

177

Kompetenzmessung in der Bildungsforschung

Zusammenfassung des Forums (7) der
5. Konferenz für Sozial- und Wirtschafts-
daten

Beatrice Rammstedt

Mai 2011

Working Paper Series des Rates für Sozial- und Wirtschaftsdaten (RatSWD)

Die *RatSWD Working Papers* Reihe startete Ende 2007. Seit 2009 werden in dieser Publikationsreihe nur noch konzeptionelle und historische Arbeiten, die sich mit der Gestaltung der statistischen Infrastruktur und der Forschungsinfrastruktur in den Sozial-, Verhaltens- und Wirtschaftswissenschaften beschäftigen, publiziert. Dies sind insbesondere Papiere zur Gestaltung der Amtlichen Statistik, der Ressortforschung und der akademisch getragenen Forschungsinfrastruktur sowie Beiträge, die Arbeit des RatSWD selbst betreffend. Auch Papiere, die sich auf die oben genannten Bereiche außerhalb Deutschlands und auf supranationale Aspekte beziehen, sind besonders willkommen.

RatSWD Working Papers sind nicht-exklusiv, d. h. einer Veröffentlichung an anderen Orten steht nichts im Wege. Alle Arbeiten können und sollen auch in fachlich, institutionell und örtlich spezialisierten Reihen erscheinen. Die *RatSWD Working Papers* können nicht über den Buchhandel, sondern nur online über den RatSWD bezogen werden.

Um nicht deutsch sprechenden Nutzer/innen die Arbeit mit der neuen Reihe zu erleichtern, sind auf den englischen Internetseiten der *RatSWD Working Papers* nur die englischsprachigen Papers zu finden, auf den deutschen Seiten werden alle Nummern der Reihe chronologisch geordnet aufgelistet.

Einige ursprünglich in der *RatSWD Working Papers* Reihe erschienen empirischen Forschungsarbeiten, sind ab 2009 in der RatSWD Research Notes Reihe zu finden.

Die Inhalte der *RatSWD Working Papers* stellen ausdrücklich die Meinung der jeweiligen Autor/innen dar und nicht die des RatSWD.

Herausgeber der RatSWD Working Paper Series:

Vorsitzender des RatSWD (2007/2008 Heike Solga; seit 2009 Gert G. Wagner)

Geschäftsführer des RatSWD (Denis Huschka)

Kompetenzmessung in der Bildungsforschung

Zusammenfassung des Forums (7) der 5. Konferenz für Sozial- und Wirtschaftsdaten

Beatrice Rammstedt

In modernen Wissensgesellschaften ist Bildung die zentrale Voraussetzung für wirtschaftliches Wachstum und Wohlstand. Umgekehrt ist auch für das Individuum Bildung entscheidend für die erfolgreiche Teilhabe an der Gesellschaft und für die Bewältigung neuer Anforderungen in einer sich rasch wandelnden, globalisierten Welt.

In den letzten Jahren basiert die Bildungsforschung zunehmend auf Daten umfangreicher und bevölkerungsrepräsentativer Umfragedaten, die somit Rückschlüsse auf die Verteilung der Bildung in den zugrundeliegenden Populationen erlauben. Um diese Verteilungen besser interpretieren und kontrastieren zu können, sind diese large-scale Bildungsforschungsstudien meist entweder in internationale Vergleichsstudien eingebettet, die somit einen kulturellen Vergleich ermöglichen, oder als Längsschnittstudien aufgesetzt mit dem Ziel, Veränderungen über Zeit und Auswirkungen von bildungspolitischen Weichenstellungen zu identifizieren.

Bildung wird in diesem Kontext meist nicht nur operationalisiert über die individuelle Ausbildung, sondern zunehmend auch über objektive, direkt erfasste Kompetenzmaße. Kompetenz wird hier sowohl als Prädiktor z.B. für gesellschaftlichen Erfolg wie aber auch als Outcome z.B. von erfolgreichen Aus- und Weiterbildungsmaßnahmen betrachtet.

Das Forum Kompetenzmessung in der Bildungsforschung im Rahmen der fünften Konferenz für Sozial- und Wirtschaftsdaten hatte zum Ziel, die verschiedenen aktuellen Studienprogramme zu Kompetenzmessung in der large-scale Bildungsforschung aus nationaler Perspektive vergleichend darzustellen. Fokus lag hierbei auf einer Darstellung des zugrundeliegenden Designs einerseits wie auch der resultierenden Daten und ihrer Zugänglichkeit andererseits. So wurden im Rahmen des Forums folgende Studienprogramme vorgestellt:

1. **Das Nationale Bildungspanel (NEPS)**, welches als nationale Längsschnittstudie im Jahr 2005 startete. Frau Dr. Fellenberg, Herr Prof. Dr. Blossfeld und Frau Dr. von Maurice stellten in ihrem Beitrag die Ziele und Herausforderungen dieses äußerst umfassenden längsschnittlichen Studienprogramms im Bereich der Bildungsforschung dar.

2. **PISA 2009.** Frau Dr. Jude und Herr Prof. Dr. Klieme vom Deutschen Institut für Internationale Pädagogische Forschung (DIPF) in Frankfurt am Main stellten in ihrem Beitrag die Ergebnisse von PISA 2009 aus nationaler Sicht vor und berichten erste Trends in der Kompetenzentwicklung sowie Kontextfaktoren basierend auf den Ergebnissen der vergangenen fünf PISA-Wellen.
3. ***The Programme for the International Assessment of Adult Competencies (PIAAC).*** Frau PD Dr. Rammstedts Vortrag zu PIAAC befasste sich - im Gegensatz zu PISA - mit einer sehr jungen Studie, zu der erst 2013 Ergebnisse vorliegen werden. Frau Rammstedt beschrieb die Ziele von PIAAC, das Studiendesign, die Herausforderungen und die Anforderungen an interkulturelle Vergleichbarkeit in PIAAC.
4. ***Die nationalen Bildungsstandards.*** Schließlich stellten Herr Dr. Knigge, Frau Prof. Dr. Stanat und Prof. Dr. Hans Anand Pant in ihrem Beitrag Ziele, Umsetzung und Problematiken der Einführung bundesländerübergreifender Bildungsstandards dar.

Flankiert wurden die Darstellungen der vier Studienprogramme durch einen übergreifenden Vortrag von Herrn Dr. Goldhammer und Herrn Dr. Reeß, der Möglichkeiten, aber auch Probleme, die sich durch eine ***computerbasierte Erhebung der Kompetenzen in large-scale Assessments*** ergeben, thematisierte.

Im Folgenden sollen die zentralen Inhalte der einzelnen Beiträge jeweils zusammengefasst dargestellt werden:

1. Dr. Fellenberg, Prof. Dr. Blossfeld, Dr. von Maurice: Das Nationale Bildungspanel (NEPS). Bedeutsamkeit, theoretischer Ansatz, methodisches Design

Das nationale Bildungspanel wurde mit dem Ziel initiiert, zentrale Bildungsprozesse und -verläufe in Deutschland über die gesamte Lebensspanne abzubilden und zu analysieren, um Expertise über Bildungserwerb und dessen Folgen zu erlangen. Das NEPS ermöglicht somit Antworten auf Fragen die Entwicklung der Kompetenzen im Lebenslauf oder die Beeinflussung von Entscheidungsprozesse an verschiedenen kritischen Übergängen der Bildungskarriere durch Kompetenzen (und umgekehrt) be-

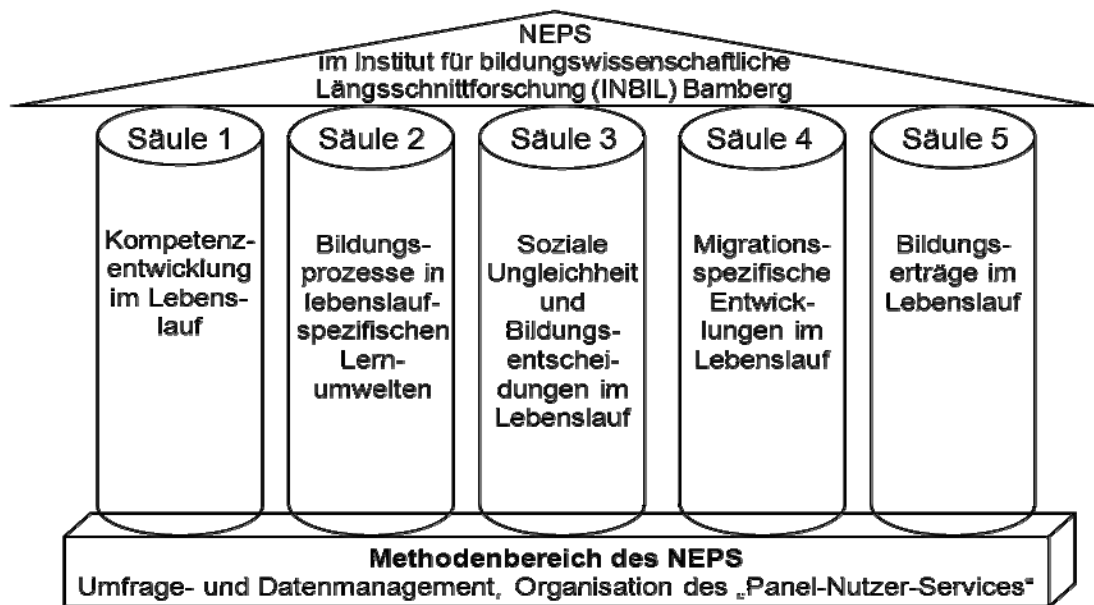
treffend oder aber wie und in welchem Umfang Kompetenzen von Lerngelegenheiten in der Familie, der Peer-Group und/oder der Gestaltung von Lehr- und Lernprozessen in Kindergarten, Schule, Hochschule und Berufsausbildung sowie Weiterbildung beeinflusst werden. Zur Beantwortung dieser Fragen ist ein Design notwendig, das (a) Kompetenzentwicklungen nicht nur im Kleinkind- und Schulalter, sondern auch in der beruflichen Ausbildung, im Studium und nach Verlassen des Bildungssystems misst und das (b) abhängige Maße zur Verfügung stellt und somit die Möglichkeit eröffnet, individuelle Verläufe zu analysieren.

Das NEPS ist in acht Alterstappen gegliedert, die von Neugeborenen bis zu Im-Beruf-Stehenden reichen (vgl. Tabelle 1), und in fünf theoretische Schwerpunktsäulen (vgl. Abbildung 1) und wird von einem bundesweiten Konsortium geleitet, die Förderung obliegt dem Bundesministerium für Bildung und Forschung.

Tabelle 1: Die Alterstappen des NEPS

Etappe 8: Berufliche Weiterbildung und lebenslanges Lernen	
Etappe 7: (Fach-) Hochschulen	→ Eintritt in Arbeitsmarkt
Etappe 6: Berufsausbildung	→ Eintritt in Arbeitsmarkt
Etappe 5: Gymnasiale Oberstufe (Sekundarstufe II)	→ Aufnahme Studium / Berufsausbildung
Etappe 4: Schulen der Sekundarstufe I	→ Sekundarstufe II / Eintritt in Arbeitsmarkt
Etappe 3: Grundschule	→ Schulen der Sekundarstufe I
Etappe 2: Kindergarten	→ Grundschule
Etappe 1: Neugeborene	→ Familie / frühkindl. instit. Betreuung

Abbildung 1: Die fünf Schwerpuntsäulen des NEPS



Auf Basis eines Sequenzkohortendesigns existieren für alle acht Altersgruppen Einstiegsstichproben, die dann jeweils jährlich weiterverfolgt werden (und zu vorher festgelegten Zeitpunkten aufgefrischt werden). Diese Ursprungsstichproben variieren in ihrer Größe zwischen 3.000 und 15.000 Befragten.

2. Dr. Nina Jude, Prof. Dr. Eckhard Klieme: PISA als Instrument des Bildungsmonitoring - zum Unterschied zwischen Querschnittsdaten und Trendanalysen

Das zentrale Anliegen des internationalen *Programme for International Student Assessment* (PISA) ist die kontinuierliche, international vergleichende Beobachtung von Bildungssystemen, das sogenannte Bildungsmonitoring. Dies beinhaltet zum einen den internationalen Vergleich auf Basis der zentralen Kompetenzen von Schülerinnen und Schülern, zum anderen ermöglicht PISA nationale wie international vergleichende Rückschlüsse auf Lernprozesse unter Zuhilfenahme der zusätzlich erhobenen Einschätzungen der Rahmenbedingungen und Funktionsmerkmale der entsprechenden Bildungssysteme. Schließlich bietet der zyklische Ansatz von PISA mit der alle drei Jahre wiederkehrenden Erhebung die

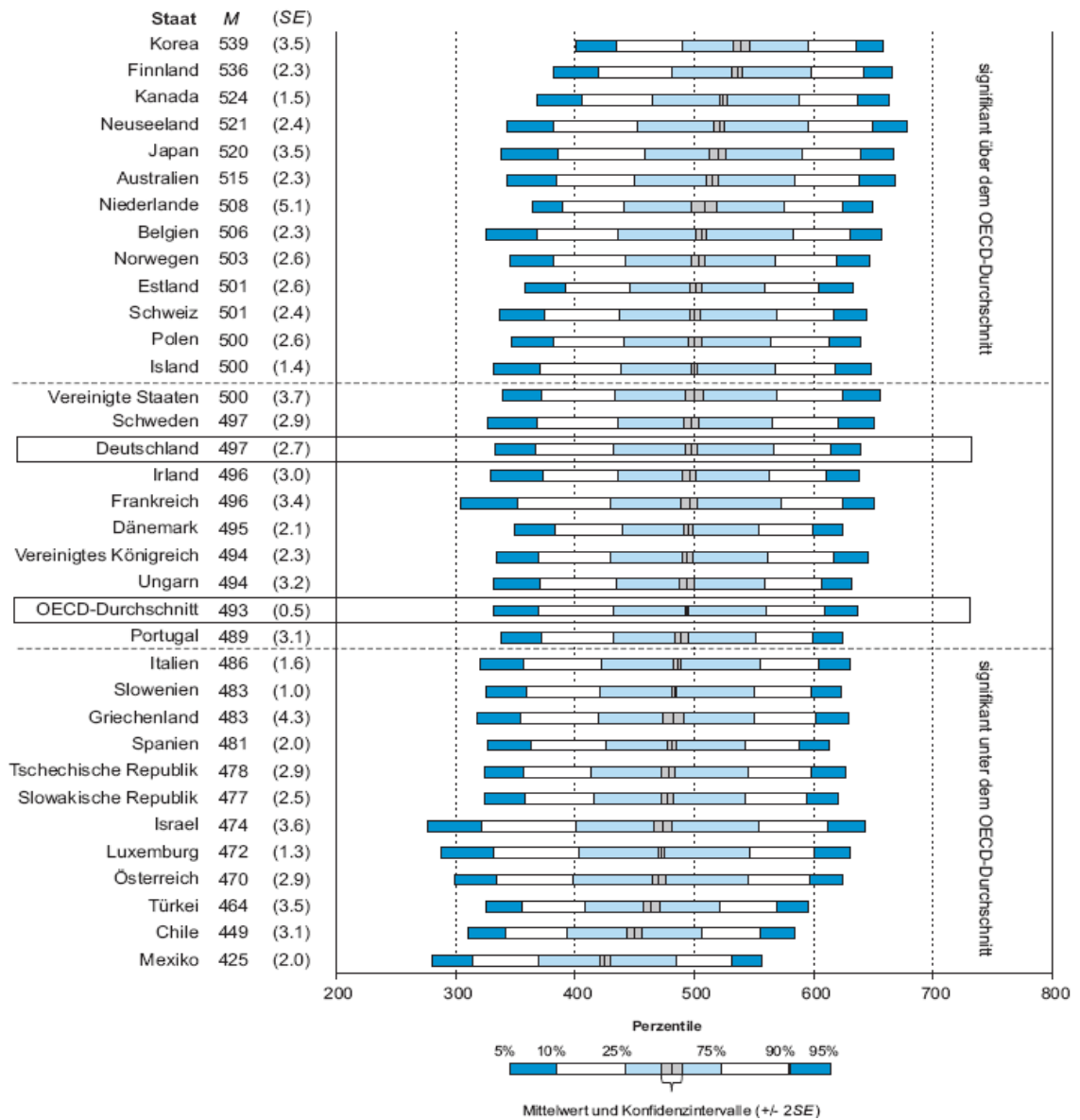
Möglichkeit, Veränderungen über die Zeit in Form von Trendentwicklungen abzubilden.

PISA ist ein OECD-gefördertes Programm an dem inzwischen 75 Staaten weltweit teilnehmen. Die Durchführung in Deutschland steht pro Zyklus unter wechselnder Leitung und wird gefördert durch die Länder und durch das BMBF. Das nationale Projektmanagement für PISA 2009 oblag dem DIPF.

PISA erfasst die zentralen Kompetenzen von Schülerinnen und Schülern im Alter von 15 Jahren. Diese zentralen Kompetenzen umfassen die drei Domänen Lesen, Mathematik und Naturwissenschaften. Diese Kompetenzdomänen rollieren über die Erhebungswellen, so dass pro Welle der Erhebungsschwerpunkt auf einer Kompetenzdomäne liegt. Ergänzt wird die Kompetenzmessung durch Schüler- und Schulfragebögen, die Rückschlüsse auf Bedingungs- und Rahmenfaktoren des Lernens und der Kompetenzentwicklung ermöglichen.

Die Ergebnisse aus PISA 2009 stimmen aus nationaler Sicht verhalten positiv: Die durchschnittliche deutsche Lesekompetenz liegt mit 497 Punkten erstmals im Bereich des OECD-Durchschnitts (vgl. Abbildung 2). Vergleicht man die nationalen Ergebnisse über die verschiedenen zurückliegenden PISA-Zyklen, ist eine kontinuierliche Verbesserung der Lesekompetenz zu verzeichnen. Der Vergleich mit zurückliegenden Zyklen bietet auch erste Hinweise auf mögliche Bedingungsfaktoren für die beobachtete Verbesserung in der Lesekompetenz. So zeigt sich beispielsweise, dass für die in PISA 2009 getestete Kohorte eine durchschnittliche frühere Einschulung zu verzeichnen war, sich die Unterrichtszeiten erhöht hatten und im Deutschunterricht weniger Störverhalten berichtet wurde.

Abbildung 2: Lesekompetenz im internationalen Vergleich basierend auf den Daten von PISA 2009



Quelle: Naumann, J.; Artelt, C.; Schneider, W. & Stanat, P. (2010). Lesekompetenz von PISA 2000 bis PISA 2009. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.). PISA 2009. Bilanz nach einem Jahrzehnt, S. 37. Münster: Waxmann.

3. Dr. Michel Knigge, Prof. Dr. Petra Stanat, Prof. Dr. Hans Anand Pant: Die Bildungsstandards und ihre Überprüfung

Mit den Konstanzer Beschlüssen von 2003 und 2004 hat die Kultusministerkonferenz (KMK) zum einen die Teilnahme an internationalen Schulleistungstudien und zum anderen für alle 16 Länder der Bundesrepublik Deutschland verbindliche Bildungsstandards für die Grundschule und für das Ende der Sekundarstufe I verabschiedet. Mit den KMK-Beschlüssen vom Oktober 2007 ist der Auftrag dahin gehend erweitert worden, auch Standards für das Ende der gymnasialen Oberstufe zu erarbeiten, und zwar erneut für die Fächer Deutsch, Französisch/Englisch, Mathematik und die drei Naturwissenschaften (Biologie, Chemie, Physik). Dies ermöglicht, dass in absehbarer Zeit in den so genannten „Kernfächern“ abschlussbezogene Standards für das gesamte allgemeinbildende Schulsystem in Deutschland vorliegen werden.

An die Verabschiedung der Bildungsstandards hat die KMK in ihren Plöner Beschlüssen vom Juni 2006 ein breites System der Qualitätssicherung gekoppelt, das u.a. stichprobenbasierte nationale Vergleiche und flächendeckende Vergleichsarbeiten (VERA) vorsieht. Mittels dieses Systems soll frühzeitig aufgezeigt werden können, ob Schülerinnen und Schüler die in den Bildungsstandards formulierten Leistungserwartungen einlösen können. Tabelle 2 bietet einen vergleichenden Überblick über Ziele, Designs und Standards (a) der internationalen Schulleistungstudien, (b) der nationalen Schulleistungstudien und (c) der bundesweiten Vergleichsarbeiten.

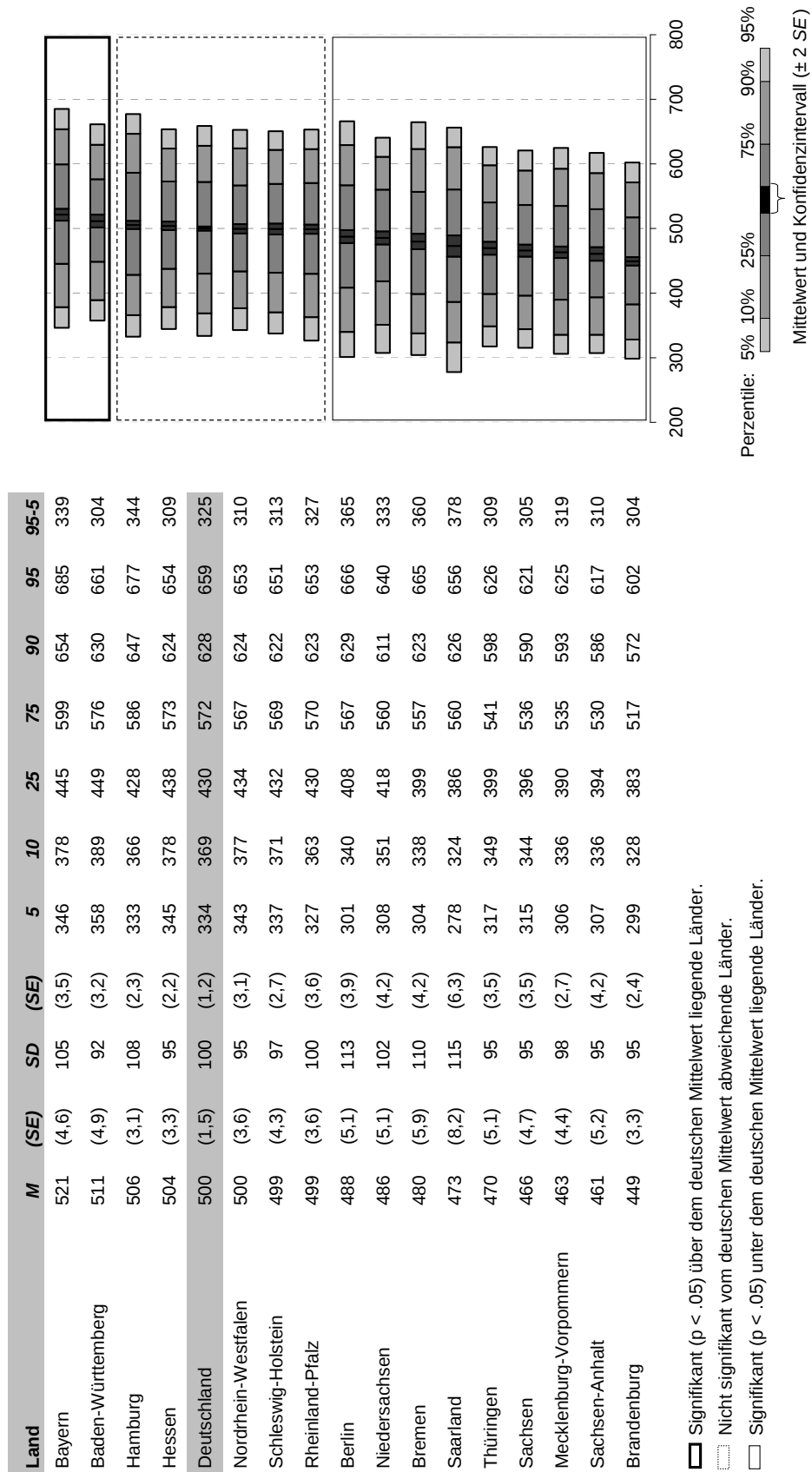
Bereits im Jahre 2008 wurden in ausgewählten Bundesländern Ländervergleiche *im Fach Französisch* mittels eines Tests auf Lese- und Hörverstehensleistungen durchgeführt. Erste bundesweite Ländervergleichstests wurden 2009 gestartet. Über 36.000 Schülerinnen und Schüler aus der 9. Jahrgangsstufe aller 16 Länder wurden in den Fächern Deutsch

Tabelle 2: Ziele, Designs und Standards (a) der internationalen Schulleistungstudien, (b) der nationalen Schulleistungstudien und (c) der bundesweiten Vergleichsarbeiten.

	Internationale Schulleistungstudien (PISA, PIRLS/IGLU, TIMSS)	Nationale Schulleistungstudien (KMK-Ländervergleiche des IQB)	Vergleichsarbeiten (VERA3 und VERA8)
Design	Stichprobenerhebung	Stichprobenerhebung	Vollerhebung der Schüler/-innen einer Jahrgangsstufe
Häufigkeit	alle 3-5 Jahre	alle 5J. (Grundschule) bzw. 6J. (Sekundarstufe)	jährlich
Zielfunktion	Systemmonitoring	Systemmonitoring	Unterrichts-/Schulentwicklung; regionales Monitoring
Rechenschaftspflicht ("High Stakes") für...	länderübergreifende Steuerungsakteure (KMK, Bildungspolitik des Bundes)	KMK, bildungspolitisch Verantwortliche in den Ländern	Schulen, Lehrkräfte, bildungspolitisch Verantwortliche in den Ländern
Bezugsnorm-orientierung	Sozialnorm/ Entwicklungsnorm	Kriteriumsnorm/Sozialnorm/Entwicklungsnorm	Kriteriumsnorm/ Sozialnorm/ Entwicklungsnorm
Durchführung	Testleiter/innen	Testleiter/innen	i.d.R. Lehrkräfte
Auswertung	zentral	zentral	dezentral durch Lehrkräfte u. Landesinstitute
Ergebnisrückmeldung:	nach ca. 1,5 Jahren	nach ca. 1 Jahr	zeitnah (Wochen)
Adressaten:	Öffentlichkeit, Scientific Community	Landespolitik, Öffentlichkeit, Scientific Community	Getestete, Lehrkräfte, Schulleitung, Eltern, Bildungsverwaltung
Bezug zu Bildungsstand. Kompetenzstufenmodell:	nein a posteriori, vorrangig anhand formaler Kriterien	ja a priori, kriterial (Bildungsstandards)	ja Übernahme der KMK-Modelle, z.T. adaptiert

und Englisch getestet. Berücksichtigte Teilkompetenzen waren für das Fach Deutsch Lesen, Zuhören und Orthografie; im Fach Englisch wurden Tests zum Lese- und Hörverstehen durchgeführt. Exemplarisch sind in Tabelle 3 die ländervergleichenden Ergebnisse für das Fach Englisch dargestellt. Wie aus der Tabelle ersichtlich schneiden die Länder Bayern und Baden-Württemberg gut ab, während die neuen Länder Bremen und das Saarland deutlich niedrigere Werte aufweisen.

Tabelle 3: Ergebnisse des IQB-Ländervergleichs 2009 im Hörverstehen Englisch



4. PD Dr. Beatrice Rammstedt: PIAAC - eine Studie zur international vergleichenden Kompetenzmessung im Erwachsenenalter

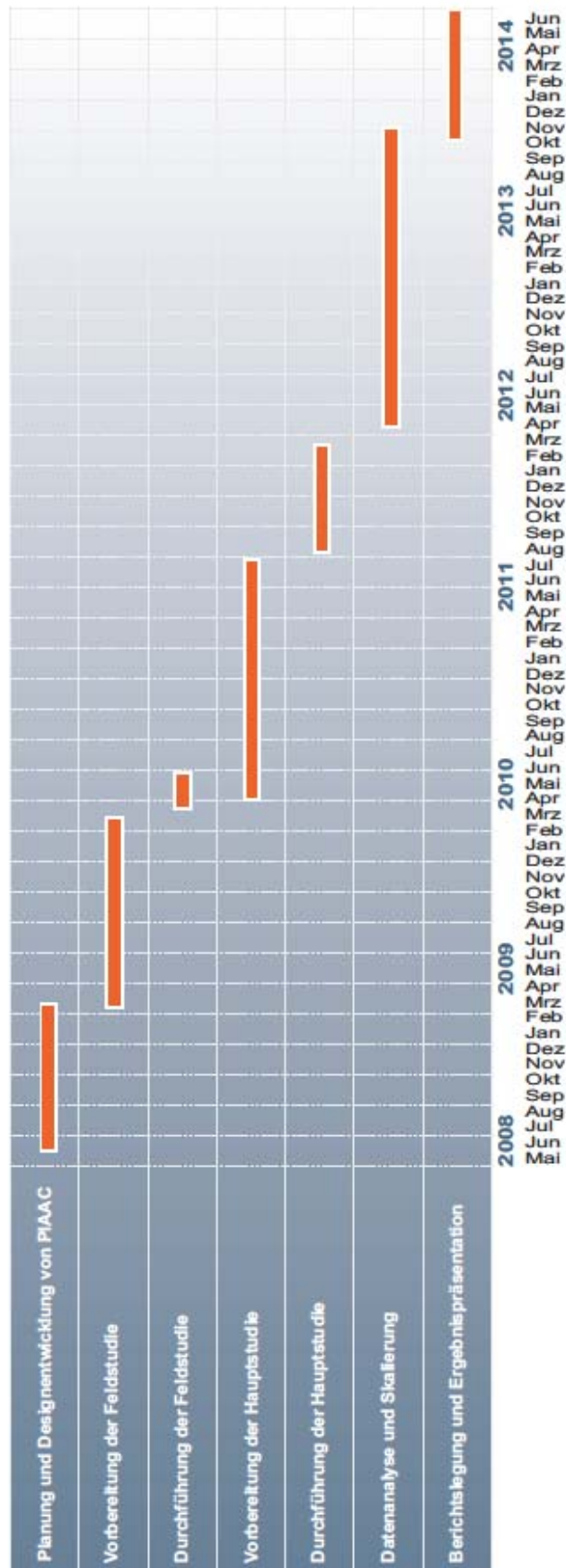
Das *Programme for the International Assessment of Adult Competencies* (PIAAC) wurde von der OECD mit dem Ziel initiiert, das Kompetenzniveau von Erwachsenen im internationalen Vergleich zu untersuchen. PIAAC soll in den 26 beteiligten Ländern als Querschnittsuntersuchung in mehreren Wellen durchgeführt werden. Die erste Welle startete 2008, die Ergebnisse werden 2013 publiziert werden. Ein detaillierter Zeitplan ist in Abbildung 3 dargestellt.

In der ersten Welle konzentriert sich PIAAC auf die Erfassung der Alltagskompetenzen *Literacy*, *Numeracy*, *Reading Components* und *Problem Solving in Technology Rich Environments*. Ergänzend hierzu ermöglicht der zusätzlich erfasste umfangreiche Hintergrundfragebogen Erkenntnisse zum Zusammenhang dieser Kompetenzen mit sozialen und ökonomischen Merkmalen und somit Rückschlüsse auf zentrale Bedingungsfaktoren für die Kompetenzentwicklung und den Kompetenzerhalt im Erwachsenenalter, speziell im Berufsalter. Der Hintergrundfragebogen wird daher neben Informationen über den Haushalt, über Aus- und Weiterbildung und über die berufliche Stellung und berufliche Karriere Informationen über allgemeine und berufliche Verwendung der eigenen Fähigkeiten, über Einstellungen zum Lernen und über soziale Outcomes erfassen.

An PIAAC werden pro Land mindestens 5.000 zufällig ausgewählte Erwachsene im Alter zwischen 16 und 65 Jahren teilnehmen. Durchgeführt wird PIAAC als persönliche Befragung im Haushalt der Zielperson. Während der Hintergrundfragebogen in Form eines persönlichen Interviews erhoben wird, werden die Kompetenzen direkt, d.h. ohne Beteiligung, aber unter Anwesenheit des Interviewers, erhoben.

PIAAC enthält zahlreiche innovative Aspekte, die es zu einer herausfordernden, aber auch zukunftsweisenden Studie machen: Von methodischer Seite ist PIAAC insbesondere insofern innovativ, weil erstmals in einer solchen large-scale Studie sowohl der Hintergrundfragebogen wie die Kompetenzmessung komplett computergestützt erfolgen. Von inhaltlicher Seite bietet PIAAC zahlreiche Neuerungen. Mit *Problem Solving in*

Abbildung 3: Der Zeitplan von PIAAC



Technology Rich Environments und *Reading Components* werden zwei neue Kompetenzdomänen in die international vergleichenden Kompetenzstudien eingeführt. Während *Problem Solving in Technology Rich Environments* mit seiner Ausrichtung an Strategien zur Navigation im Internet, der Nutzung von Online-Tools und der Bewertung von Online-Inhalten stark auf die Zukunftsfähigkeit der verschiedenen Nationen und auf Wege, diese zu erreichen, abzielt, ermöglicht die bevölkerungsrepräsentative Erfassung der grundlegenden Lesefähigkeiten (*Reading Components*) erstmals Hinweise auf den Prozentsatz und die Verteilung der Personen in den entsprechenden Nationen, die nur über sehr geringe Lesekompetenzen verfügen und somit hierdurch Schwierigkeiten im beruflichen wie im privaten Alltag haben werden.

Die PIAAC-Datenbasis, die interessierten Forschern frei zur Verfügung stehen wird, bietet daher sowohl aus nationaler Perspektive wie auch aus internationaler umfangreiche Möglichkeiten zur Analyse bisher in diesem Umfang unerforschter Fragestellungen.

Auf nationaler Ebene wird PIAAC vom Bundesministerium für Bildung und Forschung gefördert unter Beteiligung des Bundesministeriums für Arbeit und Soziales. Das nationale Projektmanagement hat GESIS inne.

5. Dr. Frank Goldhammer, Dr. Jean-Paul Reeß: Computer-basiertes Testen in Large Scale Assessments: Mehr und bessere Verhaltensdaten?

Computer-basiertes Testen hält zunehmend Einzug in nationale und internationale Large Scale Assessments, die bisher hauptsächlich in Papierbasierter Form administriert wurden. Gründe für diese zunehmende Verbreitung können vor allem in den zahlreichen Chancen gesehen werden, die Computer-basiertes Testen für solche Studien bietet: So kann durch die standardisierte Auslieferung, Instruktion und automatische Antwortbewertung (Scoring) bei Computer-administrierter Erhebung die Objektivität und damit die Vergleichbarkeit von Studienergebnissen deutlich erhöht werden. Ein zweiter zentraler Vorteil ist die Verbesserung der Messeffizienz durch computergestütztes Testen. Durch adaptive Itemauswahl und somit durch Nicht-Administration inadäquater Items kann teure Testzeit eingespart

werden (z.B. mittels adaptiven Testens in Kompetenzassessments oder mittels antwortabhängiger Routings in Fragebogen zu Hintergrundvariablen). Ein ganz entscheidender Vorteil von Computer-gestützter Erhebung ist die Möglichkeit, die zu erfassenden Konstrukte neuartig zu operationalisieren, welches somit in einer verbesserten und damit inhaltsvalideren Repräsentation des Konstrukts resultiert. So können z.B. interaktive Szenarien und Simulationen gestaltet werden, die insbesondere bei Konstrukten mit immanentem Technologiebezug eine deutlich alltagsnähere Kompetenzerfassung erlauben (z.B. Eletronic Reading). Auch von Seiten der erhobenen Daten bietet die Computer-gestützte Erhebung deutliche Vorteile gegenüber der klassischen Papier-basierten Erhebung mit anschließender Datencodierung: Neben der Objektivierung der Dateneingabe und -auswertung sind problemlos auch zusätzliche informative Daten wie die Beantwortungslatenz oder typische Bearbeitungswege fehlerhafter Antworten erfassbar.

Neben diesen zahlreichen Chancen ist ein Computer-gestütztes Erhebungsdesign in large-scale Studien auch mit Herausforderungen konfrontiert, denen es gilt adäquat zu begegnen: So ist für die meisten Studien z.Z. eine generelle computergestützte Erfassung nicht umsetzbar, da nicht alle Befragten über hinreichend Computerliteralität verfügen und somit der Test diese Personengruppen benachteiligen würde (mangelnde Test-fairness). Eine Strategie, mit diesem Problem umzugehen, besteht darin, klassische Papier-basierte wie die Computer-gestützte Version zu erstellen. Diese beiden Versionen müssen hinsichtlich ihrer Reliabilität und Validität vergleichbar sein und sich ineinander überführen lassen (Äquivalenz der Messinstrumente). Darüber hinaus existiert natürlich auch die rein praktische Herausforderung der Kosten-Nutzen-Überlegung: Die Computer-basierte Umsetzung von Befragungen ist i.d.R. mit hohen Anfangsinvestitionen in Hard- und Software sowie einem hohen personellen Aufwand in der Erstellung wie der Betreuung verbunden.

Zurzeit wird in einigen der im Forum dargestellten large-scale Studien die Computer-gestützte Umsetzung exploriert oder sogar flächendeckend umgesetzt. So wurde in PISA 2000 erstmals als Teil der nationalen Ergänzung in Deutschland Computer-based Assessment im Rahmen des dynami-

schen Problemlösens mit Erfolg getestet. Dieser Versuch mündete in ein Computer-based Assessment der naturwissenschaftlichen Kompetenzen als internationale Option in PISA 2006, das von drei Teilnehmerstaaten durchgeführt wurde, und eines Electronic Reading Assessments ebenfalls als internationale Option in PISA 2009, das diesmal bereits von 17 Teilnehmerstaaten eingesetzt wurde.

Basierend auf den positiven Erfahrungen dieser Optionen sollen in PISA 2012 erstmals flächendeckend Computer-basierte Module eingesetzt werden. Für PISA 2015 besteht die Bestrebung, die komplette Kompetenzmessung computerbasiert zu operationalisieren.

Eine noch weitergehende vollständige Computerbasierung ist in PIAAC bereits in der ersten Welle umgesetzt worden. So wird in PIAAC sowohl der Hintergrundfragebogen wie auch die Kompetenzmessung einheitlich computergestützt erfolgen. Für Personen mit geringer Computerliteralität wird die Kompetenzmessung alternativ jedoch auch als papierbasierte Version angeboten.

Die Beiträge des Forums haben somit einen guten Überblick über den Stand der Implementierung von Kompetenzmessung in die large-scale Bildungsforschung aus nationaler Sicht vermittelt. Die Diskussion nach den einzelnen Beiträgen konzentrierte sich zum einen auf die Verfügbarkeit der Daten. Dieser Aspekt verdeutlichte daher noch einmal die Nachfrage nach solchen Datensätzen aus der Scientific Community und somit den Bedarf entsprechend aufbereiteter, dokumentierter und zugänglicher Daten, um damit auch die Nachhaltigkeit und den Mehrwert der – meist sehr aufwändig – erhobenen Daten zu verbessern. Zum anderen wurde die Vergleichbarkeit der Studien diskutiert. Dieser Punkt weist auf eine zentrale Frage in der aktuellen kompetenzbasierten Bildungsforschung hin. Inwiefern sind Kompetenzen, erfasst im NEPS, in PISA oder aber in PIAAC, miteinander vergleichbar? Welche Parallelitäten, aber auch welche Unterschiede gibt es in deren Konzeptualisierung? Und inwiefern sind die Bildungsstandards hiermit vergleichbar? Das Forum im Rahmen der 5. Konferenz für Sozial- und Wirtschaftsdaten war ein gelungener Auftakt für eine solche Diskussion, die sicherlich in Zukunft weitergeführt werden wird.