmatischen (aber keinen faulen) Kompromiss zwischen einer weitgesteckten Zielperspektive und der Machbarkeit eines Berufsbildungs-PISA dar. Es mag als ein Indiz für die Ernsthaftigkeit der Alternative angesehen werden, dass beim ersten internationalen Expertenworkshop sich die Teilnehmer einmütig für eine echte Panelstudie ausgesprochen haben, während im zweiten Expertenworkshop ein Teil der Teilnehmer die forschungspragmatischen Schwierigkeiten eines Panel-Designs

höher gewichtete und den Vorschlag einer Querschnittsstudie als ernsthaft zu prüfende Alternative ins Spiel brachte.

Die Untersuchungsanlage sieht einen Messzeitpunkt einer Alterskohorte (24 Jahre zum Untersuchungszeitpunkt) vor, deren bis zum Untersuchungszeitpunkt erreichtes Kompetenzniveau in den unterschiedlichen Bereichen gemessen und deren Schul-, Ausbildungs- und Erwerbsbiografie retrospektiv rekonstruiert wird. Auch hier steht die Kompetenzmessung im Zentrum, und zwar nicht nur im Bereich allgemeiner, sondern vor allem im Bereich fachlicher Kompetenzen.

Das Erhebungsprogramm

Analog zum ersten Modell (Panel-Design) erfolgt die *Kompetenzmessung* auf den drei Ebenen der kognitiven Basiskompetenzen, der berufsübergreifenden und der berufsfachlichen Kompetenzen (vgl. Abschnitt 2.1 und Erhebungsschema „Kompetenzmessung“, S. 142).

Ebenso analog sind die institutionellen und individuellen Lernkontexte retrospektiv zu erheben (vgl. Erhebungsschema auf S. 143). Allerdings wird das Erhebungsprogramm zu den institutionellen Kontexten auf die Dimensionen zu konzentrieren sein, die in einer Individualbefragung sinnvollerweise erfasst werden können.

Die Stichprobe

Grundlage der Stichprobe ist die zum Alter von 24 Jahren ausgeübte Tätigkeit, ohne ergänzenden Bezug auf die absolvierte formale Ausbildung. Auch hierbei wird eine nach den gleichen Berufen wie im ersten Modell geschichtete Stichprobe gezogen (vgl. Stichproben-Schema auf S. 144). Für die Bestimmung der Bruttostichprobe erscheint folgendes gestufte Vorgehen sinnvoll.

- Mittels des ISCO-Codes der ausgewählten Berufe und der nationalen Arbeitskräfteerhebungen können die 24-jährigen in den entsprechenden Berufen und deren Arbeitgeber bzw. Beschäftigungsbetriebe bestimmt werden (im Fall von Deutschland z. B. über die Beschäftigtenstatistik der Bundesagentur für Arbeit).
- Diese Personen bilden die Ausgangsbasis für entweder ein betriebsstratifiziertes Auswahlverfahren oder eine reine Zufallsstichprobe (vgl. Tabelle 5.2).

Tabelle 5.2: Stichprobenplan Querschnittsstudie

	Größe rechnerische Basisstichprobe (einfache Zufallsauswahl, Fehlermarge 5%)	Erhöhung des Umfang durch Anteil Personen ohne Ausbildung (geschätzt 25%)	Erhöhung des Umfangs durch Korrekturfaktor für Clusterung (1.08)	Stichprobenumfang für Querschnittsstudie
KFZ-Mechatroniker	400	+ 100	+ 40	= 540
Industriemechaniker	400	+ 100	+ 40	= 540
Elektroniker (Industrie)	400	+ 100	+ 40	= 540
Bankkaufmann	400	+ 100	+ 40	= 540
Krankenpfleger	400	+ 100	+ 40	= 540
Fachinformatiker	400	+ 100	+ 40	= 540
Alle Tätigkeiten				= 3240

Stärken und Schwächen des Modells

Bezugspunkt für die Betrachtung der Stärken und Schwächen des Modells sind wiederum die Zielsetzungen eines Berufsbildungs-PISA (vgl. Kapitel 1.2.) und die im 4. Kapitel analysierten Probleme der Vergleichbarkeit von Berufsbildungssystemen:

- Das Modell (Geburtskohorten-Design) beseitigt das Problem der Altersvarianz.
- Es ermöglicht den Einbezug von Personen ohne formale Ausbildung, die in den ausgewählten Berufen tätig sind, und bietet somit eine Grundlage auch für ein Benchmarking zwischen formalisierten und informellen Lernprozessen.
- Die Retrospektiverhebung der Ausbildungs- und Erwerbsbiografien erlaubt die Bildung von Berufsbildungstypologien auf der ganzen Breite von formellen und informellen Lernprozessen, was bei dem ausbildungsinstitutionen-basierten Panel-Design nur eingeschränkt möglich ist.
- Das Querschnitts-Design kann schneller Informationen über den Stand der Kompetenzentwicklung einer berufsspezifischen Alterskohorte bereitstellen als eine Panel-Studie.
- Bezogen auf Kosten und Zeit ist das Querschnittsmodell sehr viel weniger aufwändig als das Panel-Modell.

- Es stellt die kostengünstigere Erhebungsmethode dar, was insbesondere in der Phase der Instrumentenentwicklung (nur für einen, nicht für drei Messzeitpunkte) seinen Niederschlag findet. Die Kosten der Einzelerhebung fallen demgegenüber wahrscheinlich höher aus, da die retrospektive Erfassung der Bildungs-, Ausbildungs- und Erwerbsbiografie am besten in einem standardisierten face-to-face-Interview zu leisten ist.

Den hier aufgeführten Vorteilen dieses Modells stehen aber erhebliche inhaltliche Nachteile gegenüber:

- Es ist keine dynamische Kompetenzmessung möglich, die durch Ausbildung und/oder Arbeit verursachte Kompetenzzuwächse zu messen gestattete. Man kann nur Kompetenz*bestände* auf zu erstellende Typen von Ausbildungs- und Arbeitsbiografien beziehen.
- Die gemessenen Kompetenzen können nicht eindeutig institutionellen Settings zugeordnet werden, da kein Kompetenzzuwachs für einzelne Bildungsabschnitte erhoben wird. Dieses Manko dürfte bei den (berufs-) fachlichen Kompetenzen weniger stark zu Buche schlagen als bei den allgemeineren Kompetenzen; allgemeine soziale und metakognitive Kompetenzen könnten in anderen als in Berufsbildungskontexten erworben sein.
- Eine methodische Schwäche liegt auch bei der Rekonstruktion der Bildungs- und Erwerbsbiografie, da für längere Zeiträume das individuelle Gedächtnis keine sehr zuverlässige Auskunftsource ist.
- Mögliche Berufswechsel aus dem spezifischen Ausbildungsberuf sind nicht erfassbar, nur Wechsel *in* den ausgewählten Beruf.³⁰

30 Gegenüber dem Forschungsbericht haben wir die Kostenkalkulation hier herausgenommen, da sie als Modell-Kalkulation zu vorläufig und (notwendigerweise) zu sehr auf die Bedingungen der beiden Institute bezogen war, die die Machbarkeitsstudie erstellt haben.

6. ZUM WEITEREN VORGEHEN

Die vorgelegte Machbarkeitsstudie für ein Berufsbildungs-PISA zeigt die Möglichkeiten und erforderlichen inhaltlichen Arbeitsschritte auf, die für die Realisierung eines international vergleichenden „large scale assessment“ zu bewältigen sind.

- Es sind die in den Kapiteln 2 bis 4 entwickelten Untersuchungsperspektiven und Indikatoren zu operationalisieren und in Untersuchungsinstrumente umzusetzen. Hierbei erfordert die Entwicklung von Tests zur Messung von Kompetenzen einen besonders hohen Aufwand, zumal dann, wenn die berufsfachlichen Kompetenzen das ihnen in dieser Machbarkeitsstudie zugesprochene Gewicht haben sollen.
- Es ist eine Entscheidung über die Untersuchungsanlage (ob Panel-Längsschnittuntersuchung oder Querschnittsstudie) zu treffen.
- Es ist eine internationale Arbeitsplattform auszubauen, von der aus ein „Programme for an International Student Assessment on Vocational Education and Training“ (PISA-VET) durchgeführt werden kann.
- Die Kosten für die hier vorgelegte Machbarkeitsstudie, einschließlich der in sie eingeschlossenen internationalen Kooperation, sind von der deutschen Regierung (Bundesministerium für Wirtschaft und Arbeit) getragen worden. Alle internationalen Experten haben auf den Workshops die Initiative des Bundesministeriums für Wirtschaft und Arbeit (BMWA) ausdrücklich begrüßt und gewürdigt. Zugleich war allen bewusst, dass eine Weiterführung dieser Initiative einer internationalen Finanzierungs- und Organisationsbasis bedarf. Interessenbekundungen liegen von offiziellen Stellen in mehreren Ländern vor.

Das durch die internationalen Workshops geschaffene Forschungsnetzwerk ist für den Aufbau einer internationalen Arbeitsplattform weiterhin zu nutzen. Alle Teilnehmer haben in der Abschlusssitzung des zweiten Workshops (30. April 2005) ihr nachdrückliches Interesse an einer Weiterarbeit bekundet.

Nach den Diskussionen auf dem 2. internationalen Workshop zeichnen sich zwei mögliche Wege der Weiterarbeit ab, die man als „kleine“ und als „große“ Lösung bezeichnen kann.

Kleine Lösung

Einzelne Länder schließen sich zu einer jeweils kleinen Gruppe zusammen und entwickeln auf dieser Basis ein Berufsbildungs-PISA. So könnten z. B. die deutschsprachigen Länder mit starken dualen Berufsbildungssystemen eine

solche Gruppe bilden (Schweiz, Österreich, Deutschland), eine andere die skandinavischen Länder mit Großbritannien. Der Vorzug einer solchen „kleinen“ Lösung läge in der Verringerung des Koordinationsaufwandes und darin, dass ein schnellerer Start möglich wäre. Auf der anderen Seite ergibt sich als Nachteil, dass die Breite institutioneller Formen der Berufsbildung für ein „benchmarking“ eingeschränkt wäre. Gleichwohl bietet diese Lösung gute Chancen für den Vergleich unterschiedlicher Formen – selbst in der kleinen Zahl von Ländern, da in ihnen unterschiedliche Ausbildungssysteme bestehen. Auch in diesem Fall wäre zu prüfen, wieweit eine gemeinsame Methoden-Entwicklung möglich wäre und Forschungs-Designs aufeinander bezogen werden könnten, wobei alternative Messmethoden nicht ausgeschlossen werden sollten.

Große Lösung

Auf der Basis einer internationalen Organisation (EU; OECD oder einer anderen) wird eine große Lösung mit vielen Ländern angestrebt. Diese hätte den Vorzug, dass ein breiter Vergleich und eine mit politischer Autorität ausgestattete Forschungsanstrengung durchgeführt werden könnte. Sie hätte möglicherweise den Nachteil, dass die erforderliche politische Koordinierung, vor allem aber die Konsensbildung über Standards und Messmethoden für Kompetenzen sehr lange dauern würde. Dies könnte sowohl den wissenschaftlichen Elan, der gegenwärtig zu aktivieren ist, als auch die politische Aufmerksamkeit nach dem „PISA-Schock“ erlahmen lassen.

LITERATURVERZEICHNIS

- Achtenhagen, F. (1990). Vorwort. In Deutsche Forschungsgemeinschaft, *Berufsbildungsforschung an den Hochschulen der Bundesrepublik Deutschland* (S. VI-VII). Weinheim, Basel: VCH.
- Achtenhagen, F. (1996). Entwicklung und Evaluation ökonomischer Kompetenz mit Hilfe handlungsorientierter Verfahren – am Beispiel der Ausbildung zum Industriekaufmann/ zur Industriekauffrau (Skizze eines Forschungsprojekts). In W. Seyd & R. Witt (Eds.), *Situation – Handlung – Persönlichkeit. Kategorien wirtschaftspädagogischen Denkens. Festschrift für Lothar Reetz* (S. 137–147). Hamburg: Feldhaus.
- Achtenhagen, F. (1998). General versus Vocational Education – Demarcation and Integration. In W. J. Nijhof & J. N. Streumer (Eds.), *Key Qualifications in Work and Education* (pp. 133–143). Dordrecht, Boston, London: Kluwer.
- Achtenhagen, F. (2004). Curriculum Development as Modeling of Complex Reality. In N. M. Seel & S. Dijkstra (Eds.), *Curriculum, Plans, and Processes in Instructional Design* (pp. 193–210). Mahwah, N. J., London: Erlbaum.
- Achtenhagen, F., & John, E.G. (Eds.). (1992). *Mehrdimensionale Lehr-Lern-Arrangements – Innovationen in der kaufmännischen Aus- und Weiterbildung*. Wiesbaden: Gabler.
- Achtenhagen, F. in Cooperation with J. W. Nijhof & D. Raffe (1995). *Feasibility Study: Research Scope for Vocational Education in the Framework of COST Social Sciences*. European Commission: Directorate-General Science, Research and Development, COST Technical Committee Social Sciences, Vol. 3. Brussels, Luxembourg: ECSC-EC-EAEC.
- Achtenhagen, F., & Lempert, W. (Hrsg.) (2000). *Lebenslanges Lernen im Beruf – Seine Grundlegung im Kindes- und Jugendalter. Vol. 1: Das Forschungs- und Reformprogramm*. Opladen: Leske + Budrich.
- Achtenhagen, F., & Grubb, W.N. (2001). Vocational and Occupational Education: Pedagogical Complexity, Institutional Diversity. In V. Richardson (Ed.), *Handbook of Research on Teaching, Fourth Edition* (pp. 604–639). Washington, DC: American Educational Research Association.
- Achtenhagen, F., & Weber, S. (2003). „Authentizität“ in der Gestaltung beruflicher Lernumgebungen. In A. Bredow, R. Dobischat & J. Rothmann (Hrsg.), *Berufs- und Wirtschaftspädagogik von A-Z* (S. 185–199). Baltmannsweiler: Schneider Verlag Hohengehren.
- Achtenhagen, F., Preiß, P., & Weber, S. (2005). *Lerntheoretische Grundlagen für die Entwicklung von Lehr-/Lernarrangements*. Oldenburg: Universität Oldenburg.
- Aebli, H. (1987). *Grundlagen des Lehrens: Eine allgemeine Didaktik auf psychologischer Grundlage*. Stuttgart: Klett-Cotta.
- Altiner, A., Wilm, S., & Sielk, M. (2002). Allgemeinmedizinische Inhalte im Medizinstudium. Implementation einer Klausur. *Zeitschrift für Allgemeine Medizin*, S. 425–430.
- Anderson, L. W., & Krathwohl, D. R. (with Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et al.) (Eds.) (2001). *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Andersson, R., & Olsson, A.-K. (1999). *Fields of Education and Training*. Luxembourg: Eurostat.
- Arbeitsgemeinschaft Betriebliche Weiterbildungsforschung (2005). *Kompetenzmessung im Unternehmen: Lernkultur- und Kompetenzanalysen im betrieblichen Umfeld*. Band 18. Münster et al.: Waxmann.
- Baethge, M., Buss, K.P., & Lanfer, C. (2003). *Konzeptionelle Grundlagen für einen Nationalen Bildungsbericht – Berufliche Bildung und Weiterbildung/Lebenslanges Lernen*. Bonn: BMBF

- Baethge, M., & Baethge-Kinsky, V. (2004). *Der ungleiche Kampf um das lebenslange Lernen*. Münster, Berlin: Waxmann.
- Baethge, M., & Baethge-Kinsky, V. (2006, im Druck). Ökonomie, Technik, Organisation: Zur Entwicklung von Qualifikationsstruktur und qualitativem Arbeitsvermögen. In R. Arnold, & A. Lipsmeier (Hrsg.). *Handbuch der Berufsbildung*. 2. Aufl. Opladen: Leske+Budrich.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J., Weiß, M. (1999): Soziale Bedingungen von Schulleistungen – Zur Erfassung von Kontextmerkmalen durch Schüler-, Schul- und Elternbefragungen. <http://www.mpib-berlin.mpg.de/pisa/Kontextmerkmale.pdf>
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., & Weiß, M. (2001). PISA 2000: *Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Baumert, J., & Schümer, G. (2002). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzwettbewerb im internationalen Vergleich. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele et al. (Hrsg.), *PISA 2000: Die Länder der Bundesrepublik Deutschland im Vergleich* (S. 159–202). Opladen: Leske + Budrich.
- Baumert, J., Artelt, C., Carstensen, C. H., Sibbers, H., & Stanat, P. (2002). Untersuchungsgegenstand, Fragestellungen und technische Grundlagen der Studie. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele et al. (Hrsg.), *PISA 2000: Die Länder der Bundesrepublik Deutschland im Vergleich* (S. 11–38). Opladen: Leske + Budrich.
- Beck, K. (1995). Die Äquivalenz von Ausbildungsabschlüssen in Europa als Voraussetzung transnationaler beruflicher Mobilität? Ein erziehungswissenschaftlicher Ansatz zur Lösung des Vergleichsproblems. In K. A. Schachtschneider (Hrsg.), *Wirtschaft, Gesellschaft und Staat im Umbruch. Festschrift der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Friedrich-Alexander-Universität Erlangen-Nürnberg 75 Jahre nach Errichtung der Handelshochschule Nürnberg* (S. 1–16). Berlin: Duncker & Humblot.
- Bendorf, M. (2005). Förderung von Metakognition und Lernstrategien am Fachgymnasium Wirtschaft. In P. Gonon, F. Klauser & R. Huisinga (Hrsg.), *Kompetenz, Kognition und neue Konzepte der beruflichen Bildung* (S. 203–218). Wiesbaden: Verlag für Sozialwissenschaften.
- BIBB (1999). *Komplexe Prüfungsaufgaben für Bürokaufleute: KoPrA- ein Leitfaden zur Erstellung von Prüfungsaufgaben für die Kammerprüfungen*. Bielefeld: Bertelsmann.
- BIBB (2005). *Ausbilden mit Lernaufgaben*. Band 1–3. Konstanz: Christiani.
- Bjornavold, J. (2000). *Making learning visible: Identification, Assessment and Recognition of Non-Formal Learning in Europe*. Luxembourg: Official Publication of the European Communities.
- BLK (2002). *Vergleichende Bildungsstatistik – Sachstand und Vorschläge zur Verbesserung*, Heft 103. Bonn.
- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten, Soziale Welt*, Bd. 2, S. 183–198.
- Bransford, J.D., Brown, A.L., & Cocking, R.R. (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington, D.C.: National Academy Press.
- Büchel, F., & Neubäumer, R. (2001). Ausbildungsinadäquate Beschäftigung als Folge branchenspezifischer Ausbildungsstrategien. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 34, S. 269–285.
- CEDEFOP (1998). *Transparenz beruflicher Befähigungsnachweise in Europa – Stand und Entwicklungsperspektiven*. Thessaloniki.
- CEDEFOP (1999ff.), *Vocational Education and Training in (diverse europäische Länder)*. Thessaloniki.

- CEDEFOP (2003). *Terminology of vocational training policy: A multilingual glossary for an enlarged Europe*.
http://www.europass.cedefop.eu.int/img/dynamic/c313/cv-1_en_US_glossary_4030_6k.pdf.
- COM/DELSA/EDU (2004). *Programme for the International Assessment of Adult Competencies – Draft Strategy Paper*. <http://www.oecd.org/dataoecd/3/3/34463133.pdf>
- Corti, S. (2000). Forschungserfordernisse der Berufsbildungspolitik der Europäischen Union. In CEDEFOP/DIPF (Hrsg.), *Vergleichende Berufsbildungsforschung in der Europäischen Union*, (S. 28–29). Frankfurt am Main, Thessaloniki.
<http://www2.trainingvillage.gr/download/publication/reference/BonnReaderDE.pdf>
- Deci, E.L., & Ryan, M.R. (1985). *Intrinsic motivation and self-determination in human behaviour*. New York, et al.: Plenum Press.
- Descy, P., & Tessaring, M. (2002). *Kompetent für die Zukunft – Ausbildung und Lernen in Europa. Zweiter Bericht zur Berufsbildungsforschung in Europa: Synthesebericht*. Reihe Cedefop reference series, Bd. 5. Luxemburg.
- Deutsche Forschungsgemeinschaft (1990). *Berufsbildungsforschung an den Hochschulen der Bundesrepublik Deutschland. Denkschrift*. Weinheim Basel: VCH.
- Dolch, J. (1965). *Lehrplan des Abendlandes: Zweieinhalb Jahrtausende seiner Geschichte*. 2. Aufl. Ratingen: Henn.
- Ellström, P.-E. (1997). The many meanings of occupational competence and qualification. In W.J. Nijhof & J.N. Streumer (Eds.), *Key qualifications in work and education* (pp. 39-50). Dordrecht, Boston, London: Kluwer.
- Erpenbeck, J., & von Rosenstiel, L. (Hrsg.) (2003a). *Handbuch Kompetenzmessung*. Stuttgart: Schäffer-Poeschel.
- Erpenbeck, J., & von Rosenstiel, L. (2003b). Einführung. In J. Erpenbeck & L. von Rosenstiel (Hrsg.), *Handbuch Kompetenzmessung* (S. IX-XL). Stuttgart: Schäffer-Poeschel.
- Euler, D. (Hrsg.) (2004). *Handbuch der Lernortkooperation, Band 1*. Bielefeld: Bertelsmann.
- Europäische Kommission (1997). *Schlüsselzahlen zur Berufsbildung in der Europäischen Union*. Luxemburg.
- Europäische Kommission (2002). *Beschäftigung in Europa 2002 – Jüngste Tendenzen und Ausblick in die Zukunft*. Brüssel:
http://europa.eu.int/comm/employment_social/news/2002/sep/employment_in_europe2002_de.pdf
- EUR-Lex (1985). 85/368/EWG: *Entscheidung des Rates vom 16. Juli 1985 über die Entsprechungen der beruflichen Befähigungsnachweise zwischen Mitgliedstaaten der Europäischen Gemeinschaften*.
<http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=CELEX:31985D0368:DE:HTML>
- Europäische Sozialagenda (2001). Rat der Europäischen Union (Hrsg.). Luxemburg: Amt für amtliche Veröffentlichungen der Europäischen Gemeinschaften.
- Europäische Union (2001). *Übergang vom Bildungswesen ins Erwerbsleben – Schlüsselzahlen zur Berufsbildung in Europa*. Eurostat, Cedefop.
- Eurostat/CEDEFOP (2002). *Overview of National VET-Systems*. Thessaloniki.
http://www.trainingvillage.gr/etv/Information_resources/NationalVet/
- Fischer, G., & Siebern-Thomas (2004). Arbeit und Arbeitsmärkte im Zeitalter wissensbasierter Wirtschaft und Gesellschaft. Gibt es noch einen nationalen Arbeitsmarkt? In E. Hönekopp, R. Jungnickel & T. Straubhaar (Hrsg.), *Internationalisierung der Arbeitsmärkte, Beiträge zur Arbeitsmarkt- und Berufsforschung*, Nr. 282 (S. 35–60). Nürnberg: Bundesagentur für Arbeit.
- Fisher, C.W., & Berliner, D.C. (Eds.) (1985). *Perspectives on Instructional Time*. New York: Longman.

- Fulst-Bleil, S., & Ebner, H.G. (2005). Duale Berufsausbildung im internationalen Vergleich. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 101, S. 176–185.
- Franke, G., & Kleinschmitt, M. (1987). *Der Lernort Arbeitsplatz*. Berlin: Beuth.
- Ganzeboom, H.B.G., & Treiman, D.J. (2003). Three Internationally Standardised Measures for Comparative Research on Occupational Status. In J.H.P. Hoffmeyer-Zlotnik & Ch.Wolf (Eds.), *Advances in Cross-National Comparison* (pp. 159–193). New York: Kluwer.
- Gomez, P., & Probst, G. (1997). *Vernetztes Denken im Management*. Bern: Schweizerische Volksbank.
- Green, F., McIntosh, S., & Vignoles, A. (2004). *Scoping Study on International Comparisons of Non-Certified learning*.
http://www.dti.gov.uk/training_development/Non_certified_learning_final.pdf.
- Greinert, W.-D. (1995). *Das duale System der Berufsausbildung in der Bundesrepublik Deutschland*. Stuttgart: Holland + Josenhans.
- Grünewald, U., Moraal, D., & Schönfeld G. (2003). Betriebliche Weiterbildung in Deutschland und Europa. Bielefeld : Bertelsmann.
- Hannan, D., Raffé, D., & Smyth, E. (1997). Cross-national research on school to work transitions: an analytical framework. In P. Werquin, R. Breen & J. Planas (Eds.), *Youth transitions in Europe: theories and evidence* (p. 409–422). Marseille: Cereq.
- Hanushek, E.A. (1989). Expenditures, efficiency, and equity in education: the Federal Government's role. *The American Economic Review*, 79, 41–51.
- Hanushek, E.A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, 19, 141–164.
- Hasselhorn, M. (2000). Lebenslanges Lernen aus der Sicht der Metakognitionsforschung. In F. Achtenhagen & W. Lempert (Hrsg.), *Lebenslanges Lernen im Beruf – seine Grundlagen im Kindes- und Jugendalter (III) Psychologische Theorie, Empirie und Therapie* (S. 41–53). Opladen: Leske +Budrich.
- Heid, H., & Harteis, C. (2005). *Verwertbarkeit – Ein Qualitätskriterium (erziehungs-)wissenschaftlichen Wissens*. Wiesbaden: Verlag für Sozialwissenschaften.
- Heikkinen, A. (2000). Ist eine stärkere Berücksichtigung der historischen und kulturellen Dimension in der Vergleichenden Berufsbildungsforschung erforderlich? In CEDEFOP/DIPF (Hrsg.), *Vergleichende Berufsbildungsforschung in Europa. Ansätze, Politikbezüge und Innovationstransfer* (S. 118–134). Thessaloniki.
- ILO (2005): *International Standard Classification of Occupations*.
<http://www.ilo.org/public/english/bureau/stat/class/isco.htm>
- International Technology Education Association (2002). *Standards for Technological Literacy Content for the Study of Technology, Second Edition*. <http://www.iteawww.org>
- Jeanneret, P. R. (1969). *A study of the job dimensions of "worker-oriented" job variables and of their attribute profiles*. Dissertation Abstracts International.
- Jeanneret, P. R., Borman, W. C., Kubisiak, U. C., & Hanson, M. A. (2002). Generalized Work Activities. In N. G. Peterson, M. D. Mumford, W. C. Borman, P. R. Jeanneret & E. A. Fleishman (Eds.), *An Occupational Information System For The 21st Century: The Development of O*NET* (pp. 105–125). Washington, D. C.: American Psychological Association.
- Kanning, U.P. (2003). *Diagnostik sozialer Kompetenzen*. Göttingen: Hogrefe.
- Kearns, P. (2001). *Generic skills for the new economy*. Australian National Training Authority. <http://www.ncver.edu.au/research/proj/nr0024.pdf>.
- Neubrand, M., Klieme, E., Lüdtke, O., & Neubrand, J. (2002). Kompetenzstufen und Schwierigkeitsmodelle für den PISA-Test zur mathematischen Grundbildung. *Unterrichtswissenschaft*, 30, 100–119.

- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E., & Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards – Eine Expertise*. Bonn: Bundesministerium für Bildung und Forschung.
- KMK (1996/2000). *Handreichungen für die Erarbeitung von Rahmenlehrplänen der Kultusministerkonferenz (KMK) für den berufsbezogenen Unterricht in der Berufsschule und ihre Abstimmung mit Ausbildungsordnungen des Bundes für anerkannte Ausbildungsberufe* <http://www.kmk.org/doc/publ/handreich.pdf>.
- Koch, R., & Reuling, J. (1998). Öffentliche Qualitätssteuerung der Berufsausbildung in Deutschland, Frankreich und England. In: CEDEFOP (Hrsg.), *Berufsbildung, Bd. 15*. Thessaloniki.
- Kogan, I. & Müller, W. (Hrsg.) (2003). *School-to-work Transitions in Europe: Analyses of the EU LFS 2000 Ad hoc Module*. Mannheim: http://www.mzes.uni-mannheim.de/publications/books/school_to_work_2ed.pdf.
- Kommission der Europäischen Gemeinschaften (2005): *Auf dem Weg zu einem Europäischen Qualifikationsrahmen für lebenslanges Lernen*. Arbeitsunterlagen der Kommissionsstellen. http://ec.europa.eu/education/policies/2010/doc/consultation_eqf_de.pdf
- Kommission der Europäischen Gemeinschaften (2005). *Auf dem Weg zu einem Europäischen Qualifikationsrahmen für Lebenslanges Lernen*. Brüssel
- Konsortium Bildungsberichterstattung (2006): *Bildung in Deutschland*. Bielefeld.
- Kuhl, J. (1984). *Motivation, Konflikt und Handlungskontrolle*. Heidelberg: Springer
- Kuhl, J. (1998). Wille und Persönlichkeit. *Psychologische Rundschau*, 49, 61–77.
- Kunter, M. et al. (2002). *PISA 2000 – Dokumentation der Erhebungsinstrumente*. Max-Planck-Institut für Bildungsforschung, Berlin. <http://www.kmk.org/schul/pisa/Datensaetze/Gesamt-Druckversion.pdf>
- Kuratorium der deutschen Wirtschaft (2005). *Berufliche Bildung für Europa: Europäischer Qualifikationsrahmen (EQF) und Leistungspunktesystem (ECVET)*. http://www.kwb-berufsbildung.de/pdf/2005_Positionspapier_EQF_ECVET.pdf
- Lamb, S., Davies, M., Polesel, J. & Teese, R. (2003). *International indicators for vocational education and training*. Leabrook.
- Lang, B., & Pointinger, M. (2002). Das selbst gesteuerte Lernen österreichischer Jugendlicher. In C. Reiter & G. Haider (Hrsg.), *PISA 2000 – Lernen für das Leben. Österreichische Perspektiven des internationalen Vergleichs* (S. 101–104). Innsbruck: Studienverlag Innsbruck.
- Lang-von Wins, T. (2004). *Der Unternehmer: arbeits- und organisationspsychologische Grundlagen*. Berlin et al.: Springer.
- Lasonen, J., & Manning, S. (2001). How to improve the standing of vocational compared to general education. A collaborative investigation of strategies and qualifications across Europe. In P. Descy & M. Tessaring (Eds.), *Training in Europe. Second report of vocational training research in Europe* (p. 115–167), Vol. I. Thessaloniki.
- Lassnigg, L. (2000). *Steuerung, Vernetzung und Professionalisierung in der Berufsausbildung*. Wien: Institut für Höhere Studien (IHS). http://www.lebenslangeslernen.at/downloads/IHS_SteuerungBerufsausbildung_1100.pdf
- Lassnigg, L. (2003). *Indicators for Quality in VET – European Experience*. Vienna: Institute for Advanced Studies.
- Lempert, W. (1979). Zur theoretischen und empirischen Analyse von Beziehungen zwischen Arbeiten und Lernen. Grundprobleme und Lösungsstrategien. In P. Groskurth (Hrsg.), *Arbeit und Persönlichkeit. Berufliche Sozialisation in der arbeitsteiligen Gesellschaft* (S. 87–111). Reinbek bei Hamburg.

- Leney, T., & Deluca, M. (2000). Einheit und Vielfalt der Systeme beruflicher Bildung der Mitgliedsstaaten der Europäischen Union – Konvergenz- und Divergenztendenzen im Bildungs- und Berufsbildungswesen der Mitgliedstaaten der Europäischen Union. In CEDEFOP/DIPF (Hrsg.), *Vergleichende Berufsbildungsforschung in Europa. Ansätze, Politikbezüge und Innovationstransfer* (S. 103–118). Thessaloniki.
- Leschinsky, A., & Mayer, K.U. (1999). *The Comprehensive School Experiment Revisited*. Frankfurt: Lang.
- Lisbon-to-Copenhagen-to-Maastrich Consortium Partners (2004). *Achieving the Lisbon goal: The contribution of VET - Final report to the European Commission*.
http://europa.eu.int/comm/education/policies/2010/studies/maastricht_en.pdf
- Mainguet, C. (1999). *Der Übergang vom Bildungssystem ins Erwerbsleben – Auswertung der nationalen statistischen Daten*. Luxembourg: CEDEFOP.
- Mertens, D. (1974): Schlüsselqualifikationen: *Thesen zur Schulung für eine moderne Gesellschaft. Mitteilungen aus Arbeitsmarkt- und Berufsforschung* 7, S. 36–43.
- Münch, J. (1999). *Qualifikation als Standortfaktor*. Hochheim am Main: Neres.
- Murray, T.S., Clermont, Y., & Binkley, M. (2004). *The International Adult Literacy and Life Skills Survey: Aspects of Design, Development and Validation*. Unpublished paper.
- National Center for Education Statistics (NCES) (1998). *Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey*. Washington, D.C.: NCES.
- National Council on Economic Education (1997). *Voluntary National Content Standards in Economics*. <http://www.ncee.net/ea/program.php?pid=19>
- National Framework of Qualifications (2003). *Determinations for the Outline National Framework of Qualifications*. <http://www.nqai.ie/determinations.pdf>
- National Training Information Service (2005). www.ntis.gov.au
- NCVER (2003). *International indicators for vocational education and training – An Australian perspective*. Adelaide.
- NCVER (2004). *International benchmarking of vocational education and training*. Adelaide.
- Nieuwenhuis, L. (2004). *The learning potential of the workplace*.
http://edu.fss.uu.nl/ord/fullpapers/Nieuwenhuis-L_1%204_2.DOC
- Nijhof, W.J. (1998). Qualifying for the future. W.J. Nijhof & J.N. Streumer (Eds.), *Key qualifications in work and education* (pp. 19–38). Dordrecht: Kluwer.
- Norris, N. (1991). The trouble with competence. *Cambridge Journal of Education*, Vol. 21(3), p. 1–11.
- Oates, T. (2004). The Role of outcome-based National Qualifications in the Development of an Effective Vocational Education and Training System: The Case of England and Wales. *Policy Futures in Education*, Volume 2 (1).
- OECD (2001). *Lernen für das Leben. Erste Ergebnisse der internationalen Schulleistungsstudie PISA 2000*, Paris.
[http://www.pisa.oecd.org/Docs/Download/PISA_2001\(deutsch\).pdf](http://www.pisa.oecd.org/Docs/Download/PISA_2001(deutsch).pdf).
- OECD (2003): *Education at a glance*, Paris.
- OECD (2004). *PIAAC Draft Strategy Paper – Policy Objectives, Strategic Options and Cost Implications*. <http://www.oecd.org/dataoecd/3/3/34463133.pdf>
- Office of Apprenticeship Training, Employer and Labor Services (OATELS) (2004). *The national Registered Apprenticeship System – Programmes and Apprentices*. Fiscal Year 2003.
- Onstenk, J.H.A.M. (2003). *Work based learning in continuing vocational and professional education, Paper for WACE – First Draft*.
http://www.wacerotterdam2003.nl/documents/final_papers_abstracts/140.doc

- Pellegrino, J.W.; Chudowsky, N., & Glaser, R. (2001). *Knowing what Students Know: The Science and Design of Educational Assessment*. Washington, D.C.: National Academy Press.
- Pintrich, P.R., Smith, J.A., Garcia, T., & McKeachie, W. (1991). A manual for the use of motivated strategies for learning questionnaire (MSLQ). University of Michigan: National Center for Research to Improve Postsecondary Teaching and Learning.
- Preiß, P. (1992). Komplexität im Betriebswirtschaftslehre-Anfangsunterricht. In Frank Achtenhagen & Ernst G. John (Hrsg.), *Mehrdimensionale Lehr-Lern-Arrangements – Innovationen in der kaufmännischen Aus- und Weiterbildung* (S. 58–78). Wiesbaden: Gabler.
- Preiß, P. (1999). Didaktik des wirtschaftsinstrumentellen Rechnungswesens. München: Oldenbourg.
- Preiß, P. (2004). *Strukturen der Beschreibung von Kompetenzen in England, Frankreich und Deutschland*. Vortrag auf der Herbsttagung der Sektion Berufs- und Wirtschaftspädagogik der Deutschen Gesellschaft für Erziehungswissenschaft, Siegen.
- Prenzel, M., Kramer, K., & Drechsel, B. (2001). Selbstbestimmt motiviertes und interessiertes Lernen in der kaufmännischen Erstausbildung – Ergebnisse eines Forschungsprojekts. In K. Beck & V. Krumm (Hrsg.), *Lehren und Lernen in der beruflichen Erstausbildung. Grundlagen einer modernen kaufmännischen Berufsqualifizierung* (S. 37–63). Opladen: Leske + Budrich.
- Rauner, F., & Spöttl, G. (2002). *Aufgabenanalyse für die Neuordnung der Berufe im Kfz-Sektor. Abschlussbericht: Wissenschaftliche Begleitung zur Neuordnung der fahrzeugtechnischen Berufe*.
<http://www.itb.uni-bremen.de/downloads/Sonstiges/Kfz-Neuordnung-Bericht4.pdf>.
- Rammstedt, B., Koch, K., Borg, I., & Reitz, T. (2004): *Entwicklung und Validierung einer Kurzskaala für die Messung der Big-Five-Persönlichkeitsdimensionen in Umfragen*.
http://www.gesis.org/Publikationen/Zeitschriften/ZUMA_Nachrichten/documents/pdfs/55/05_rammstedt.pdf.
- Reetz, L. (1999). Zum Zusammenhang von Schlüsselqualifikationen – Kompetenzen – Bildung. In T. Tramm, D. Sembill, F. Klauser & E. G. John (Hrsg.), *Professionalisierung kaufmännischer Berufsbildung. Beiträge zur Öffnung der Wirtschaftspädagogik für die Anforderungen des 21. Jahrhunderts. Festschrift zum 60. Geburtstag von Frank Achtenhagen* (S. 32–51). Frankfurt a. M.: Lang.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. 2. Aufl. Bern: Huber.
- Rost, J., Carstensen, C.H., & Davier, M. von (1999). Sind die Big Five Rasch-skalierbar? Eine Reanalyse der NFO-FFI-Normungsdaten. *Diagnostica*, 45, S. 119–127.
- Roth, H. (1971). *Pädagogische Anthropologie, Bd. 2. Entwicklung und Erziehung. Grundlagen einer Entwicklungspädagogik*. Hannover: Schroedel.
- Rychen, D.S., & Salganik, L.H. (2003). A holistic model of competence. In D.S. Rychen & L.H. Salganik (Eds.), *Key Competencies for a Successful Life and a Well-Functioning Society* (pp. 41–62). Cambridge, Göttingen: Hogrefe & Huber.
- Sandberg, J. (2000). Competence – the basis for a smart workforce. In R. Gerber and C. Lankshear (Eds.), *Training for a Smart Workforce*, pp. 47–72. London: Routledge.
- Scottish Credit and Qualifications Framework (2003). *An Introduction to the Scottish Credit and Qualification Framework*, 2nd edition. Edinburgh.
http://www.sqa.org.uk/files_ccc/IntroductiontoSCQF-2ndEdition.pdf
- Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, 1, 61–80.
- Scheerens, J. & Bosker, R.J. (1997). *The Foundations of Educational Effectiveness*. Oxford: Pergamon Press.
- Scheerens, J., (2004). *Review of school and instructional effectiveness research – Contribution to chapter 3 of the 2004 EFA Global Monitoring Report*.

- http://portal.unesco.org/education/en/file_download.php/23880a1c51a21961999fe362fcefde1dScheerens.pdf
- Scholz, S. (2003). *PISA 2000: Die Studie im Fokus der Statistik*. http://www.sozialforschung.uni-bremen.de/dip_scholz.pdf.
- Schnell, R., Hill, P., & Esser, E. (1989). *Methoden der empirischen Sozialforschung*. München, Oldenbuorg.
- Secretary's Commission on Achieving Necessary Skills (1991). *What Work Requires of Schools. A SCANS Report for America 2000: Executive Summary*. Washington, DC: The Secretary's Commission on Achieving Necessary Skills, U.S. Department of Labour.
- Sellin, B. (1999). *Europäische Trends in der Berufs- und Qualifikationsentwicklung, Band III*. Thessaloniki: CEDEFOP.
- Sellin, B. (2005). *Europäischer Qualifikationsrahmen (EQF) – ein gemeinsames Bezugssystem für Bildung und Lernen in Europa*. http://www.bwpat.de/ausgabe8/sellin_bwpat8.shtml
- Shavelson, R.J., & Ruiz-Primo, M. (2005). Windows into the mind. *International Journal of Higher Education and Educational Planning*, 48(1), 260–278.
- Smith, E., & Comyn, P. (2003). *The development of employability skills in novice workers*. Australian National Training Authority. <http://www.ncver.edu.au/research/proj/nr2005.pdf?PHPSESSID=0f303f7a365c7efd09da16d87747e0bd>.
- Smits, M.H.S.B., Sluijsmans, D.M.A., & Jochems, W.M.G. (2005). *Performance Assessment in Nursing Education: A Case Study*. Open University of the Netherlands. Paper zum Vortrag auf der EARLI Konferenz, Zypern August 2005.
- Solga, H. (2005). *Ohne Abschluss in die Bildungs-Gesellschaft*. Opladen: Barbara Budrich.
- Spencer, L. M., & Spencer, S. M. (1993). *Competence at Work*. New York et al.: John Wiley & Sons.
- State of Oregon, Department of administrative Services, Human Resource Service Division (2005): <http://www.hr.das.state.or.us/hrsd/class/6255.HTM>.
- Stasz, C., Ramsey, K., Eden, R., Melamid, E., & Kaganoff, T. (1996). *Workplace Skills in Practice – Case Studies of Technical Work*. Santa Monica: National Centre for Research in Vocational Education, RAND.
- Straka, G.A. (2003): *Zertifizierung non-formell und informell erworbener beruflicher Kompetenzen*. Münster: Waxmann.
- Stoof, A. (2005). *Tools for identification and description of competencies*. Herleen: Open Universiteit Nederland. <http://www.ou.nl/otecresearch/publications/Angela%20Stoof/Proefschrift%20Angela%20Stoof.pdf>
- Tramm, T. (1992). *Konzeption und theoretische Grundlagen einer evaluativ-konstruktiven Curriculumstrategie – Entwurf eines Forschungsprogramms unter der Perspektive des Lernhandelns*. Göttingen: Seminar für Wirtschaftspädagogik der Georg-August-Universität Göttingen, Berichte: Band 17.
- Treiber, B. (1982). Lehr- und Lernzeiten im Unterricht. In B. Treiber & F.E. Weinert (Hrsg.), *Lehr-Lern-Forschung* (S. 12–36). München: Urban&Schwarzenberg.
- Toolsema, B. (2003). *Werken met competenties. Naar een instrument voor de identificatie van competenties*. Enschede: Universiteit Twente.
- Tuomi, I. (1999). *Corporate Learning*. Helsinki: Metaxis.
- U.S. Department of Education; U.S. Department of Labor (1996). *Occupational Skill Standards Project*. Washington, D.C.
- UNESCO (2004). *Global education digest 2004*. Montreal: UNESCO Institute for Statistics.
- UNESCO (2005). *International Standard Classification of Education*. http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm

- US Bureau of Labor Statistics (2005). *High Growth Industry Profile*. <http://www.careervoyages.gov/>
- Van den Berghe, W. (1997). *Indicators in Perspective – The use of indicators in vocational education and training*. Luxembourg: CEDEFOP.
- Verein Deutscher Ingenieure e. V. (2004). *Bildungsstandards im Fach Technik für den mittleren Schulabschluss. Arbeitspapier zum Politikdialog „Kompetent durch Technikunterricht“*. Tagung in Berlin.
<http://www.vdi-jutec.de/medienarchiv/ablage/original/tagungsunterlage.pdf>
- Weber, S. (2005). *Intercultural learning as identity negotiation*. Frankfurt am Main: Lang.
- Weber, S., Achtenhagen, F., Bendorf, M., & Getsch, U. (2003). Überlegungen zur Gestaltung einer wirksamen Kooperation von Institutionen und Lernorten der beruflichen Bildung. In F. Achtenhagen (Hrsg.), *Neue Wirtschaftspädagogische Forschungs- und Entwicklungsarbeiten* (S. 298–331). Göttingen: Seminars für Wirtschaftspädagogik der Georg-August-Universität Göttingen, Bd. 28.
- Weinert, F. E. (2001). Concept of competence: A Conceptual Clarification. In D. S. Rychen & L. H. Salganik (Eds), *Defining and Selecting Key Competencies* (pp. 45–65). Seattle et al.: Hogrefe & Huber.
- Weinert, F.E. (2002). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F.E. Weinert (Ed.), *Leistungsmessung in Schulen*, (S. 17–31). Weinheim und Basel: Beltz.
- Weiß, M., & Preuschoff, C. (2004). *Kosten und Effizienzanalysen im Bildungsbereich. Referenzpapier für das Projekt „Bildungssteuerung“*. Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.
- Wigdor, A.K., & Green, B.F. (1991). *Performance Assessment for the Workplace, Volume I: Committee on the Performance of Military Personnel*. Commission on Behavioural and Social Science and Education. National Research Council. Washington, D.C.: National Academy Press.
- Wild, K.-P., & Schiefele, U. (1994). Lernstrategien im Studium. Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 15, S. 185–200.
- Winther, E. (2005). Trait- und state-Komponenten der Motivation in Lernprozessen zwischen Interventions- und Forschungsanspruch. In P. Gonon, F. Klauser & R. Huisinga (Hrsg.), *Kompetenz, Kognition und neue Konzepte der beruflichen Bildung* (S. 219–236). Wiesbaden: Verlag für Sozialwissenschaften.
- Winterton, J., Delamare-Le Deist, F., & Stringfellow, E. (2005). *Typology of knowledge, skills and competences: clarification of the concept prototype: Final draft*. Thessaloniki: CEDEFOP.
- Wirth, J. (2003). *Selbstreguliertes Lernen in dynamischen Systemen. Unpublished Dissertation*. Berlin: Humboldt-Universität zu Berlin.
- Wirth, J., & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education*, 10, 329–345.
- Wolter, H.-J., & Petersen, H. (2003). *Ankauf und Aufbereitung von Unfallfahrzeugen für den CG-Fuhrpark: Expertenprojekt im Modellversuch GAB; Berufsgruppe Automobilmechaniker*. http://www.nibis.de/nli1/bbs/archiv/rahmenrichtlinien/gab/am_ep_fertig.pdf

A N H A N G

- I Anhang Kapitel 1: Teilnehmerliste:
1. Internationaler Workshop für
Berufsbildungs-PISA
(Göttingen, 18. bis 20. November 2004)
Teilnehmerliste:
2. Internationaler Workshop für
Berufsbildungs-PISA
(Göttingen, 28. bis 30. April 2005)
- II Anhang Kapitel 2: Konzept „Generalized Work Activities“
- Dokument 1: „GWA“ Reparieren und
Instandhaltung mechanischer
Ausstattung
- Dokument 2: Taxonomie der 42
„Generalized Work Activities“
- Dokument 3: Das Konzept der Lernaufgabe
- Dokument 4: Klassifikation von
Handlungsfunktionen
- Dokument 5: Fachdidaktische
Systematisierung
- III Anhang Kapitel 4: Ausbildungsprogramme mit beruflichem
Einschlag nach Ländern
Expertise Steffen Kühnel
Expertise Jürgen Rost

Teilnehmerliste: 1. Internationaler Workshop für ein Berufsbildungs-PISA (Göttingen, 18. bis 20. November 2004)

Name	Institution	Department	Participation Workshop	
			accepted	linked up with business
Achtenhagen, Frank	University of Göttingen	Institute for Economic and Business Education and Management Training	X	
Arends, Lena	University of Göttingen	Institute for Economic and Business Education and Management Training	X	
Babic, Edvin	Sociological Research Institute at the University of Göttingen (SOFI)		X	
Baethge, Martin	Sociological Research Institute at the University of Göttingen (SOFI)		X	
Baethge-Kinsky, Volker	Sociological Research Institute at the University of Göttingen (SOFI)		X	
Bailey, Thomas	University of Columbia	Teachers College/Institute of Education and Economy		X
Birzea, César	University of Bucarest	Institut of Educational Sciences		X
Brown, Alan	University of Warwick	Social Studies Building/Warwick Institute for Employment Research	X	
Descy, Pascaline	CEDEFOP		X	
Grootings, Peter	European Training Foundation	Department: Eastern Europe and Central Asia/Villa Gualino		X
Halász, Gábor	Országos Közoktatási Intézet	National Institute for Public Education		X
Hämäläinen, Kauko	University of Helsinki	Palmenia Centre for Continuing Education		X
Hoffmann, Walter	Federal Ministry of Economics and Labour (BMWA)		X	
Kalous, Jaroslav				X
Kauppi, Antti	University Helsinki	Palmenia Centre for Continuing Education	X	
Lopez, Alberto	CEREQ (Marseille)		X	
Munk, Peter	Federal Ministry of Education and Research (BMBF)		X	
Nijhof, Wim J.	University of Twente	Department of Education	X	
Olsen, Ole J.	University of Bergen	Department of Sociology	X	
Paulsen, Bent	Federal Institute for Vocational Education and Training (BIBB)		X	
Renold, Ursula	Schweizerisches Bundesamt für Berufsbildung und Technologie (BBT)			X
Schweri, Jürg	University of Bern	Centre for Research in Economics of Education	X	
Sørensen, John Houman	Aalborg University	FIB 1-92A	X	
Stevenson, John	Griffith University/Mt Gravatt Campus	Centre for Learning Research	X	
Svetlik, Ivan	University of Ljubljana	Faculty of Social Sciences		X
Wagner, Sandra	Federal Ministry of Education and Research (BMBF)		X	
Weber, Susanne	University of Göttingen	Institute for Economic and Business Education and Management Training	X	
Winther, Esther	University of Göttingen	Institute for Economic and Business Education and Management Training	X	

Teilnehmerliste: 2. Internationaler Workshop für ein Berufsbildungs-PISA (Göttingen, 28. bis 30. April 2005)

Name	Institution	Department	Mail	Country
Achtenhagen, Frank	University of Göttingen	Institute for Economic and Business Education and Management Training	fachten@wipaed.wiso.uni-goettingen.de	Germany
Arends, Lena	University of Göttingen	Institute for Economic and Business Education and Management Training	larends@wipaed.wiso.uni-goettingen.de	Germany
Babic, Edvin	Sociological Research Institute at the University of Göttingen (SOFI)		etabac@gwdg.de	Germany
Baethge, Martin	Sociological Research Institute at the University of Göttingen (SOFI)		mibaethg@gwdg.de	Germany
Baethge-Kinsky, Volker	Sociological Research Institute at the University of Göttingen (SOFI)		vbaethg@gwdg.de	Germany
Badesou, Mircea	European Training Foundation, Turin	Department for Enlargement and South Eastern Europe	Mircea.Badesou@eft.eu.int	Romania
Bailey, Thomas	University of Columbia	Teachers College/Institute of Education and Economy	tb3@columbia.edu	USA
Bouder, Annie	CEREQ (Marseille)		bouder@cerq.fr	France
Brown, Alan	University of Warwick	Social Studies Building/Warwick Institute for Employment Research	Alan.Brown@warwick.ac.uk	United Kingdom
Coffield, Frank	LONDON University	Institute of Education	F.Coffield@be.ac.uk	United Kingdom
Haapakorpi, Arja	University of Helsinki	Palmenia Centre for Continuing Education	arja.haapakorpi@helsinki.fi	Finland
Hamilton, Stephen F.	Cornell University, Ithaca, NY	Associate Provost for Outreach Professor Human Development	sfh3@cornell.edu	USA
Hoffmann, Walter	Federal Ministry of Economics and Labour (BMWA)		walter.hoffmann@bmwa.bund.de	Germany
Homs, Oriol	OIREM-Foundation (Barcelona)		oriol.homs@ciem.org	Spain
Munk, Peter	Federal Ministry of Education and Research (BMBF)		Peter.Munk@bmbf.bund.de	Germany
Nijhof, Wim J.	University of Twente	Department of Education	n.jhof@edte.utwente.nl	Netherlands
Oates, Tim	Qualification and Curriculum Authority (QCA)	Research and Statistics	oates@qca.org.uk	United Kingdom
Olsen, Ole J.	University of Bergen	Department of Sociology	ole.j.olsen@soc.uib.no	Norway
Paulsen, Bent	Federal Institute for Vocational Education and Training (BIBB)		paulsen@bibb.de	Germany
Schweni, Jürg	University of Bern	Centre for Research in Economics of Education	juerg.schweni@vwi.unibe.ch	Switzerland
Säljö Roger	University of Gothenburg	Department of Education	rsaeljo@gwdg.de	Sweden
Sørensen, John Houman	Aalborg University	FIB 1-92A	houman@socsci.aau.dk	Denmark
Stevenson, John	Griffith University/Mt Gravatt Campus	Centre for Learning Research	J.Stevenson@griffith.edu.au	Australia
Svetlik, Ivan	University of Ljubljana	Faculty of Social Sciences	ivan.svetlik@uni.lj.si	Slovenia
Thang, Per-Olof	University of Gothenburg		per-olof.thang@ped.gu.se	Sweden
Verhoeven, Jef	Belgien		Jef.Verhoeven@kuleuven.ac.be	Belgium
Weber, Susanne	University of Göttingen	Institute for Economic and Business Education and Management Training	sweber@wipaed.wiso.uni-goettingen.de	Germany
Wagner, Sandra	Federal Ministry of Education and Research (BMBF)		sandra.wagner@bmbf.bund.de	Germany

Anhang Kapitel 2

Das Konstrukt der „Generalized Work Activities“ basiert auf einer psychologisch begründeten taxonomischen Struktur, wonach Arbeitsverhalten durch die folgenden drei Komponenten bestimmt ist:

- „*Stimulus*“: *die vom Individuum empfangene Information*
- „*Organism*“: *mentale Prozesse des Individuums*
- „*Response*“: *Arbeitstätigkeit/Verhalten des Individuums als Reaktion auf einen Stimulus*

Da Arbeitstätigkeiten/Verhaltensformen („Activities“) auch von externen Faktoren beeinflusst werden, wurde diese Taxonomie um *Interaktionen mit anderen* und *Arbeitskontexte* ergänzt. Das Modell besagt also, dass die interaktiven Komponenten Information, mentale Prozesse und Output der Arbeit innerhalb von Interaktionsprozessen sowie Arbeitskontexten stattfinden. Durch Ergebnisse einer Faktorenanalyse wurde bestätigt, dass die Analyse von individuellen Arbeitstätigkeiten auf der Basis von Fragebögen hinreichend Einblicke in die unterliegende Struktur menschlicher Arbeit gewährleistet. Da dieser Ansatz nicht spezifisch hinsichtlich Technologien oder Aufgaben ist, können allgemeine, arbeitsübergreifende Strukturen analysiert werden.

Die Entwicklung des „GWA“-Fragebogens basiert auf der Grundlage der S-O-R-Taxonomie, ergänzt um Interaktionsprozesse. Um ein Modell für das „GWA“-Konstrukt sowie Definitionen und Skalen für die Levels von Ratingskalen für jede „GWA“ zu entwickeln, wurden eine Analyse von bestehenden taxonomischen Strukturen vorgenommen sowie Arbeitsanalysen durchgeführt. Bei der Entwicklung des „GWA“-Konstrukts wurden die folgenden Kriterien zugrundegelegt:

- *Fundierung durch mehrere Forschungsergebnisse;*
- *auf der Grundlage unterliegender Inhalte;*
- *umfassende und bestmögliche Repräsentation aller Jobs in den USA;*
- *Bereitstellen einer einzigartigen deskriptiven Informationsdatenbank.*

Diese Strategie berücksichtigt auch die Frage der Spezifität: „GWA“ können sowohl auf einem sehr allgemeinen, aber auch einem enger definierten Level definiert werden (auf einer Skala von allgemein bis spezifisch). Wenn Arbeitsanalysen auf einem mehr spezifischen Level durchgeführt werden, sind sie eher dazu geeignet, berufsspezifische Informationen mit einem feineren Differenzierungsgrad zu erbringen.

Insgesamt wurden innerhalb der Taxonomie 42 „GWA“ ermittelt, die im Folgenden anhand der einzelnen Bereiche beispielhaft erläutert werden: Der Input von Informationen (Stimulus) beinhaltet zwei Faktoren:

- *Recherchieren und Empfangen von arbeitsbezogenen Informationen;*
- *Identifizieren/Evaluieren von arbeitsrelevanten Informationen.*

Diese GWAs wurden auf der Grundlage von nomothetischen Fragebögen gewonnen (PAQ, Position Analysis Questionnaire; OAI, Occupational Analysis Inventory; GWI, General Work Inventory). Die folgenden „GWA“ beschreiben die Arbeitstätigkeiten bzw. das Verhalten („Activities“) von Individuen beim Recherchieren und Empfangen von arbeitsbezogenen Informationen:

- *Empfangen von relevanten Informationen, um die Arbeit ausführen zu können*
- *Prozesse, Materialien und Umgebungen überwachen*

Nachdem Informationen empfangen werden, müssen sie identifiziert und evaluiert werden. In der Forschungsliteratur wurden zur Beschreibung dieser Arbeitstätigkeiten bzw. dieses Verhaltens die folgenden drei GWA ermittelt:

- Identifizieren von Objekten, Handlungen und Ereignissen;
- Inspizieren von Ausstattungen, Strukturen und Materialien;
- Einschätzen der Eigenschaften von Materialien, Produkten, Ereignissen oder Informationen.

Mentale Prozesse beginnen, nachdem arbeitsrelevanter Input empfangen wurde. Die in Anspruch genommenen kognitiven Fähigkeiten von Individuen werden hinsichtlich der Faktoren Informations- oder Datenprozesse und Schlussfolgern/Entscheiden klassifiziert. Vier „GWA“ wurden zur Beschreibung von Informations- oder Datenprozessen in der Forschungsliteratur ermittelt:

- *Beurteilen der Qualität von Objekten, Dienstleistungen oder Personen;*
- *Informationen evaluieren zur Einhaltung von Standards;*
- *Informationen verarbeiten;*
- *Daten oder Informationen analysieren.*

Diese „GWA“ tauchten bei der Analyse für alle Levels der Arbeit auf. Im Gegensatz dazu waren bei den „GWA“ zur Beschreibung von Schlussfolgern/Entscheiden nur zwei von sechs „GWA“ zur Analyse aller Levels geeignet:

- *Entscheidungen treffen und Probleme lösen;*
- *Aktualisieren und Einsetzen von arbeitsrelevantem Wissen.*

Die restlichen vier tauchten hingegen in der Literatur hauptsächlich bei Studien zu Vorgesetzten, Managern etc. auf:

- *Kreatives Denken;*
- *Objekte und Strategien entwickeln;*
- *Arbeit und Arbeitstätigkeiten/Verhalten planen;*
- *Organisieren, Planen und Prioritäten setzen bei der Arbeit.*

Die vollständige Übersicht der Taxonomie mit 42 „GWA“ befindet sich in Dokument 2.

Jeder „GWA“ ist eine operationale Definition sowie eine Skala zugeordnet. Die „GWA“ *Empfangen von relevanten Informationen, um die Arbeit ausführen zu können* wird definiert als: Beobachten, Empfangen und darüber hinaus Erhalten von Informationen aus allen relevanten Quellen. Die

Skala hat sieben Stufen: 7 = Beschaffung neuer Informationen von vielen Quellen – oftmals durch aktive Interaktionen mit den Quellen. 1 = Normaler Gebrauch derselben Informationsarten aus einer Informationsquelle.

Für die Entwicklung des „GWA“-Fragebogens wurden drei Arten von Ratingskalen eingesetzt: Level-, Wichtigkeits- und Häufigkeitsskalen. Jede GWA hat eine eigene Definition der Skalen, welche die Komplexität messen. Die Levelskala ist eher für ein Rating zur Repräsentation berufsübergreifender Komplexität und weniger für die Erfassung von Komplexität innerhalb bestimmter Berufsfelder geeignet. Sie umfasst sieben Stufen mit einer zusätzlichen Option für „nicht relevant“. Zusätzlich wurden jeweils drei verhaltensbezogene Aussagen entwickelt, um die hohen, mittleren und unteren Levels zu verankern. Die Wichtigkeitsskala umfasst fünf Stufen mit den verbalen Anker: nicht wichtig; etwas wichtig, wichtig, sehr wichtig und überaus wichtig. Die Häufigkeitsskala ist eine 7-Punkte Skala mit den verbalen Anker: einmal im Jahr oder weniger; häufiger als einmal im Jahr; häufiger als einmal im Monat; häufiger als einmal in der Woche; täglich; mehrmals täglich; stündlich oder öfter. Sowohl die Wichtigkeits- als auch die Häufigkeitsskala sind Ratings hinsichtlich der Komplexität innerhalb von Berufsfeldern. Das folgende Dokument zeigt einen Fragebogen für die „GWA“ *Reparieren und Instandhaltung mechanischer Ausstattungen* als Komponente des Outputs der Arbeit:

Dokument 1: „GWA“ Reparieren und Instandhaltung mechanischer
Ausstattungen

Level	
What level of this activity is needed to perform this job?	
Performing complex or non-routine repair, maintenance, or adjustment of mechanical equipment, often involving overhauls or rebuilding.	7
	6
Performing straightforward repair, maintenance, or adjustment of mechanical equipment using established, easy to understand procedures.	5
	4
	3
	2
	1
	NR

Importance	
How important is this activity to performance on this job?	
Not Important	Somewhat Important
1	2
	3
	4
	5
	Extremely Important

Frequency	
How often is this activity performed on this job?	
Once per year or less	More than once per year
1	2
	3
	4
	5
	6
	7

Dokument 2: Taxonomie der 42 „Generalized Work Activities“

Descriptions and Definitions of Generalized Work Activities		
Construct label	Operational definition	Level scale
Looking For and Receiving Job-Related Information		
1. Getting information needed to do the job.	Observing, receiving, and otherwise obtaining information from all relevant sources.	High—Getting new information from many sources, often by actively interacting with the sources. Low—Making regular use of the same types of information from a single source.
2. Monitoring processes, materials, or surroundings.	Monitoring and reviewing information from materials, events, or the environment, often to detect problems or to find out when things are finished.	High—Monitoring very complex processes, events, or circumstances. Low—Monitoring processes, events, or circumstances that are not complex.
Identifying and Evaluating Job-Related Information		
3. Identifying objects, actions, and events.	Identifying information received by making estimates or categorizations, recognizing differences or similarities, or sensing changes in circumstances or events.	High—Making extremely difficult identifications based on very complex information. Low—Making easy identifications based on information that is not complex.
4. Inspecting equipment, structures, or materials.	Inspecting or diagnosing equipment, structures, or materials to identify the causes of errors or other problems or defects.	High—Making inspections or diagnoses of a complex system that may have many interrelated parts and determining whether conditions exist within a range of acceptable limits. Low—Making easy judgments about the quality or importance of things or people when there are many guidelines.
5. Estimating the characteristics of materials, products, events, or information.	Estimating sizes, distances, and quantities, or determining time, costs, resources, or materials needed to perform a work activity.	High—Making very difficult estimates of characteristics, time, or resources where there is limited guidance or supporting information. Low—Making straightforward estimates of characteristics, time, or resources where there is considerable guidance and supporting information.
Information/Data Processing		
6. Judging the qualities of objects, services, or persons.	Making judgments about or assessing the value, importance, or quality of things or people.	High—Making very difficult judgments about the quality or importance of things or people for which there is limited guidance or supporting information. Low—Evaluating information against a simple criterion.
7. Evaluating information for compliance to standards.	Evaluating information against a set of standards and verifying that it is correct.	High—Evaluating complex information for compliance with regulations, laws, or technical criteria, where compliance decisions require significant interpretation or judgment. Low—Evaluating information against a simple criterion.
8. Processing information.	Compiling, coding, categorizing, calculating, tabulating, auditing, verifying, or processing information or data.	High—Processing very different and complicated data or information, where there are several ways in which the information can be processed. Low—Processing data or information that is standardized and easy to understand, where there is only one way to process the information.
9. Analyzing data or information.	Identifying underlying principles, reasons, or facts by breaking down information or data into separate parts.	High—Analyzing very different and complicated data or information that can be used for making critical decisions. Low—Analyzing data or information that is easy to understand.
Reasoning/Decisions Making		
10. Making decisions and solving problems.	Combining, evaluating, and reasoning with information and data to make decisions and solve problems. These processes involve making decisions about the relative importance of information and choosing the best solution.	High—Reaching conclusions after considering a large number of choices that are often ambiguous or abstract, where there are competing viewpoints and alternatives that must be considered before reaching final decisions and the solutions decided upon will have very significant impact. Low—Reaching conclusions after considering a few choices that are usually well defined, where there are a limited number of possible actions, and the decisions or solutions will have minor impact.

Table 8-1 continues

TABLE 8-1 (Continued)

Construct label	Operational definition	Level scale
11. Thinking creatively.	Originating, inventing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions.	High—Creating or inventing new and yet-to-be-proven practices, technologies, materials, products, or strategies, where the creative effort will have widespread impact and will result in substantial improvements for both an organization and its customers. Low—Offering suggestions for some change or improvement to immediate work functions or products.
12. Updating and using job-relevant knowledge.	Keeping up-to-date technically and knowing one's own job's and related jobs' functions.	High—Learning, retaining, and staying current with complex, often highly technical information. Low—Learning, retaining, and staying current with relatively easy-to-master information.
13. Developing objectives and strategies.	Establishing long-range objectives and specifying the strategies and actions to achieve these objectives.	High—Doing complex, future-oriented strategic planning. Low—Doing strategic or long-term planning that is not complex.
14. Scheduling work and activities.	Scheduling events, programs, activities, as well as the work of others.	High—Engaging in complex and difficult scheduling activities. Low—Engaging in simple or straightforward scheduling activities.
15. Organizing, planning, and prioritizing work.	Developing plans to accomplish work, and prioritizing and organizing one's own work.	High—Doing a high degree of complex planning, organizing, and prioritizing of one's own work. Low—Doing uncomplicated planning, organizing, or prioritizing of one's own work.
Performing Physical and Manual Work Activities		
16. Performing general physical activities.	Performing physical activities that require moving one's whole body, such as in climbing, lifting, balancing, walking, stooping, where the activities often also require considerable use of the arms and legs, such as in the physical handling of materials.	High—Making repetitive and often fatiguing extensive use of the whole body in completing work activities that are done with or without the use of tools. Low—Making nonfatiguing use of the whole body in completing work activities that are done with or without the use of tools.
17. Handling and moving objects.	Using one's own hands and arms in handling, installing, forming, positioning, and moving materials, or in manipulating things, including the use of keyboards.	High—Using one's hands and arms to do the same functions almost continually. Low—Making little use of one's hands and arms.
18. Controlling machines and processes.	Using either control mechanisms or direct physical activity to operate machines or processes (not including computers of vehicles).	High—Controlling machines or processes that are very difficult to operate. Low—Controlling machines or processes that are easy to operate.
19. Operating vehicles, mechanized devices, or equipment.	Running, maneuvering, navigating, or driving vehicles or mechanized equipment, such as forklifts, passenger vehicles, aircraft, or water craft.	High—Operating equipment or vehicles that are very difficult to run. Low—Operating equipment or vehicles that are easy to run.
Performing Complex/Technical Activities		
20. Interacting with computers.	Controlling computer functions by using programs, setting up functions, writing software, or otherwise communicating with computer systems.	High—Using computers to develop very complex, high speed data linkages and operating systems. Low—Using computers to produce standard correspondence, graphic materials, and business-related information.
21. Drafting, laying-out, and specifying technical devices, parts, and equipment.	Providing documentation, detailed instructions, drawings, or specifications to inform others about how devices, parts, equipment, or structures are to be fabricated, constructed, assembled, modified, maintained, or used.	High—Drafting and specifying the components or technical relationships for complicated devices, parts, or equipment. Low—Drafting or specifying the components or technical relationships for devices, parts, or equipment that are easily understood.
22. Implementing ideas, programs, systems, or products.	Conducting or carrying out work procedures and activities in accord with one's own ideas or information provided through directions/instructions for purposes of installing, modifying, preparing, delivering, constructing, integrating, finishing, or completing programs, systems, structures, or products.	High—Performing highly complex and very difficult work activities with very limited guidelines to follow. Low—Performing activities that have clear-cut directions and are easy to carry out.

Table 8.1 continues

TABLE 8-1 (Continued)

Construct label	Operational definition	Level scale
23. Repairing and maintaining mechanical equipment.	Fixing, servicing, aligning, setting up, adjusting, and testing machines, devices, moving parts, and equipment that operate primarily on the basis of mechanical (not electronic) principles.	High—Performing complex or nonroutine repair, maintenance, or adjustment of mechanical equipment, often involving overhauls or rebuilding. Low—Performing straightforward repair, maintenance, or adjustment of mechanical equipment using established, easy to understand procedures.
24. Repairing and maintaining electronic equipment.	Fixing, servicing, adjusting, regulating, calibrating, fine-tuning, or testing machines, devices, and equipment that operate primarily on the basis of electrical or electronic (not mechanical) principles.	High—Performing complex or nonroutine repair, maintenance, or adjustment of electronic equipment, where repairs are often made to complex internal components or circuitry. Low—Performing straightforward repair, maintenance, or adjustment of electronic devices or equipment using established, easy to understand procedures.
25. Documenting and recording information.	Entering, transcribing, recording, storing, or maintaining information in either written form or by electronic/magnetic recording.	High—Documenting or recording very complex information using new, unstandardized procedures. Low—Documenting or recording straightforward information using predetermined forms and procedures.
Communicating/Interacting		
26. Interpreting the meaning of information for others.	Translating or explaining what information means and how it can be understood or used to support responses or feedback to others.	High—Making very difficult interpretations of information with limited, if any, guidance to follow. Low—Making easy interpretations of information with a high degree of guidance to follow.
27. Communicating with supervisors, peers, or subordinates.	Providing information to supervisors, fellow workers, and subordinates. This information can be exchanged face-to-face, in writing, or via telephone/electronic transfer.	High—Providing complex oral and written communications to others in the organization. Low—Providing straightforward oral or written communications to others in the organization.
28. Communicating with persons outside the organization.	Communicating with persons outside the organization, representing the organization to customers, the public, government, and other external sources. This information can be exchanged face-to-face, in writing, or via telephone/electronic transfer.	High—Presenting complex oral and written communications to persons outside the organization. Low—Presenting routine and simple oral and written communications to persons outside the organization.
29. Establishing and maintaining interpersonal relationships.	Developing constructive and cooperative working relationships with others.	High—Developing very good interpersonal relationships with highly diverse individuals or stakeholders in difficult situations. Low—Developing very few working relationships with others.
30. Assisting and caring for others.	Providing assistance or personal care to others.	High—Providing care or assistance to others in highly stressful or difficult situations. Low—Needing to provide minimal help or assistance to others.
31. Selling or influencing others.	Convincing others to buy merchandise/goods, or otherwise changing their minds or actions.	High—Doing a lot of high-level persuading to accomplish work objectives, involving persuading a very difficult to convince audience. Low—Doing little persuading to accomplish work objectives, because there is little need to convince others in any area.
32. Resolving conflicts and negotiating with others.	Handling complaints, arbitrating disputes, and resolving grievances, or otherwise negotiating with others.	High—Handling complaints and negotiations in very challenging situations, involving complex matters and significant conflict and pressure. Low—Handling negotiations that involve very simple matters that are easily resolved or involve complaint-handling or negotiation.
33. Performing for or working directly with the public.	Performing for people or dealing directly with the public, including serving persons in restaurants and stores, and receiving clients or guests.	High—Handling interactions with the public, where the audience is hard to please or other conflict is involved. Low—Having little interaction with the public, or needing to have only brief interactions.
Coordinating/Developing/Managing/Advising Others		
34. Coordinating the work and activities of others.	Coordinating members of a work group to accomplish tasks.	High—Coordinating the work of many employees, where a complex sequencing of others' tasks is required. Low—Needing to do little coordinating of others.

Table 8-1 continues

TABLE 8-1 (Continued)

Construct label	Operational definition	Level scale
35. Developing and building teams.	Encouraging and building mutual trust, respect, and cooperation among team members.	High—Managing large teams and building cooperation among diverse team members toward accomplishment of highly complex or poorly defined activities/projects. Low—Doing little team building.
36. Teaching others.	Identifying educational needs, developing formal training programs or classes, and teaching or instructing others.	High—Teaching and explaining difficult tasks, concepts, or material and conducting complex training. Low—Doing little training or educating of others.
37. Guiding, directing, and motivating subordinates.	Providing guidance and direction to subordinates, including setting performance standards and monitoring subordinates.	High—Directing and motivating several organization members and building and maintaining morale in difficult or unpleasant work settings. Low—Doing little directing or motivating of subordinates.
38. Coaching and developing others.	Identifying developmental needs of others and coaching or otherwise helping others to improve their knowledge or skills.	High—Identifying effective ways of developing others to perform highly complex or difficult tasks and coaching them under these difficult conditions. Low—Doing little coaching or developing of others.
39. Providing consultation and advice to others.	Providing consultation and expert advice to management or other groups on technical, systems-related, or process-related topics.	High—Providing expert guidance on complex matters regarding the design, development, or implementation of major programs. Low—Providing little advice or consultation to others.
Administering		
40. Performing administrative activities.	Approving requests, handling paperwork, and performing day-to-day administrative tasks.	High—Overseeing administrative activities for a large workforce, with a complex set of administrative procedures. Low—Doing very straightforward administrative activities.
41. Staffing organizational units.	Recruiting, interviewing, selecting, hiring, and promoting persons for an organization.	High—Overseeing the staff of a large and diverse workforce, with complex staffing needs. Low—Doing very straightforward staffing activities.
42. Monitoring and controlling resources.	Monitoring and controlling resources and overseeing the spending of money.	High—Monitoring and controlling a large number of resources, including managing a large budget. Low—Needing to do little monitoring or controlling of resources or money.

Dokument 3: Das Konzept der Lernaufgabe

Ein übergreifendes Prinzip stellt auch die Erarbeitung von Lernaufgaben dar, wie sie vor allem vom BIBB vorgeschlagen werden:

Bei Lernaufgaben handelt es sich um eine Systematik zur Organisation von Lerntätigkeiten an definierten Lerngegenständen. Theoretische Grundlagen werden nur so weit vermittelt, wie sie für das Verständnis konkreter Lernhandlungen erforderlich sind. Lernaufgaben orientieren sich an der Wirklichkeit beruflichen Lernens, also an berufstypischen Arbeitsaufgaben innerhalb der betrieblichen Prozesse. Die typische Struktur der Arbeitsaufgabe gibt die Struktur der Lernaufgabe vor.

Die Gesamtaufgabe wird über Teilaufgaben ermittelt, die anhand von übergeordneten Strukturen klassifiziert und in eine sinnvolle und nachvollziehbare Abfolge gebracht werden. Anschließend werden – unabhängig von einem konkreten Auftrag – unterschiedliche Anforderungsniveaus als Grundlage für die Zuordnung von Teilaufgaben definiert. Am Ende liegt ein System aus gestuften Lernaufgaben vor; dabei repräsentiert die Gesamtlernaufgabe zugleich die höchste Stufe des gesamten Systems. Schließlich werden für jede Lernaufgabe ein passender betrieblicher Auftrag gesucht und der Lerngehalt beurteilt. Die Wissensvermittlung wird bei jeder Lernaufgabe hinsichtlich des deklarativen und prozeduralen Wissens differenziert.

Bei der Auswahl von typischen Arbeitsaufgaben sollen vornehmlich drei Aspekte berücksichtigt werden:

1. Ermittlung von typischen Arbeitsaufgaben im Betrieb;
2. Analyse der Anzahl von Arbeitsaufträgen pro Arbeitsaufgabe;
3. Ermittlung der Ausschnitte aus dem Geschäftsprozess, die bei der Durchführung der Arbeitsaufgabe durchlaufen werden.

Eine Sammlung von Lernaufgaben liegt vor (vgl. die drei vom BIBB 2005 herausgegebenen Hefte); für die Itementwicklung bietet es sich an, die mit diesem Vorgehen gewonnenen Erfahrungen sowie die Aufgabenformate zu prüfen, zu übernehmen bzw. spezifische Alternativen zu konstruieren. Die Übertragbarkeit der zugrundeliegenden Konstruktionsannahmen auf andere berufliche Bereiche als die bisher bearbeiteten ist zu evaluieren.

Dokument 4: Klassifikation von Handlungsfunktionen

Franke 2005 hat „Facetten der Kompetenzentwicklung“ äußerst kenntnisreich und detailliert zusammengetragen und dabei insbesondere Fragen einer weiteren möglichen Aufgliederung von Wissen und Handeln und ihrer Komponenten – vor allem im Hinblick auf Bestimmungen ihrer Schwierigkeiten – diskutiert. Die folgende Abbildung zeigt eine taxonomische Klassifikation intrinsischer Handlungsfunktionen – wie er es nennt – im Bereich der Reparaturtätigkeit.

Tabelle 5-1: Taxonomische Klassifikation intrinsischer Handlungsfunktionen im Bereich der Reparaturtätigkeit (Ausschnitt)

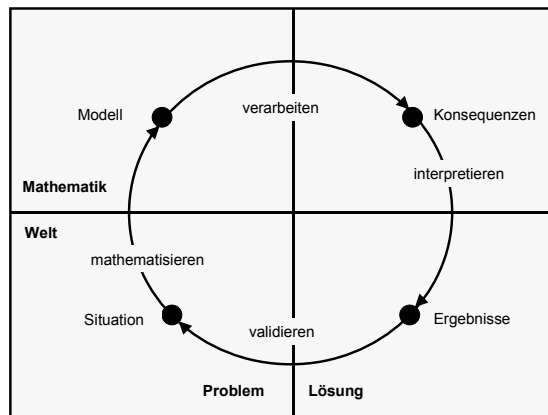
Stufe 3	Stufe 2	Stufe 1
1. Klärung der Rahmenbedingungen	• • • • •	• • • • •
2. Störungsanalyse	2.1. Symptomverifikation	Befragung des Maschinenbedieners
		Sichtkontrolle
		Durchführung von Probeläufen (evtl. unter verschiedenen Bedingungen)
	2.2. Rekonstruktion des Störfalls	Prüfung, ob Fehler durch den Maschinenbediener verursacht wurde
		Klärung der Bedingungen des Auftretens der Störung
		Zeitpunkt für das Auftreten der Störung
		Berücksichtigung möglicher Sabotage (z. B. vertauschte Schläuche in der Pneumatik/Hydraulik)
3. Intuitive Fehlersuche	• • • • •	• • • • •
4. Gewinnung eines Systemverständnisses	• • • • •	• • • • •
5. Entscheidung über die Auftragsübernahme	• • • • •	• • • • •
6. Hypothesenbildung (Suche nach möglichem Fehler)	6.1. Zurückverfolgen der Wirkungskette	Zurückverfolgen - des Stoffflusses
		- des Energieflusses
		- des Informationsflusses
	6.2. Aufbau einer Wirkungskette durch Vorwärtsverkettung	
	6.3. Überprüfung der Schwachstellen der Maschine	Einsicht in Diagnosepapiere und Instandhaltungsberichte
		Achten auf Verschleißteile
	6.4. Abarbeiten von störungsspezifischen Checklisten (Fehlerlisten)	
	6.5. Einbeziehung des Maschinenbedieners in die Hypothesenbildung	
7. Hypothesenprüfung (Entscheidung über Reihenfolgen der Hypothesenprüfung und über Prüfmethoden)	7.1. Prüfung der Hypothesen unter dem Gesichtspunkt der Fehlerwahrscheinlichkeiten/Fehlerstatistiken	• • • • •

Stufe 3	Stufe 2	Stufe 1
	7.2. Blockung der Wirkungskette, Durchführung einer Komplexdiagnose vor einer Tiefendiagnose	
	7.3. Prüfung der Hypothesen nach der Aufwandkalkulation	
	7.4. Diagnostizieren, ob Zustandsparameter des jeweiligen Teilsystems in den zulässigen Toleranzbereichen liegen	
	7.5. Durchführung indikativer Veränderungen und Beobachtung von Signalverläufen	
8. Suchraumerweiterung	8.1. Selbstkontrolle / Rekapitulation der vollzogenen Handlungsschritte	
	8.2. Erweiterung des zu analysierenden Systemausschnittes (bis hin zum Testen eines „Apriori-Fehlerbaumes“)	
	8.3. Unsystematisches Herumsuchen („trial and error“)	
	8.4. Berücksichtigung „komplizierter“ Hypothesen (z. B. temporäre Fehler)	
	8.5. Konsultation von Fachleuten	
	8.6. Konsultation eines Instandhaltungsprogrammes (eines wissensbasierten Systems)	
	8.7. Vom Problem Abstand nehmen	
9. Organisation der Arbeitsmittel		
10. Eingriff in die Maschine zwecks Fehlersuche und -beseitigung		
11. Suche nach den Ursachen für die Fehlerentstehung		
12. Effektkontrolle		
13. Leistungsdokumentation		
14. Verbesserung der Wartungsplanung		
15. Evaluation der Arbeit	15.1. Vergewärtigen des Vorgehens bei der Leistungserbringung	
	15.2. Nachdenken über alternative Vorgehensweisen	
	15.3. Suche nach den Ursachen von Handlungsfehlern	
	15.4. Effizienzbeurteilung der praktizierten Vorgehensweise	
	15.5. Bewertung der Arbeit im Hinblick auf die Persönlichkeitsentwicklung	
	15.6. Entwicklung von Perspektiven für das künftige Handeln	

Dokument 5: Fachdidaktische Systematisierungen

Der Rückbezug auf fachdidaktische Überlegungen ermöglicht den systematischen Zugriff auf die Auswahl von Items sowie deren Einordnung nach deren Anforderungsgehalt. Dabei lassen sich durchaus Modelle formulieren, die einen Rahmen für die Items abgeben und dabei Prinzipien verdeutlichen, die der Ziel- und Inhaltsauswahl zugrunde liegen sollten. Die folgende Abbildung zeigt die Modellierung des Prozesses des Mathematisierens, wie er vom PISA-Konsortium zugrunde gelegt wurde (Klieme, Neubrand & Lüdtke, 2001, S. 143 f.).

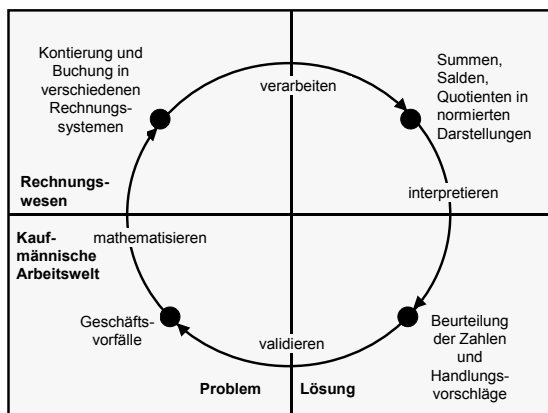
Abbildung: Prozess des Mathematisierens



Der dort beschriebene Prozess kann als ein Prozess des ökonomischen Mathematisierens auf die kaufmännische Arbeitswelt und das Rechnungswesen übertragen werden (Preiß, 1999).

Ausgangs- und Endpunkt der Prozesse ist die Situation, der im ökonomischen Bereich die Geschäftsvorfälle einer Periode entsprechen. Die einzelnen Geschäftsvorfälle sind im ersten Schritt aus der jeweiligen Repräsentationsform (Beleg, Belegbeschreibung, Standardformulierung) zu rekonstruieren und im Hinblick auf ein Rechnungswesen-Modell zu formalisieren. Den mathematischen Modellen (Formeln) entsprechen dabei die verschiedenen Instrumente des Rechnungswesens (Kontierungsstempel, Haupt- und Nebenchkonten, Kosten- und Finanzstellen, Aufträge, Ergebnisobjekte). Ähnlich den Formeln werden diese im Laufe des Unterrichts kontinuierlich erweitert.

Abbildung: Prozess des ökonomischen Mathematisierens



Die Verarbeitungsschritte sind dann im Wesentlichen Übertragungsoperationen und rechnerische Transformationen. Diese können sich über viele Teilschritte erstrecken und zu mehreren verdichteten Darstellungen führen. Die Besonderheit des Rechnungswesens besteht darin, dass dabei auch ökonomische Kalküle (Fristigkeiten, Chancen und Risiken) und rechtliche Normen im Hinblick auf die Interessen verschiedener Adressaten der Rechnungslegung zu beachten sind.

Im dritten Schritt sind dann die errechneten Werte im Hinblick auf die Ausgangssituation/-periode als mögliches Ergebnis zu interpretieren. Im Bereich des Rechnungswesens bedeutet dies, dass eine ökonomische Beurteilung der im jeweiligen Rechnungslegungsinstrument abgebildeten Situationen abzugeben ist. Dies schließt die Wertung der Situationen einer Periode ein. In einem möglichen vierten Schritt können dann die aus dem Rechnungswesen heraus interpretierten Ergebnisse in Handlungsvorschläge für neue Situationen, d. h. neue Geschäftsvorfälle übersetzt werden. Dabei ist das Modell, das die Ergebnisse geliefert hat, kritisch zu betrachten. Dies bedeutet, dass es sich im Zuge der aus ihm gezogenen Schlussfolgerungen zu bewähren hat, d. h. es wird validiert.

Auf der Grundlage dieser Modellannahmen lassen sich dann Ziele und Inhalte für unterscheidbare Kompetenzstufen formulieren:

Abbildung: Kompetenzstufen für Mathematik und Rechnungswesen/Controlling

Stufe	Mathematik	Rechnungswesen/Controlling
I	Rechnen auf Grundschulniveau - arithmetisches und geometrisches Wissen auf Grundschulniveau - Abrufen von Wissen und unmittelbares Anwenden, wenn die Aufgabenstellung von vornherein eine bestimmte Standard-Mathematisierung nahe legt - begriffliche Modellierungen sind nicht leistbar	Rechnen und Argumentieren auf Alltagsniveau - Rechentechnisches und ökonomisches Wissen auf Alltagsniveau - Abrufen von Wissen und unmittelbares Anwenden bei einfachen Darstellungen von Standardgeschäftsvorfällen - begriffliche Modellierungen sind nicht leistbar
II	Elementare Modellierungen - einfachste begriffliche Modellierungen, die in einen außermathematischen Kontext eingebettet sind - Finden unter mehreren möglichen Lösungsansätzen den richtigen, wenn durch Graphiken, Tabellen, Zeichnungen usw. eine Struktur vorgegeben ist - Auf dieser Stufe sind nur drei Wissensinhalte der Grundschulmathematik verfügbar	Elementare Buchungen und Interpretationen von Abweichungen - einfachste begriffliche Modellierungen, die in einen ökonomischen Kontext eingebettet sind - richtige Buchung und Interpretation, wenn durch Graphiken, Tabellen, Zeichnungen, Merksätze usw. ein Vorstrukturierung gegeben ist - auf dieser Stufe sind nur die Wissensinhalte der Grundrechenarten, Prozent- und Zinsberechnungen bei Aggregationen, Aufteilungen und Vergleichen in Verbindung mit ökonomischen Interessen verfügbar
III (Standard-Grundbildung)	Modellieren und begriffliches Verknüpfen auf dem Niveau der Sekundarstufe I - Verfügen über einfache Wissensinhalte der Sekundarstufe I, also über den Standardstoff der Lehrpläne aller Schulformen - Verknüpfen von Konzepten aus unterschiedlichen Bereichen und Nutzen zur Lösung von Problemstellungen, wenn visuelle Darstellungen den Lösungsprozess unterstützen	Buchen, Aufbereiten und Interpretieren auf dem Niveau der kaufmännischen Erstausbildung - Verfügen über einfache Wissensinhalte der kaufmännischen Erstausbildung, also über den Standardstoff in Buchführung, Kostenrechnung und Controlling aller kaufmännischer Ausbildungsberufe - Verknüpfen von Konzepten aus unterschiedlichen Rechnungswesen-Teilbereichen und der BWL, wenn visuelle Darstellungen den Lösungsprozess unterstützen
IV	Umfangreiche Modellierungen auf der Basis anspruchsvoller Begriffe - Bewältigen im technischen Bereich umfangreiche Verarbeitungsprozesse - Aufbauen einer Lösung über mehrere Zwischenergebnisse hinweg - Bewältigung auch offener Modellierungsaufgaben, bei denen man unter vielfältigen Lösungswegen einen eigenen finden muss - Verstärkte Modellierung innermathematischer begrifflicher Zusammenhänge	Umfangreiche Buchungen und Auswertungen auf der Basis anspruchsvoller Begriffe - Bewältigung umfangreicher informationstechnischer Verarbeitungsprozesse - Aufbauen einer betriebswirtschaftlichen Interpretation über mehrere Rechnungsperioden oder Rechnungswesen-Teilbereiche bzw. andere Teile integrierter Informationssysteme hinweg - Entwicklung von alternativen Lösungswegen zu vorgegebenen Standardlösungen - Verstärkte Berücksichtigung von Theorien zur Rechnungslegung und Kostenrechnung
V	Komplexe Modellierungen und innermathematisches Argumentieren - Verfügbarkeit anspruchsvollen curricularen Wissens - Bewältigen von sehr offen formulierten Aufgaben, bei denen ein Modell frei gewählt bzw. selbst konstruiert werden muss - Begründen und Beweisen sowie Reflektieren über den Modellierungsprozess selbst	Komplexe Buchungen und rechnungsziel- und -instrumentbezogenes Argumentieren - Verfügbarkeit anspruchsvollen curricularen Wissens - Bewältigung von sehr offen formulierten Aufgaben, bei denen eine zweckmäßige Verrechnungstechnik im Rahmen zulässiger Normen angewandt bzw. gefunden werden muss - Begründen und Beweisen sowie Reflektieren über Techniken, Prämissen und Normen des Rechnungswesens und betriebswirtschaftlicher Modelle

Anhang Kapitel 4

Tabelle 1: Ausbildungsprogramme mit beruflichem Einschlag nach Ländern:
 Programmbezeichnung, Art des Programms
 (P=Prevocational, V=Vocational, G=General), ISCED-Level des
 Ausbildungsziels, typisches Eintrittsalter und Teilnehmerzahlen

Land	Programmbezeichnung	Art des Pro-gramms	ISCED-Level	Typisches Eintritts-alter	Teilneh-merzahlen
Australien					
	Traditional apprenticeship	V	3b	15–18	
	Modern apprenticeship (long term)	V	3b	15–45	
	Upper secondary school-based VET (TAFE)	V	3	15–45	
Dänemark					
	Basic vocational Education (EGU)	P	3c	16–20	2.000
	Social and health service assistant (SOSU)	V	3c	17–30	12.000
	Upper secondary higher technical oder commercial education	G	3a	16–17	34.000
	Individual organized Youth education	P	3	16–20	4.000
	Upper secondary VET (alternative)	V	3c	16–20	114.000
Deutschland					
	Berufsgrundbildungsjahr	V	3b	16–18	40.000
	Berufsfachschulen (berufliche Grund-kenntnisse)	V	3b	16–17	
	Kollegschulen	V	3b	17–18	70.000
	Berufsfachschulen (mit beruflichem Abschluss)	V	3b	17–17	286.000
	Duales System	V	3b	16–18	1.337.000
	Schulen des Gesundheitswesens		5b	19–20	
Finnland					
	Upper Secondary VET-programs (including apprenticeships)	V	3a	16–18	162.000
Frankreich					
	Vocational Training for young people without qualification (Lehrlingsausbildung) (CAP, BEP, BP)	V	3c	16–25	100.000
	Secondary education (2 nd circle-CAP)	V	3c	15–16	80.000
	Secondary education (2 nd circle-BEP)	V	3c	15–17	510.000
	Secondary education (2 nd circle-CAP) combined with apprenticeship	V	3c	16–18	185.000
	Secondary education (2 nd circle-BEP) combined with apprenticeship	V	3c	16–18	42.000
	Specific schools (health and social care)	V	3c	16–20	16.000
	Secondary education (2 nd circle-BP) combined with apprenticeship	V	3b	18–20	20.000
	Secondary education (2 nd circle-BP)-school-based programm	V	3b	18–19	150.000
	Specific schools (health and social care)	V	4a	18–20	
Niederlande					
	Different types of VET including apprenticeship	V	3	16	449.00
	Basic vocational training-level 2	V	3c	16	
	Professional training level 3	V	3c	16	
	Middle management training level 4	V	3a	16	

Norwegen					
	Upper secondary vocational (including apprenticeships)	V	3c	15–17	125.000
Rumänien					
	Vocational school	V	3c	15	
	Apprenticeship school	V	3c	15	
Schweiz					
	Anlehre (basic VET)	V	3c	15–17	3.000
	Berufslehre (apprenticeship or full time-schooling)- 2 years	V	3c	15–17	20.000
	Berufslehre (apprenticeship or full time-schooling)- 3 or 4 years	V	3b	15–17	149.000
	Berufsmaturität (combined with apprenticeship) 2 years	G/V	3a	15–7	13.000
	Ausbildung für Krankenpflege		4b	18–20	8.000
Schweden					
	Upper secondary school	G/V	3a	16	304.000
	Upper secondary school - individual program	V	3c	16	20.000
	Upper secondary education for pupils with learning disabilities	V	3c	16–17	3.000
	Tertiary education (incl. Nurses)		5a	19–	
Slowenien					
	VET-program of short duration	V	3c	15	
	VET-Program (medium duration) - school type organisation	V	3c	15	
	VET-Program (medium duration) - dualsystem/apprenticeship	V	3c	15	
Ungarn	Secondary VET-school -Prematura	P	3a	14–15	188.000
	Vocational programmes preparing for NVQL examinations	V	3c	16–17	110.000 (geschätzt)
	College graduate education		5a	18–20	
	College first programmes		5a	18–20	
United Kingdom					
	GNVQ –Intermediate level	V	3c	15+	
	Activities leading to NVQ level 2	V	3c	16	
	Activities leading to NVQ level 1	V	3c	16	
	Traditional apprenticeship	V	3c	16–18	
	Work-based training	V	3c	16–17	
	Activities leading to NVQ level 3	V	3a	16+	
	GNVQ advanced level	V	3a	16+	
	Modern apprenticeship	V	3a	16–19+	
	Diploma in Higher Education (incl. Nurse training)		5b	18*	
USA					
	Vocational High Schools	P/V	?	16–17	
	Registered Apprenticeships	V	3c	16–?	
	School-based post-secondary VET (2-years college)	V	3c	18–20	

Quellen: Eurostat/CEDEFOP (2002), Office of Apprenticeship Training, Employer and Labor Services (2004), Lamb u. a. 2003, NCVER (2004)

Expertise Steffen Kühnel

Stichprobenplan für ein deutsches Pilotprojekt einer internationalen Vergleichsstudie zur beruflichen Ausbildung

Ganz offensichtlich hängt die Aussagekraft einer Vergleichsstudie zur beruflichen Ausbildung von der Qualität der Daten ab, die Grundlage dieses Vergleichs bilden. Datenqualität bedeutet, dass die berücksichtigten Daten relevant, valide, zuverlässig und bei Stichproben über die erfassten Elemente hinaus verallgemeinerbar sind.

In der Vergleichsstudie werden verschiedene Arten von Untersuchungseinheiten auf unterschiedlichen Ebenen betrachtet. Im Folgenden geht es ausschließlich um die Ebene der Teilnehmer und Absolventen beruflicher Ausbildung. Es werden alternative Stichprobenpläne vorgestellt, die es erlauben sollen, durch die Befragung einer noch zu bestimmenden Anzahl von Personen generalisierbare Aussagen über den Erfolg spezifischer beruflicher Ausbildungen zu gewinnen

Bezogen auf diese Gruppe der Teilnehmer und Absolventen beruflicher Ausbildung sollen folgende Untersuchungsfragen beantwortet werden:

- Welche Kompetenzen erwerben die Teilnehmer in ihrer beruflichen Ausbildung?
- Können sie diese Kompetenzen im späteren Berufsalltag nutzen?

Es wird davon ausgegangen, dass der Kompetenzerwerb nicht nur vom institutionellen Setting der beruflichen Ausbildung innerhalb eines Staates und von der Qualität des Bildungsangebots und der Bildungsvermittlung abhängt, sondern auch davon, welche Voraussetzungen die Teilnehmer zu Beginn ihrer Ausbildung mitbringen, wie motiviert sie sind und wie sehr sie sich in ihrer beruflichen Ausbildung engagieren. Die Nutzung der erworbenen Kompetenzen hängt wiederum davon ab, unter welchen Kontextbedingungen die Kompetenzen eingesetzt werden. Gerade im Hinblick auf eine international vergleichende Bewertung beruflicher Ausbildung erscheint es wenig sinnvoll, allein auf die zu erwerbenden Kompetenzen zu schauen, sondern diese in den jeweiligen Berufskontext zu setzen. Im Hinblick auf die andauernden Veränderungen der Berufskontexte durch technisch-wissenschaftliche Entwicklungen, aber auch soziale und politische Veränderungen ist zudem die Robustheit von Kompetenzen bei Veränderungen des Kontextes zu berücksichtigen. Aus der Sicht der Teilnehmer und Absolventen beruflicher Bildung bedeutet dies ihre Fähigkeit, sich auch unter veränderten Arbeitssituationen behaupten zu können.

A. Paneldesign

Für den Stichprobenplan folgt daraus, dass ein geeignetes Design sich nicht allein auf die Ausbildungssituation beschränken kann, sondern den Verlauf

vom Beginn der beruflichen Ausbildung an bis hin zu den späteren Berufstätigkeiten verfolgen sollte. Optimal wäre daher eine langfristige Panelstudie, bei der Informationen über die gleichen Personen zu Beginn ihrer Ausbildung, am Ende ihrer Ausbildung und während ihres Berufslebens erfasst werden. Dieses Design erlaubt die Kontrolle sowohl der variierenden Voraussetzungen zu Beginn der Ausbildung, die Kontrolle individueller Determinanten bei der Erfassung des Kompetenzzuwachses während der beruflichen Ausbildung und damit eine in höherem Maße valide Erfassung des Einflusses der beruflichen Ausbildung, sowie die Erfassung der Relevanz der erworbenen Kompetenzen im späteren Berufsalltag.

Ein nicht unerhebliches Problem dieses Designs besteht jedoch im Auftreten von Ausfällen. Nicht alle Personen, die zu Beginn einer Ausbildung befragt wurden, können auch am Ende der Ausbildung wiederbefragt werden. Noch stärker wird dieses Problem bei einer späteren Verfolgung während des Berufslebens auftreten. Ausfälle haben im wesentlichen zwei Konsequenzen. Zum einen verringert sich die Fallzahl, zum anderen kann es zu Verzerrungen bei den Ergebnissen kommen.

Wenn zwischen zwei Messzeitpunkten ein Anteil q ($0 < q < 1$) von Untersuchungseinheiten ausfällt, und zum letzten Messzeitpunkt ein Nettostichprobenumfang von n angestrebt wird, bedeutet dies, dass zum vorletzten Zeitpunkt $n/(1-q)$ Untersuchungseinheiten befragt werden müssen. Generell berechnet sich in einem mehrwelligen Panel mit K Erhebungszeitpunkten und Ausfällen q_k zwischen den Erhebungszeitpunkten k und $k-1$ die notwendige Ausgangsfallzahl n_1 zum ersten Messzeitpunkt bei einer angestrebten Fallzahl n in der letzten (K -ten) Erhebung nach:

$$n_1 = \frac{n}{\prod_{k=2}^K (1 - q_k)} = \frac{n}{(1 - q_2) \cdot (1 - q_3) \cdot \dots \cdot (1 - q_K)}$$

Wie hoch der Ausfall ist, hängt auch davon ab, wie intensiv die Panelpflege betrieben wird, d. h. ob mögliche Adressänderungen der Teilnehmer regelmäßig verfolgt werden, Kontakt mit den Teilnehmern gesucht wird und Anreize geboten werden, dass die Teilnehmer ihre Teilnahmebereitschaft aufrechterhalten und Adressveränderungen möglichst von sich aus melden.

Im GFM-Panel von Heitmeyer u. a.³¹ konnten von den insgesamt 2722 deutschen Befragten ohne Migrationshintergrund, die 2002 erstmals befragt wurden, 2003 noch 1175 und 2004 noch 825 wiederbefragt werden. Die Ausfallquoten betrugen somit 56,8 % zwischen ersten und zweiten und 29,8 % zwischen zweiten und dritten Erhebungszeitpunkt.³² Der Befragungsabstand zwischen den Telefoninterviews betrug ein Jahr. Es gab keine Maßnahmen zur

31 Vgl. zu dieser Studie W. Heitmeyer (Hg.): Deutsche Zustände. Folge 1 ff. Frankfurt a.M.: Suhrkamp 2002 ff.

32 Typischerweise ist der Ausfall zwischen erster und zweiter Befragung am höchsten.

Reduzierung der Ausfallquote. Dem gegenüber wurde in der Wahluntersuchung von Falter u. a. eine sehr intensive Panelpflege betrieben. Von den zur Bundestagswahl 2002 in persönlichen Interviews befragten 3000 Personen liegen nach Auskunft der Projektmitarbeiter z. Z. noch 1600 recherchierte Adressen von teilnahmewilligen Befragten vor, was einer Ausfallquote von 46.7 % entspräche.³³

Wenn in einer Panelumfrage von Teilnehmern beruflicher Ausbildung der erste Erhebungszeitpunkt zu Beginn der Ausbildung, der zweite am Ende der Ausbildung und der dritte drei Jahre nach Beendigung der Ausbildung erfolgen würde, dann wäre bei Befragung in den ausbildenden Einrichtungen oder Betrieben der Ausfall zwischen erster und zweiter Befragung im wesentlichen mit Abbrüchen der Ausbildungen identisch, während zwischen dem Ende der Ausbildung und der Befragung drei Jahre später ein großes Ausfallrisiko bestünde. Um dies zu reduzieren, ist eine intensive Panelpflege notwendig, die u. a. regelmäßige Kontaktaufnahmen und Adressrecherchierungen beinhaltet. Gleichwohl wäre mit einer Ausfallquote von 50 % zu rechnen.

Problematischer als die reine Zahl der Ausfälle, die durch eine entsprechend große Anfangsstichprobe aufgefangen werden kann, ist jedoch die Gefahr von systematischen Ausfällen. Bezogen auf Absolventen beruflicher Ausbildung bedeutet dies, dass eine Fehleinschätzung des Erfolgs beruflicher Ausbildung erfolgen würde, wenn überproportional Befragte ausfallen, die keinen Berufserfolg haben. Das gleiche würde gelten, wenn Personen mit besonders großem Berufserfolg überproportional ausfallen.

Um diese Fehlerquelle einzugrenzen, sind bei einer Panelbefragung zum einen Ausfallanalysen notwendig, die auf der Basis der vorliegenden Daten prüfen, ob systematische Ausfälle vorliegen. Für eine umfassende Kontrolle ist es darüber hinaus notwendig, parallel zu Wiederholungsbefragungen zusätzliche Querschnittsbefragungen aus der gleichen Population zu betrachten.

Wiederholungsbefragungen sind somit deutlich kostenintensiver als einmalige Querschnittsbefragungen, da

- eine größere Ausgangsstichprobe notwendig ist (s. obige Formel zur Fallzahlberechnung),
- Zusatzkosten durch die Panelpflege entstehen
- möglichst zusätzliche Querschnittsbefragungen zur Ausfallkontrolle erhoben werden sollten
- datenanalytischer Mehraufwand für Ausfallanalysen notwendig sind.

Der Kostenaufwand steigt dabei mit der Zahl der Befragungszeitpunkte und dem Abstand zwischen ihnen.

Auf der anderen Seite lassen sich nur bei Wiederholungsbefragungen zuverlässige Aussagen über den individuellen Erfolg beruflicher Bildung treffen.

33 Bei einer tatsächlichen Wiederbefragung ist damit zu rechnen, dass zusätzliche Ausfälle auftreten, so dass eher mit einer Ausfallquote von 50% nach 3 Jahren zu rechnen wäre.

Zu berücksichtigen ist ferner, dass die wesentlichen Untersuchungsergebnisse erst nach der letzten Erhebungswelle vorliegen. Im Rahmen einer international vergleichenden Studie zur beruflichen Ausbildung ist kaum damit zu rechnen, dass der Projektverbund sechs Jahre und länger warten will, bis wirklich zuverlässige Ergebnisse vorliegen.

Vor diesem Hintergrund stellt sich die Frage nach alternativen Erhebungsdesigns. Eine Möglichkeit ist eine rotierende Wiederholungsbefragung. Dabei werden zum Zeitpunkt t_1 in voneinander unabhängigen Stichproben zum einen Personen befragt, die gerade ihre berufliche Ausbildung beginnen, zum zweiten Personen, die gerade am Ende ihrer beruflichen Ausbildung stehen und zum dritten Personen, deren Ausbildungsabschluss bereits solange zurückliegt, wie die Ausbildung dauerte.³⁴ Die Wiederholungsbefragung findet nach einem Zeitraum statt, der der Ausbildungslänge entspricht. Gleichzeitig wird zu diesem Zeitpunkt eine neue Serie von Stichproben gezogen. Ein solches Design erlaubt eine umfassende Analyse des Erfolgs beruflicher Ausbildung in einem relativ kurzen Zeitraum auf der Individualebene. Gleichzeitig ermöglicht es eine Dauerbeobachtung, die z. B. die Folge von Veränderungen des Ausbildungs- wie des Berufssystems ermöglicht.

Voraussetzung für ein rotierendes Panel sind gleich lange Ausbildungszeiten. Ansonsten müssten getrennte Stichproben für die verschiedenen Ausbildungslängen erfolgen. Ein Vergleich des Berufserfolgs über das Ausbildungsende ist dann erst zum Zeitpunkt des kleinsten gemeinsamen Vielfachen der Ausbildungszeiten möglich. Wenn etwa einjährige, zweijährige und dreijährige Ausbildungsgänge A, B und C bestehen, würden zum Zeitpunkt t_1 in A, B und C Befragungen von Personen erfolgen, die ihre Ausbildung beginnen. Gleichzeitig würden in jedem Ausbildungsgang unabhängige Stichproben von Personen erfolgen, die am Ausbildungsende stehen. Weiter würden Personen befragt, die Ausbildungsgang A und B vor 2 und vor 6 Jahren beendet haben, und Ausbildungsgang C vor 3 und vor 6 Jahren beendet haben. Zur Kostensparung wäre es auch denkbar, bei den Absolventen nur solche zu befragen, deren Ausbildung A, B bzw. C sechs Jahre zurückliegt. Nach einem Jahr würden dann diejenigen wiederbefragt, die ihre Ausbildung A ein Jahr zuvor begonnen hatten. Entsprechend würden nach 2 Jahren Personen befragt, die ihre Ausbildung B zwei Jahre zuvor beendet hatten. Schließlich würden nach 3 Jahren Personen wiederbefragt, die drei Jahre zuvor ihre dreijährige Ausbildung begonnen hatten. Entsprechen würden die erstmals am Ende ihrer Ausbildung befragten, nach 2, 3 bzw. 6 Jahren wiederbefragt. Die folgende Abbildung verdeutlicht dieses Erhebungsschema:

³⁴ Das Design kann verlängert werden um zusätzliche Stichproben, deren Ausbildungsende noch länger zurückliegt und Vielfache der Ausbildungszeit beträgt.

Stichprobe	12	3	4	5	6	7	8	9	10	11	12	
Ausbildungsgang	A	B	C	A	B	C	A	B	C	A	B	C
Ausbildungsbeginn	t ₁	t ₁	t ₁									
Ausbildungsende	t ₂	t ₂	t ₂	t ₁	t ₁	t ₁						
1 Jahr nach A.-Ende												
2 Jahre nach A.-Ende	t ₃	t ₃		t ₂	t ₂		t ₁	t ₁				
3 Jahre nach A.-Ende			t ₃			t ₂			t ₁			
4 Jahre nach A.-Ende												
5 Jahre nach A.-Ende												
6 Jahre nach A.-Ende	t ₄	t ₄	t ₄	t ₃	t ₃	t ₃	t ₂	t ₂	t ₂	t ₁	t ₁	t ₁

A =: einjährige Ausbildung, B = zweijährige Ausbildung, C = dreijährige Ausbildung;
t_i =: i-ter Messzeitpunkt; alle Stichproben beginnen zum selben kalendarischen Zeitpunkt, die Abstände zwischen den Messzeitpunkten ergeben sich aus den Angaben in der 1. Spalte.

B. Retrospektivdesign

Eine sehr kostengünstige Alternative zur Panelbefragung bietet ein Retrospektivdesign. In einem Retrospektivdesign werden nur Berufstätige befragt. Die Informationen über die Ausbildung werden ex-post in der Befragung erfasst. Erfasst werden können jedoch nur die gegenwärtigen Berufskompetenzen. Über Gruppenvergleiche lassen sich auf der Aggregatebene Unterschiede bei den Kompetenzen erfassen, wobei eine kausale Interpretation von Erklärungsfaktoren für diese Unterschiede jedoch weitgehend spekulativ wäre.

C. Trendstudiendesign

Anstelle eines Paneldesigns käme ein Trendstudiendesign in Frage. Das Design entspricht dem obigen Paneldesign mit der Ausnahme, dass jeweils anstelle von Wiederholungsbefragungen unabhängige Stichproben mit anderen Befragten erhoben würde. Der Vorteil einer Trendstudie liegt darin, dass es keinen Panelausfall geben kann und daher der Kostenaufwand deutlich geringer liegt. Auf der anderen Seite sind nur Veränderungen auf der Aggregatebene valide erfassbar. Kausale Erklärungen von Veränderungen weisen ähnliche Probleme auf, wie bei Retrospektiverhebungen. Es lassen sich somit sehr gut Unterschiede zwischen Ausbildungsgängen innerhalb und zwischen Ländern untersuchen, weniger gut aber die Ursachen für diese Unterschiede.

D. Ein modulares Erhebungsdesign

Solange der Kostenrahmen für eine international vergleichende Untersuchung der beruflichen Ausbildung nicht fest steht, bietet sich ein modulares Erhe-

bungsdesign an. Die tatsächliche Durchführung kann dann die Module beinhalten, die bei gegebenen Kostenrahmen den Informationsgewinn maximieren.

Ausgangspunkt des Untersuchungsdesigns ist die obige Übersicht des rotierenden Panels. Im einfachsten Fall werde nur dreijährige berufliche Ausbildungen betrachtet, wobei die Ausbildung ausschließlich schulisch, ausschließlich betrieblich, oder kombiniert erfolgen kann.

Modul 1: Absolventenbefragung 6 Jahre nach Beginn der Ausbildung

Ausgangselement wäre dann die Stichprobe 9 (S9a), bei der Befragte erstmals sechs Jahre nach Beginn einer Ausbildung befragt würden. Die Stichprobe würde sich also auf Befragte beziehen, die sechs Jahre zuvor eine dreijährige Ausbildung begonnen hatten. Innerhalb dieser Gruppe sind Vergleiche möglich zwischen Personen mit unterschiedlichen dreijährigen Ausbildungsgängen und zwischen Personen, die eine Ausbildung begonnen, aber nicht abgeschlossen haben.

Durch externe Zusatzinformationen über die Ausbildungsgänge, insbesondere ihre Abbrecherquoten wäre es möglich, ein umfassenderes Bild über den Erfolg der Ausbildungsgänge zu erhalten.

Sinnvoll wäre zudem eine Vergleichsgruppe, die sechs Jahre zuvor hinsichtlich Alter u. möglichst auch schulischer Bildung vergleichbar mit den Befragten der Stichprobe 9 wären, aber ohne Berufsausbildung erwerbstätig sind.

Modul 2: Befragung von Personen in der beruflichen Ausbildung

Ergänzend zum Modul 1 werden in Modul 2 die Stichproben 3 (S3a) und 6 (S6a) gezogen. Bei Stichprobe 3 wäre es optimal, wenn Befragte nach Abschluss eines Ausbildungsvertrags aber vor Beginn der Ausbildung befragt würden.

Das Modul 2 erlaubt eine detailliertere Bewertung des Kompetenzanstiegs durch die berufliche Ausbildung, sowie Rückschlüsse über die Ausgangsvoraussetzungen und den Kompetenzstand vor bzw. am Beginn der Ausbildung sowie am Ende der Ausbildung. Unter der Annahme, dass sich die Kompetenzverteilung vor Beginn der Ausbildung und die Kompetenzvermittlung in der Ausbildung in den drei Jahren zuvor nicht (relevant) geändert hat, ist zudem eine valide Erfassung des Kompetenzanstiegs während der Ausbildung und der Kompetenzveränderung nach Ende der Ausbildung möglich. Darüber hinaus ist unter dieser Annahme eine Bewertung möglich, inwieweit die in der Ausbildung vermittelten Kompetenzen für die Erwerbstätigkeit relevant sind.

Modul 3a: Wiederholungsbefragung nach 3 Jahren

Die Befragten der Stichprobe S3a und S6a werden bei diesem Modul nach 3 Jahren wiederbefragt. Dies ermöglicht auf der Individualebene die Erfassung des Lernerfolgs in der beruflichen Ausbildung (Vergleich von t_2 mit t_1 in

Stichprobe 3) sowie die Nutzung und Veränderung der Kompetenzen in den ersten drei Berufsjahren (Vergleich von t_2 mit t_1 in Stichprobe 6).

Unter der Annahme unsystematischer Panelausfälle kann zudem untersucht werden, ob innerhalb von drei Jahren Veränderungen bei den Kompetenzen am Ende der Ausbildung auftreten.

Wenn Modul 3 realisiert werden kann, sollte bei den Stichproben 3 und 6 eine sorgfältige Panelpflege eingeplant werden. Zu überlegen ist auch die Fallzahlen der Stichproben 3 und 6 von vornherein größer zu bilden, um hinreichend Fälle in der Wiederholungsbefragung zu haben.

Modul 3b: Kontrollstichprobe am Ende der Ausbildung und 3 Jahre nach Ausbildungsende

Ergänzend zu Modul 3a erscheint es mir äußerst sinnvoll, zur gleichen Zeit zusätzliche Stichproben von Befragten zu ziehen, die drei Jahre (Stichprobe 6 neu, S6b) bzw. 6 Jahre (Stichprobe 9 neu, S9b) zuvor eine Ausbildung begonnen hatten bzw. hätten beginnen können.

Mit Hilfe dieser beiden Stichproben kann zum einen ein systematischer Panelausfall in Modul 3b kontrolliert werden. Vor allem ist es aber möglich, Veränderungen in der Ausbildung (Vergleich von Stichprobe S6a und Stichprobe S6b) und Veränderungen im Erwerbsleben (Vergleich von Stichprobe S9a und Stichprobe S9b) zu erfassen.

Modul 3c: Kontrollstichprobe zu Beginn der Ausbildung

Sinnvoll erscheint es auch, ergänzend zu Modul 3a gleichzeitig eine zusätzliche Stichprobe von Personen zu ziehen, die eine Ausbildung beginnen (Stichprobe 3 neu, S3b). Dies ermöglicht eine valide Erfassung der Veränderungen der Voraussetzungen.

Modul 4a: Wiederholungsbefragung nach 6 Jahren

6 Jahre nach Studienbeginn werden die Befragten der Stichproben 3a ein drittes Mal und gegebenenfalls die der Stichproben 6b, und 3b ein zweites Mal befragt. Für die Stichprobe S3a liegt dann ein vollständiger Verlauf vom Ausbildungsbeginn über 6 Jahre vor. Darüber hinaus liegen für zwei Kohorten im Abstand von 3 Jahren Angaben über den individuellen Kompetenzzuwachs während der Ausbildung vor.

Modul 5: Kontrollstichprobe am Ende der Ausbildung und 3 Jahre nach Ausbildungsende

Es empfiehlt sich, nach 6 Jahren neben den Wiederholungsbefragungen weitere Kontrollstichproben 3 (Stichprobe S6c) und 6 (Stichprobe 9c) Jahre nach Ausbildungsbeginn zu ziehen. Dies ermöglicht neben der Kontrolle der

Panelmortalität auch einen Kohortenvergleich über jeweils 3 Kohorten 3 Jahre beim Ausbildungsbeginn, beim Ausbildungsende und 3 Jahre nach Ausbildungsende.

Die folgende Tabelle gibt einen Überblick über das gesamte Erhebungsdesign:

Beginn der Erhebung	+ 3 Jahre	+ 6 Jahre
Modul 1: S9a(6 Jahre nach A.-B.)		
Modul 2: S6a (3 Jahre nach A.-B.)	M 3 a: S6a (W2, 6 Jahre nach A.-B.)	
Modul 2: S3a (Ausbildungsbeginn)	M 3a: S3a (W2, 3 Jahre nach A.-B.)	M 4: S3a (W3, 6 Jahre nach A.-B.)
	M 3b: S9b (6 Jahre nach A.-B.)	
	M3b: S6b (3 Jahre nach A.-B.)	M 4: S6b (W2, 6 Jahre nach A.-B.)
	M3c: S3b (A.-B.)	M 4: S3b (W2, 3 Jahre nach A.-B.)
		M 5: S9c (6 Jahre nach A.-B.)
		M 5: S6c (3Jahre nach A.-B.)

Dieses Design lässt sich problemlos auf einen längeren Zeitpunkt erweitern. Möglich ist zudem den Berufsverlauf länger zu verfolgen, wenn die Befragten der Stichproben S9a nach 3 und 6 Jahren des Untersuchungsbeginns und die Stichprobe S9b 6 Jahre nach Beginn wiederbefragt werden.

E. Rekrutierung

Das Erhebungsdesign ist noch hinsichtlich der Zielpopulation und der Methode der Stichprobenziehung offen. Da das Untersuchungsziel die vergleichende Evaluation beruflicher Ausbildung ist, bietet es sich an, als Zielpopulation Personen auszuwählen, die unterschiedliche Ausbildungsgänge durchlaufen (haben). Angesichts der Vielzahl von Berufsausbildungen allein in der Bundesrepublik empfiehlt es sich, sich auf typische unterschiedliche Berufsfelder zu konzentrieren und die Folgen der Ausbildung bzw. fehlende Ausbildung in diesen Berufsfeldern zu betrachten.

Für meine Überlegungen zum Stichprobenplan der Studie wurde vorgegeben, sich auf die Bereiche Pflege, KFZ-Technik, Informationstechnik und Verwaltung zu konzentrieren und bezogen auf die Bundesrepublik einen Stichprobenplan vorzuschlagen, der die beruflichen Ausbildungen zur Krankenschwester/Krankenpfleger, zum KFZ-Mechatroniker (früher KFZ-Mechaniker, KFZ-Elektriker), zum Informationselektroniker und zur Bürokauffrau/Bürokaufmann berücksichtigt.

Prinzipiell lassen sich die Befragungspersonen über allgemeine Haushaltsstichproben rekrutieren. Dies führt bei entsprechender Auswahl der Haushaltsstichprobe zu einer repräsentativen Stichprobe sowohl innerhalb der vier Ausbildungsberufe als auch zwischen den Berufen und sowohl von Personen zu Beginn, am Ende und auch einer beliebigen vorgegebenen Anzahl von Berufsjahren. Auf der anderen Seite muss, wenn hinreichende Fallzahlen für alle

vier Berufsgruppen erreicht werden sollen, eine sehr große Ausgangsstichprobe gezogen werden, bei der eine Vielzahl von Personen erfasst werden, die nicht zur Zielpopulation gehören und daher nicht weiter befragte werden sollten.

Eine weitere Möglichkeit besteht darin, sich an bereits bestehende Stichproben „anzuhängen“ und Zusatzerhebungen im Rahmen dieser bestehenden Studien durchzuführen. In Frage kommt hier primär das Sozio-ökonomische Panel (SOEP). Es ist allerdings anzunehmen, dass die Fallzahlen in den interessierenden Gruppen zu gering sind. Dies ist noch zu überprüfen. Eine Repräsentativität auf Personenebene lässt sich vermutlich nicht erreichen. Denkbar wäre auch ein Zugriff über die großen Access-Panel der Befragungsinstitute.

Ein weiterer Nachteil der Rekrutierung über allgemeine Haushaltsstichproben besteht darin, dass die Befragten Bezug zu einer Vielzahl von unterschiedlichen Institutionen und Betrieben haben. Die Erfassung des institutionellen Kontextes ist für die Forschungsfragen von großer Bedeutung. Daher müssen neben den hier betrachteten Teilnehmern und Absolventen beruflicher Bildung auch Informationen über die Ausbildungsstätten und Betriebe erhoben werden. Bei Haushaltsstichproben werden in der Befragung die Adressen der Institutionen erhoben und diese in eine Einrichtungsstichprobe zusammengeführt. Es ist damit zu rechnen, dass die Angaben der befragten Personen allein aus Erinnerungsgründen über die Einrichtungen unvollständig und fehlerhaft sind und darüber hinaus die Zahl der Einrichtungen möglicherweise die Zahl der befragten Personen sogar übersteigt. Der hohe Aufwand, der hierdurch entsteht, führt dazu, dass Haushaltsstichproben nach meiner Auffassung für die Vergleichsstudie zur beruflichen Ausbildung nicht in Frage kommen.

Von daher bietet sich ein (mindestens) zweistufiger Stichprobenplan an, bei dem auf der ersten Stufe nicht Haushalte, sondern Ausbildungsstätten bzw. Betriebe erhoben werden und auf der zweiten Ebene Teilnehmer an Ausbildungsgängen bzw. Betriebsangehörige. Wenn die Einrichtungen zufällig ausgewählt werden, ist auch dieser Stichprobenplan repräsentativ für die Angehörigen der Einrichtungen.

Nicht berücksichtigt werden bei einem solchen Stichprobenplan Personen der Zielpopulation, die keiner Einrichtung angehören. Für die Fragestellung der Studie relevant sind allein solche Personen, die in der Vergangenheit eine Ausbildung für die vier betrachteten Berufsgruppen begonnen oder abgeschlossen haben und unfreiwillig nicht mehr erwerbstätig sind. Ergänzend zur Befragung der Teilnehmer, Abbrecher und Absolventen von beruflicher Bildung, die über die Ausbildungsstätte bzw. Betriebe erfasst werden, wäre es daher angeraten, vergleichend Arbeitslosigkeitsrisiken von Personen dieser Berufsgruppen zu analysieren.

Für die Rekrutierung der Untersuchungspersonen bietet es sich somit an, bei Befragung von Teilnehmern beruflicher Ausbildungsprogramme, die von Schulen angeboten werden, eine Zufallsstichprobe der entsprechenden Schulen zu ziehen, und/oder entsprechende Zufallsstichproben von Betrieben, in denen die Absolventen dieser Ausbildungen üblicherweise arbeiten, ein-

schließlich von Kleingewerbebetrieben bzw. Selbständigen, bei denen der Inhaber in der Regel eine entsprechende Ausbildung aufweist.

Für Teilnehmer von Ausbildungen im bundesdeutschen dualen Ausbildungssystem ist zu überlegen, ob die Befragten eher in Betrieben oder in Schulen rekrutiert werden sollten. Der Vorteil von Schulen besteht darin, aufgrund der relativ hohen Klassenstärken schnell eine größere Zahl von Personen zu befragen. Denkbar ist allerdings, dass die Schulen nicht sehr willig sind, Unterrichtszeit für die Befragungen zu opfern. Möglicherweise können die Befragungen im Anschluss an den regulären Unterricht innerhalb der Klassenräume erfolgen. Vermutlich ist in dieser Situation ein (z. B. finanzieller) Anreiz zur Teilnahme an der Untersuchung notwendig, da sonst die Teilnahmebereitschaft gering sein dürfte und insbesondere „schwache“ Schüler(innen) eher nicht befragt werden.

Wenn die Erhebung von Berufskompetenzen eine Infrastruktur voraussetzt, die nur in den auszubildenden Betrieben vorhanden ist, muss die Befragung in diesen Betrieben erfolgen. Auch hier stellt sich die Frage einer finanziellen Kompensation, möglicherweise allerdings nicht zugunsten der Befragten, sondern zugunsten der Betriebe.

Bei der Befragung in Betrieben kann die Rekrutierung sowohl über die Schulen als auch die Betriebe erfolgen. Eine Stichprobenziehung über die Betriebe erleichtert die Erfassung von früheren Abbrechern von Ausbildungsgängen sowie von Betriebsangehörigen mit vergleichbaren Tätigkeiten, aber ohne Ausbildung. Daher scheint es mir ratsam, zu versuchen, Untersuchungspersonen über Betriebe zu rekrutieren.

Bei der Auswahl der Betriebe ist zu überlegen, eine Teilstichprobe aus dem Betriebspanel des IAB zu ziehen. Für kleinere Betriebe und Gewerbetreibende, die nicht erfasst werden, sollten über Kammern u.ä. teilnehmende Betriebe gewonnen werden.

F. Stratifizierung

Bei der Auswahl von Betrieben dürfte eine Stratifizierung entsprechend den Vorgaben des IAB-Betriebspanels sinnvoll sein. Überlegenswert ist es, innerhalb der Strata die Auswahlwahrscheinlichkeiten so zu setzen, dass in jeder Schicht eine gleiche Zahl von Betrieben erfasst wird. Wenn kein Ausgleich bei der Ziehung auf der zweiten Ebene erfolgt, würde dies bedeuten, dass Angehörige größerer Betriebe eine größere Chance haben, in die Stichprobe aufgenommen zu werden.

G. Stichprobenumfang

Aussagen über den notwendigen Stichprobenumfang sind ohne Informationen über die verwendete Analyseverfahren, die Variation der betrachteten Merkmale in der Population und die angestrebte Genauigkeit kaum zu treffen.

Als grobe Faustregel kann der notwendige Stichprobenumfang eines dichotomen Merkmals mit maximaler Varianz bei einer Irrtumswahrscheinlichkeit von 5 % und einer Fehlermarge von ebenfalls 5 % herangezogen werden. Bei einer einfachen Zufallsauswahl führt dies zu einer Fallzahl von ca. $n=400$ Fällen, bei einer Genauigkeit von 3 % auf ca. $n=1100$ Fälle. Generell berechnet sich die Fallzahl bei einfachen Zufallsauswahlen nach:³⁵

$$n = \frac{(z_{1-\alpha/2})^2 \cdot s^2}{\varepsilon^2}$$

wobei $z_{1-\alpha/2}$ das $(1-\alpha/2)$ -Quantil der Standardnormalverteilung, α die Irrtumswahrscheinlichkeit, s die Standardabweichung der interessierenden Größe in der Population und ε die maximal akzeptierte Fehlermarge ist.

Bei einer Clusterstichprobe ist die Intraklassenkorrelation zu berücksichtigen, die sich dadurch ergibt, dass sich die Befragten in einem Cluster (hier: Betrieb bzw. Ausbildungsstätte) ähnlicher sind als Befragte aus verschiedenen Clustern. Als Faustformel für die Berechnung der Fallzahl kann die Formel:³⁶

$$1 + (M - 1) \cdot ICC$$

verwendet werden, wobei ICC die Intraklassenkorrelation und M die Clustergröße bezeichnet. Die Formel gibt an, um welchen Faktor die Stichprobengröße erhöht werden muss, wenn in einer ausgewählten Einrichtung alle in Frage kommenden Personen befragt werden

Wenn also z. B. die durchschnittliche Intraklassenkorrelation innerhalb der Schulklassen 0.1 beträgt, die Klassenstärke $M=20$ Schüler(innen) beträgt und klassenweise alle Klassenmitglieder befragt werden, ergibt sich ein Faktor von $1+19 \cdot 0.1 = 2.9$. Anstelle von $n=400$ Fällen bzw. $n=1100$ Fällen bei einer einfachen Zufallsauswahl müssten dann $n=1160$ bzw. $n=3190$ Schüler(innen) befragt werden, was zur Auswahl von $N=55$ bzw. $N=160$ Klassen führt.

Der Stichprobenplan sieht bei Verwendung der Module 3 und 4 Wiederholungsbefragungen vor. Zur Schätzung der Panelmortalität innerhalb der Ausbildung kann die durchschnittliche Abbrecherquote in einem Ausbildungsgang herangezogen werden. Wenn diese z. B. 20 % beträgt, muss die Stichprobenzahl bei einer ausschließlichen Querschnittsbefragung um den Faktor 1.25 erhöht werden, statt 400 wären also 500 Personen statt 1100 wären 1375 Personen zu befragen.

35 Sharon L. Lohr: Sampling: Design and Analysis. Bonn u. a.: Duxbury Press, 1999: S. 40.

36 ebda, S. 139f.

Unter den Befragten, die drei Jahre nach Ende der Berufsausbildung wiederbefragt werden, ist auch bei sorgfältiger Panelpflege mit einem ein Ausfall von 50 % zu rechnen (s. o.). Entsprechend sollte die Fallzahl der am Ende der Ausbildung erstmalig befragten Absolventen doppelt so hoch sein, wie die benötigte Zahl, also statt $n=400$ bzw. 1100 $n=800$ bzw. 2200 Fälle umfassen.

Um verallgemeinerungsfähige Aussagen in jedem der vier Ausbildungsbe-
reiche zu ermöglichen, sollte für jedes Berufsfeld eine eigene Stichprobe ge-
zogen werden, wobei eine Fehlermarge von 5 % akzeptabel erscheint. In jeder
Gruppe wären bei einer einfachen Zufallsauswahl 400 Personen zu befragen,
insgesamt also 1600 Personen.

In Modul 1 wird eine Querschnittsbefragung von Personen 6 Jahre nach
Beginn der üblichen Ausbildungszeiten vorgeschlagen. In die Stichprobe soll-
ten alle Betriebsangehörige aufgenommen werden, die in etwa der gleichen
Altergruppe angehören und gleiche oder ähnliche Tätigkeiten ausüben. Aus
Sekundärdaten ist zu recherchieren, wie hoch der Anteil der Betriebsangehöri-
gen ist, die keine abgeschlossene berufliche Ausbildung haben oder aus einem
anderen Beruf kommen. Wenn dieser Anteil z. B. 25 % beträgt, sollte die
Ausgangsfallzahl um den Faktor 100 % plus dem Anteil erhöht werden, im
Beispiel also um den Faktor 1.25, was zu einer Fallzahl von $n=500$ pro Be-
rufsguppe führt.

Im Unterschied zu Schulklassen ist hier mit einer eher geringen Intraklas-
senkorrelation zu rechnen. Über die verschiedenen Betriebsgrößen hinweg
mag der Wert bei vielleicht 0.02 liegen. Für die Berechnung des Korrektur-
faktors ist zu recherchieren, wie viele Betriebsangehörige in (gewichteten)
Durchschnitt über alle Betriebe dem Tätigkeitsbereich der entsprechenden Al-
tersgruppe angehören. Wenn dieser Wert z. B. 5 beträgt, würde sich ein Kor-
rekturfaktor von $1+(5-1) \cdot 0.02 = 1.08$ ergeben. Die Fallzahl würde sich dann
von $n=500$ auf $n=540$ zu befragende Personen in jedem der vier Tätigkeitsbe-
reiche erhöhen.

Die Anzahl der auszuwählenden Betriebe richtet sich danach, wie viele
Zielpersonen im Mittel in einem Betrieb vorgegebener Größe arbeiten. Wenn
auch diese Zahl über alle Betriebsgrößenklassen 5 beträgt wären in jedem Be-
rufsfeld etwa 108 Betriebe auszuwählen. Über die vier Gruppen hinweg ergibt
sich für die Stichprobe S9a im Modul 1 eine Fallzahl von $N=432$ Betrieben
mit insgesamt $n=2160$ Befragten.

Für die Berechnung der Umfänge der Stichproben S3a und S6a kommt es
darauf an, ob die Rekrutierung über Betriebe oder Schulen erfolgt. Soll die
Rekrutierung über Betriebe erfolgen, ist ein analoges Vorgehen wie bei der
Berechnung der Fallzahl für S9a vorzusehen. Fraglich ist hier vor allem, wie
viele Personen am Beginn und wie viele am Ende ihrer Ausbildungszeiten im
Durchschnitt in den Betrieben beschäftigt sind. Wenn die Zahlen wieder 5 pro
Betrieb sind und der ICC 0.02 beträgt, würden sich hier in jedem Ausbil-
dungsgang Stichprobenumfänge von $n=432$ Personen am Beginn und von
weiteren $n=432$ Personen am Ende ihrer Ausbildung ergeben.

Anders sieht es aus, wenn die Rekrutierung über Schulen erfolgt, wie dies bei der Berufsgruppe der Krankenpfleger(innen) der Fall ist. Hier ist mit einer höheren Besetzungszahl in den Clustern und mit einer höheren Intraklassenkorrelation zu rechnen. Wenn die Annahmen von oben zutreffen ($ICC=0.1$, im Durchschnitt $m=20$ Personen pro Klasse) sind, dann beträgt die Befragtenzahl jeweils $n=1160$ Personen in $N=58$ Klassen zu Beginn und am Ende der Ausbildung.

Insgesamt ergeben sich somit für die Stichprobe S3a und S9a jeweils 2456 Befragte, die aus 260 Betrieben und 58 Schulklassen rekrutiert werden.

Sollen die Stichproben S3a und S6a für eine Wiederholungsbefragung nach drei Jahren aufgestockt werden, würde sich die Fallzahl bei S3a bei einer angenommenen Abbrecherquote von 20 % auf $n=3070$ Befragte am Beginn ihrer Ausbildung und bei einem Panelausfall von 50 % nach Beendigung der Ausbildung auf $n=4912$ Befragte am Ende ihrer Ausbildung erhöhen.

Unter den getroffenen Annahmen ergeben sich insgesamt für das vollständige Erhebungsdesign mit allen Modulen folgende Fallzahlen:

Geschätzte Fallzahlen in den Teilstichproben

Beginn der Erhebung	+ 3 Jahre	+ 6 Jahre
S9a(6 Jahre nach A.-B.) $n=2160$		
(3 Jahre nach A.-B.) $n=4912/2456$	(W2, 6 Jahre nach A.-B.) $n=2456$	
(Ausbildungsbeginn) $n=3070/2456$	(W2, 3 Jahre nach A.-B.) $n=2456$	(W3, 6 Jahre n. A.-B.) $n=1228$
	S9b (6 Jahre nach A.-B.) $n=2160$	
	(3 Jahre nach (A.-B.) $n=2456$	(W2, 6 Jahre nach A.-B.) $n=1228$
	S3b (A.-B.) $n=2456$	(W2, 3 Jahre nach A.-B.) $n=1228$
		(6 Jahre nach A.-B.) $n=2160$
		(3Jahre nach A.-B.) $n=2456$

Bei zwei Fallzahlen: 1. Zahl mit Ausgleich für Panelausfall, 2. Zahl ohne Ausgleich

H. Interviewkosten

Die Interviewkosten sind ohne zusätzliche Informationen nur schwer zu kalkulieren. Geht man davon aus, dass zur Erfassung der benötigten Größen bei allen Befragten 4 Zeitstunden notwendig sind, ist dies kaum in einem Interview zu schaffen. Es müssten daher zwei Interviews a 2 Stunden pro Befragungsperson und Messzeitpunkt berücksichtigt werden. Pro Befragten müssten etwa 120.-- € reine Interviewkosten kalkuliert werden. Hinzu kommen Kosten fixe Kosten für Pretests, Erstellung der Erhebungsinstrumente, Datenaufbereitung, Dokumentation und Leitung der Erhebung in Höhe von ca. 75.000.-- € pro Stichprobe.³⁷ Bezogen auf die Module ergeben sich folgende Schätzungen:

37 Diese Angaben sind grobe Schätzungen basierend auf Angeboten professioneller Umfrageinstitute für „typische“ Umfragen von ca. 1 Stunde Dauer und 2000-3000 Befragten.

Geschätzte Erhebungskosten in den Teilstichproben

Beginn der Erhebung	+ 3 Jahre	+ 6 Jahre
n=2160, 334.200,-- €		
a) n=4912, 664.440,-- €	n=2456, 369.720,-- €	
b) n=2456, 369.720,-- €		
a) n=3070, 443.400,-- €	n=2456, 369.720,-- €	n= 1228, 222.360,-- €
b) n=2456, 369.720, -- €		
	n=2160, 334.200,-- €	
	n=2456, 369.720,-- €	n=1228, 222.360,-- €
	n=2456, 369.720,-- €	n=1228, 222.360,-- €
		n=2160, 334.200,-- €
		n=2456, 369.720,-- €
1.073.640,-- bis 1.442.040,-- €	1.813.080,-- €	1.371.000,-- €

Zu diesen Kosten hinzukommen würden noch Kosten für die Stichprobenziehung. Deren Höhe hängt davon ab, ob es Verzeichnisse der Einrichtungen gibt, die in die Stichprobe aufgenommen werden und welche Teilnahmebereitschaft die Einrichtungen zeigen.

Nicht berücksichtigt sind Kosten, die die Teilnahmebereitschaft erhöhen können. Hierzu gehört die Panelpflege bei Wiederholungsbefragungen sowie evtl. notwendige Kosten wie Entlohnungen für Befragten bzw. Betriebe.

I. Einschränkungen der Aussagekraft

Das vorgeschlagene Design erlaubt eine für Deutschland repräsentative Analyse beruflicher Ausbildung in den vier Tätigkeitsfeldern. Ohne eine zusätzliche Analyse von Arbeitslosenrisiken kann sich die Aussagekraft nur auf Personen beziehen, die in einem der vier Tätigkeitsfelder erwerbstätig sind. bzw. bei Auszubildenden Erwerbstätigkeit anstreben.

Darüber hinaus ist mit diesem Design nur eingeschränkt festzustellen, wie der Berufserfolg derjenigen aussieht, die nach ihrer Ausbildung den Tätigkeitsbereich wechseln, soweit sie nicht in einen der den anderen drei Tätigkeitsfelder wechseln. Erfasst werden nur Bereichswechsel von Personen aus den Stichproben S3a/b und S6a/b. Diese Einschränkung lässt sich durch eine Erweiterung auf weitere Felder überwinden.

Expertise Prof. Dr. Jürgen Rost**Zum Einsatz der Item-Response-Theorie für die Messung
berufsbezogener Kompetenzen im Rahmen der Studie
„Berufsbildungs-PISA“**

Es liegt die Machbarkeitsstudie für ein „Berufsbildungs-PISA“ einschließlich diverser Anhänge vor, insbesondere die Expertise von Steffen Kühnel zu alternativen Erhebungsdesigns. In der Expertise von Kühnel wird die in der Machbarkeitsstudie aufgestellte Alternative zwischen einem Querschnittsdesign und einer Panelstudie um zahlreiche Alternativvorschläge erweitert. Die Alternativdesigns sind sehr wohl begründet, so dass ich davon ausgehe, dass irgendein Mischdesign zwischen Querschnitt- und Längsschnittstudie, z. B. die von Steffen Kühnel unter Punkt A seiner Expertise genannte rotierende Wiederholungsbefragung zur Realisierung gelangen wird. Die folgenden Überlegungen zum Einsatz der Item-Response-Theorie im Rahmen einer Berufsbildungs-PISA-Studie sind jedoch unabhängig von der Auswahl eines bestimmten Erhebungsdesigns.

1. Zwei alternative Analysemethoden

Als Ergebnis der Kompetenzmessung sollten *Messwerte* zur Verfügung stehen, die – je nach dem zugrunde gelegten Kompetenzstrukturmodell – Aussagen über die erreichten Kompetenzstufen oder den qualitativen Ausprägungsgrad der jeweiligen Kompetenz erlauben. Die Messwerte sollten eine kriteriumsorientierte Interpretation haben, d. h. über die Normorientierung in Form von Mittelwertsabweichungen hinausgehen.

Die Lösungswahrscheinlichkeiten können zum einen durch relative Häufigkeiten geschätzt werden, ein Verfahren, das häufig als klassische Testtheorie bezeichnet wird. Oder die Lösungswahrscheinlichkeiten der Testaufgaben werden zum Gegenstand einer Modellierung gemacht, die über die Anpassung eines geeigneten Item-Response-Modells zur Ableitung von theoretisch erwarteten Lösungswahrscheinlichkeiten führt. Das ist der Ansatz der Item-Response-Theorie.

Die Schätzung von Lösungswahrscheinlichkeiten durch relative Häufigkeiten ist ein sehr direkter und anschaulicher Weg zu kriteriumsorientierten Aussagen zu gelangen. Auch kann man über die durch relative Lösungshäufigkeiten geschätzten Lösungswahrscheinlichkeiten zu kriteriumsorientierten Messwerten gelangen (z. B. über die mittlere Lösungshäufigkeit aller Testaufgaben). Dieser Weg ist jedoch nur unter der Voraussetzung gangbar, dass alle Personen dieselben Testaufgaben bearbeiten.

Die Methode der Item-Response-Theorie besteht darin, das Antwortverhalten der Respondenten durch ein wahrscheinlichkeitstheoretisches Modell zu modellieren und die bei Geltung dieses Modells zu erwartenden Lösungswahrscheinlichkeiten zur Grundlage von kriteriumsorientierten Aussagen zu machen. Dieser Umweg über eine Modellierung der Antwortwahrscheinlichkeiten hat insbesondere dann Vorteile, wenn

- es darum geht einen Test neu zu entwickeln,
- man mit größeren Mengen an ausfallenden Daten (missing-data) zu rechnen hat,
- Veränderungsmessung, also die wiederholte Testung derselben Personen ein Bestandteil des Designs ist,
- oder Aussagen über Teilkompetenzen anhand von Untergruppen der Testaufgaben abgeleitet werden sollen.

Diese Vorteile ergeben sich daraus, dass über die Item-Parameter, die im Rahmen des Ansatzes der Item-Response-Theorie geschätzt werden, Testteile sowohl zwischen Personenstichproben als auch zwischen Item-Mengen und zwischen Testzeitpunkten verlinkt werden können. Eine solche Verlinkung ist mit den relativen Häufigkeiten als Lösungswahrscheinlichkeiten nicht möglich.

2. Die Item Response Theorie im Projekt einer Berufsbildungs-PISA-Studie

Es ist davon auszugehen, dass es im Rahmen eines komplexen Untersuchungsdesigns auch Anteile einer *Messwiederholung*, d. h. der Vorgabe desselben Tests an dieselben Personen geben wird. Bei der Auswertung derartiger Längsschnittdaten steht zunächst die Klärung an, inwieweit strukturelle oder quantitative Veränderungen auf der Ebene des Gesamttests den Kompetenzzuwachs bestimmen. Diese Fragen lassen sich am besten im Rahmen der Item-Response-Theorie klären. Auch die Berechnung von Differenzwerten und die Abschätzung der Reliabilität der Tests ist auf dem Hintergrund der IRT wesentlich genauer möglich.

Die *Testinstrumente* zur Kompetenzmessung in verschiedenen Bereichen stehen noch zur Entwicklung an und werden diversen Präpilotierungen und Pilotierungen zu unterwerfen sein, zumal Einigkeit darüber besteht, dass Arbeitsproben in vielen Bereichen die valideren Ergebnisse versprechen. Es geht also darum, Itemformate zu entwickeln, die aussagekräftig genug sind, um die Ergebnisse von Arbeitsproben ersetzen oder gar simulieren zu können. Vorerfahrungen bestehen mit Essays als Antwortformat, die sich nach verschiedenen Aspekten kodieren lassen.

Aus dieser Situation ergibt sich, dass das in der Pilotierungsphase eingesetzte Testmodell Ergebnisse liefern sollte, die etwas über die mögliche mehrdimensionale Struktur aussagen. Das betreffende Modell sollte also heuristi-

sche Qualitäten haben. Das ist beim normalen Rasch-Modell und auch seinen mehrdimensionalen Verallgemeinerungen nicht der Fall. Die mehrdimensionale Struktur muss bei der Anwendung des mehrdimensionalen Rasch-Modells (anders als bei der Faktorenanalyse) a priori spezifiziert sein. Zum anderen sollte das Modell erlauben, den Test in Richtung auf die (für jede Dimension getrennte) Anwendung des Rasch-Modells hin zu optimieren.

Ein solches Modell stellt das mixed Rasch-Modell dar, das auf der Annahme beruht, dass das Rasch-Modell – wenn schon nicht in der gesamten untersuchten Population – so doch in latenten Subpopulationen auf die Daten passt. Das Programm sucht quasi nach Teilpopulationen, in denen das Rasch-Modell gilt, wobei es keine weiteren Indikatoren für diese Subpopulationen gibt. Es handelt sich um sog. ‚latente Klassen‘.

Das heuristische Element bei der Anwendung dieses Modells besteht darin, dass man sieht in welchen Teilpopulationen eine Aufgabe relativ leichter oder schwerer ist als in der übrigen Stichprobe. Auch Mehrdimensionalität bildet sich in diesem Modell ab, so dass anhand der Ergebnisse einer Datenanalyse mit dem mixed Rasch-Modell

- Items selektiert werden können, um den Test in Richtung auf Eindimensionalität zu optimieren
- Items so gruppiert werden können, dass sie jeweils eine eigene Dimension erfassen
- Personengruppen identifiziert werden, deren Kompetenzen nicht mit dem Rasch-Modell skaliert werden können
- Personengruppen unterschieden werden können, in denen die gemessene Kompetenz mit unterschiedlichen Lösungsprofilen definiert wird.

Das mixed Rasch-Modell stellt ein großes Informationsangebot über den analysierten Test und die getesteten Personen zur Verfügung, das in der Phase der Testentwicklung zur Optimierung des Tests in Richtung auf Rasch-Skalierbarkeit genutzt werden kann.

Bei einem derart komplexen Unternehmen wie der Erfassung von berufsbezogenen Kompetenzen wird es selbst im Rahmen eines einfachen Querschnittsdesigns darum gehen, die Aufgabenschwierigkeiten in Abhängigkeit von verschiedenen *Einflussfaktoren* zu analysieren. Hierfür eignen sich die durch relative Häufigkeiten geschätzten Lösungswahrscheinlichkeiten nicht, da sie als Wahrscheinlichkeitswerte nicht einer linearen Dekomposition zugänglich sind. Im Rahmen der Item-Response-Theorie können derartige Effekte auf die Aufgabenschwierigkeiten durch eine weitere lineare Modellierung der logistischen Item-Parameter erfasst werden.

3. Empfehlung der Rasch-Messtheorie

Die genannten Argumente für eine Verwendung der Item-Response-Theorie (Längsschnittdaten, Testentwicklung, Verlinkung von Teilstichproben, Effekte

auf die Aufgabenschwierigkeiten) sprechen dafür, eine Auswertung der Kompetenzdaten mit Hilfe der Item-Response-Theorie zu präferieren. Unter den zur Wahl stehenden Alternativen (Birnbaum-Modell, Mokken-Modell oder Rasch-Modell) würde ich Verwendung des Rasch-Modells empfehlen, da es als parametrisches Messmodell am ehesten geeignet ist, die Abhängigkeit der Lösungswahrscheinlichkeiten von der zu messenden Kompetenz zu modellieren.

Dabei würde ich empfehlen, die Komplexität der Datenauswertung im Vergleich zu den beiden PISA-Studien 2000 und 2003 dadurch zu reduzieren, dass kein Multi-Matrix-Design angewendet wird und der Stichprobenplan auch ohne Quotierung größtmögliche Repräsentativität sicherstellt. Unter diesen Voraussetzungen kann auf die Verwendung des Programms CONQUEST verzichtet werden (das eine relativ aufwendige Einarbeitung erfordert) und z. B. mit dem Programm WINMIRA gearbeitet werden, das eine wesentlich geringere Einarbeitungszeit erfordert. Ich empfehle eine hinreichend große Aufgabenanzahl zur Messung jeder Kompetenz vorzusehen, so dass auf die plausible-values-Technik verzichtet und mit normalen WLE-Schätzern gearbeitet werden kann. Auf die Verwendung sogenannter Hintergrundmodelle kann verzichtet werden, wenn zusätzlich zu den WLE-Schätzern die Messgenauigkeit und Reliabilität der Tests erfasst wird, so dass messfehlerbedingt überschätzte Varianzen adjustiert werden können. Insbesondere empfehle ich das Rasch-Modell schon bei der Testentwicklung einzusetzen (siehe oben), da auch bei kleinen Stichproben die vom Modell angebotenen Fitmaße und Möglichkeiten der Kontrolle der Eindimensionalität genutzt werden können.

Jürgen Rost

Das Zusammenwachsen des europäischen Güter- und Arbeitsmarktes stellt die Bildungs- und insbesondere die Berufsbildungssysteme vor eine Belastungsprobe: Sie müssen die qualifikatorischen Voraussetzungen dafür schaffen, dass die Arbeitskräfte den Herausforderungen grenzüberschreitender beruflicher Mobilität und internationaler Kooperation gewachsen sind. Wieweit sie diese Kompetenzen vermitteln, weiß heute niemand.

Deshalb hat die deutsche Regierung (BMWA) die Initiative ergriffen und eine Machbarkeits-Studie für ein Berufsbildungs-PISA – analog zum Schul-PISA – ausgeschrieben. Das hier vorgelegte, mit führenden internationalen Experten bera-

tene Konzept eines Berufsbildungs-PISA zeigt auf, wie Kompetenzmessung im Bereich der Berufsbildung möglich ist, welche Kriterien zur Beurteilung von Ausbildungsprogrammen herangezogen werden müssen und wie sich die Leistungsfähigkeit verschiedener Lernorte einschätzen lässt.

Die Studie arbeitet Bedingungen und Prozesse für eine Verbesserung der beruflichen Bildung in den beteiligten Ländern aus. Damit leistet sie einen zentralen Beitrag für die politische Auseinandersetzung, aber auch für die wissenschaftliche Diskussion – beispielsweise über die Herstellung eines europäischen Berufsbildungsraums.

www.steiner-verlag.de

Franz Steiner Verlag

ISBN-10: 3-515-08947-0
ISBN-13: 978-3-515-08947-0

