

Wer hat einen Nutzen von Kompetenzmodellen?

Hans Anand Pant

Zusammenfassung: Die in Deutschland 2003/2004 verbindlich eingeführten Bildungsstandards stellen inhaltliche und normative Kompetenzerwartungen dar. In ihnen wird anhand konkreter Könnensbeschreibungen für schulische Kernfächer formuliert, welche Kompetenzen Schüler/-innen zu bestimmten Zeitpunkten ihrer Bildungslaufbahn aufweisen sollen. Um das Erreichen dieser normativen Vorgaben überprüfen zu können, wurde von der Konferenz der Kultusminister der Länder (KMK) eine Strategie zum Bildungsmonitoring aufgesetzt, die neben einem testbasierten Schulleistungsvergleich auf Länderebene auch jährliche standardbasierte Lernstandserhebungen für alle dritten und achten Klassen in Deutschland vorsieht. Die Tests basieren auf fachdidaktisch entwickelten und psychometrisch validierten Kompetenzmodellen. Es wird anhand von Beispielen aufgezeigt, welche unterschiedlichen Potenziale die Ergebnismeldungen aus solchen Kompetenztestungen auf der jeweiligen Aggregatebene (Bundesland, Schule, Lehrkraft, Schüler) haben. Es wird argumentiert, dass das Potenzial kompetenzbasierter Rückmeldungen am ehesten ausgeschöpft wird, wenn diese adressatengerecht kommuniziert und mit kohärenten Unterstützungsangeboten verknüpft werden.

Schlüsselwörter: Bildungsstandards · Systemmonitoring · Kompetenzniveaumodelle · Testrückmeldungen

Who benefits from competence models?

Abstract: National educational standards (NES) were released in Germany 2003/2004 as obligatory performance standards substantively and normatively defining an acceptable degree of proficiency. The NES describe tangible competencies in the form of can-do statements for core subjects to be acquired by students at the end of defined stages of schooling. To evaluate standards-based performances the Standing Conference of the Ministers of Education and Cultural Affairs of the Federal States in Germany (KMK) implemented a test-based monitoring system that comprised a centralized evaluation in a nationwide study sample comparing the German federal states (Ländervergleich), as well as state-wide testing schemes at the end of grade 3 and grade 8 linked to the NES. All test items were designed by didactic experts on the basis of domain specific competence models that were psychometrically evaluated. The potential practical benefit from competence testing is discussed for the various system levels of feedback (state, school, teacher, and student). It is argued that, in order to tap the full potential of standards-based per-

formance feedback, it must be appropriately communicated to the target system level and should be accompanied by suitable measures of action support.

Keywords: Educational standards · Performance feedback · Proficiency level models · System monitoring

Die Entwicklung, Überprüfung und Anwendung von Kompetenzmodellen und entsprechenden Messinstrumenten erfordern umfangreiche Ressourcen sowohl auf wissenschaftlicher als auch auf schulischer Seite. Der diesem Ressourcenbedarf gegenüberstehende wahrgenommene oder tatsächliche Nutzen hängt jedoch wesentlich vom jeweiligen Akteur bzw. von der jeweiligen Akteurin innerhalb des Bildungssystems und dessen Interessen beziehungsweise Handlungsnotwendigkeiten ab.

Zurzeit werden standardisierte und zyklisch stattfindende Kompetenzmessungen im deutschen Bildungssystem durchgeführt im Rahmen

- flächendeckender einzelschulbezogener Lernstandserhebungen (Vergleichsarbeiten in den Jahrgangsstufen 3, 6 und 8, im Folgenden VERA-3, VERA-6 bzw. VERA-8),
- der Überprüfungen des Erreichens der länderübergreifenden Bildungsstandards der Kultusministerkonferenz (kurz: Ländervergleich) sowie im Rahmen
- internationaler Schulleistungsstudien (PISA, IGLU/PIRLS, TIMSS) auf Staatenebene.

Die Implementation dieses Systems schulextern vorgegebener Kompetenzfeststellungen beruht auf einer Gesamtstrategie der Kultusministerkonferenz (KMK) zum Bildungsmonitoring (KMK 2006). Die Sicherung und Entwicklung von Bildungsqualität sollte demnach durch eine Konzentration auf die Bildungserträge der Schüler/-innen, der Schulen und des schulischen Bildungssystems insgesamt erzielt werden und nicht mehr nur über die Modifikation von Lehrplänen, die Verbesserung der Lehrkräfteausbildung oder die Entwicklung und Erprobung neuer didaktischer Modelle. Das Kernstück dieser Strategie sind die in den Jahren 2003 und 2004 von der KMK verabschiedeten länderübergreifenden Bildungsstandards, die für alle Schulen als verbindlich festgelegt wurden.

Bildungsstandards stellen eine Spezifizierung allgemeinerer Bildungsziele dar und können als normative Kompetenzerwartungen an Schüler/-innen verstanden werden. Sie definieren, welche – auf Wissen und Kenntnissen fußenden – Fähigkeiten und Fertigkeiten zu verschiedenen Zeitpunkten der Schullaufbahn von den Lernenden entwickelt worden sein sollen. Der Kompetenzbegriff, so wie er in den Bildungsstandards und den aktuellen Schulleistungsuntersuchungen Verwendung findet, bezieht sich dabei auf ein relativ pragmatisches Verständnis: Kompetenzen werden als Fähigkeiten konzeptualisiert, die sich in konkreten Anforderungssituationen relativ „oberflächennah“ als ein Können manifestieren. In den KMK-Bildungsstandards zeigt sich dies in der durchgängigen Formulierung der Kompetenzerwartungen als Könnensbeschreibungen (Can-do-Statements), zum Beispiel: „Schülerinnen und Schüler können die Aussagen einfacher literarischer Texte verstehen“, und „in klar geschriebenen argumentativen Texten zu vertrauten Themen die wesentlichen Schlussfolgerungen erkennen, z. B. in Zeitungsartikeln“ (KMK 2004, S. 12). Die Kompetenzmodelle der KMK-Bildungsstandards, die auch den flächendeckenden Lernstandserhebungen zugrunde liegen, weisen einerseits domänenpezifische

Kompetenzstrukturmodelle und andererseits Domänen übergreifende Kompetenzniveau-Modelle mit derzeit fünf Stufen je fachbezogenem Standard auf. Die Kategorisierung von gemessenen Leistungsergebnissen als Niveaustufen wird normativ mit einem Standard etikettiert („unter Mindeststandard“, „Mindeststandard“, „Regelstandard“, „Regelstandard plus“, „Optimalstandard“; vgl. Pant et al. 2012).

Die Frage nach dem Nutzen der Kompetenzmodelle ist demnach eng verbunden mit der Frage, in welchem Umfang Kompetenzmodelle Funktionen im Rahmen des Bildungsmonitorings beziehungsweise des schulaufsichtlichen Rechenschaftswesens, der Unterrichts- und Schulentwicklung oder der individuellen Leistungsfeststellung und -beurteilung erfüllen sollen (Tab. 1; vgl. Pant et al. 2011). Selbst bei bestehenden Large-Scale-Untersuchungen gibt es wichtige Unterschiede hinsichtlich der jeweiligen Zielfunktion (Wozu sollen die Ergebnisse dienen?), der Aussage- oder Evaluationsebene (Wer wird evaluiert?) und der Bezugsnormorientierung (Woran orientiert sich die Beurteilung der Testergebnisse?).

1 Nutzen auf Systemebene

Auf systemischer Ebene besteht der Nutzen von Kompetenzmodellen darin, dass überhaupt ein einheitliches Referenzmodell zur Verfügung steht, auf das sich politische, fachdidaktische und schulpraktische Vertreter/-innen in einem langwierigen und kontroversen Prozess verständigen konnten und auf das sie sich in Folge konsistent beziehen können. Die psychometrische Umsetzung dieser Kompetenzmodelle in Kompetenzskalen mit einer einheitlichen Metrik ermöglicht es, dass Leistungsrückmeldungen an Lehrkräfte, Schulen, Schulverwaltungen und Bildungspolitik auf unterschiedlichsten Aggregatstufen in einer einheitlichen „Währung“ (Kompetenzpunkte bzw. Kompetenzstufen) erfolgen kann. Durch Linking-Studien wird die Metrik der nationalen Studien mit denjenigen der internationalen Schulleistungsstudien abgestimmt. Die inhaltliche und psychometrische Anbindung von VERA-3 und VERA-8 an das nationale und internationale Systemmonitoring wird dadurch gewährleistet, dass die Aufgabenentwicklung und Pilotierung für alle Studien „aus einer Hand“ erfolgt (vgl. Pant et al. 2011).

Die primäre Zielfunktion von Kompetenzmodellierung und Kompetenztestung aus systemischer Perspektive ist es, politisches oder bildungsadministratives Steuerungswissen zu generieren. Zwei Beispiele: Wößmann (2005, 2006) demonstriert das Nutzenpotenzial international vergleichender Kompetenzstudien (hier: TIMSS), indem er Effekte institutioneller Merkmale von Bildungssystemen, die innerhalb eines Bildungssystems nicht variieren, auf die Mathematikkompetenzen von Schülerinnen und Schülern untersucht. Er kommt auf Staatenebene unter anderem zu folgenden steuerungsrelevanten Aussagen: Schüler/-innen schneiden in Kompetenztests besser ab

- in Ländern mit mehr Wettbewerb durch privat geleitete Schulen,
- wenn rechtliche Rahmenbedingungen Schulen mit Autonomie in Prozess- und Personalentscheidungen ausstatten,
- wenn Lehrkräfte sowohl Anreize als auch Möglichkeiten haben, angemessene Lehrmethoden auszuwählen,

Tab. 1: Vergleich der aktuell laufenden Large-Scale-Studien im Rahmen der Gesamtstrategie der KMK zum Bildungsmonitoring

	Vergleichsarbeiten/Lernstandserhebungen (VERA-3/VERA-8)	Nationale Schulleistungsstudien (KMK-Ländervergleiche)	Internationale Schulleistungsstudien (PISA, PIRLS/IGLU, TIMSS)
Design	z. T. verpflichtende Vollerhebung, z. T. freiwillige Teilnahme	Stichprobenerhebung („Zufallsauswahl“)	Stichprobenerhebung („Zufallsauswahl“)
Häufigkeit	Jährlich	Alle 5 (Grundschule) bzw. 6 Jahre (Sekundarstufe und Fach)	Alle 3–5 Jahre
Zielfunktion	Unterrichts-/Schulentwicklung; regionales Monitoring	Systemmonitoring	Systemmonitoring
Rechenschaftspflicht („high-stakes“) für	Schulen, Lehrkräfte, bildungspolitisch Verantwortliche in den Ländern	KMK, bildungspolitisch Verantwortliche in den Ländern	Länderübergreifende Steuerungsakteure (KMK, Bildungspolitik des Bundes)
Bezugsnormorientierung	Kriterialnorm (Bildungsstandards)/Sozialnorm/Entwicklungsnorm	Kriterialnorm (Bildungsstandards)/Sozialnorm/Entwicklungsnorm	Sozialnorm/Entwicklungsnorm
Durchführung	Lehrkräfte	Testleiter/-innen	Testleiter/-innen
Auswertung	Dezentral durch Lehrkräfte/Landesinstitute	Zentral	Zentral
Ergebnisrückmeldung			
Zeitpunkt	Zeitnah (Wochen)	Nach ca. 1 Jahr	Nach ca. 1,5 Jahren
Adressaten	Getestete, Lehrkräfte, Schulleitung, Eltern, Bildungsverwaltung (Schulaufsicht, Schulinspektion)	Landespolitik, Öffentlichkeit, Scientific Community	Öffentlichkeit, Scientific Community
Bezugnahme auf Bildungsstandards	Ja	Ja	Nein
Kompetenzstufenmodell	Z. T. Ankoppelung an die Modelle der Bildungsstandards, ggf. adaptiert	A priori, inhaltlich-kriterial (Bildungsstandards), formales Standard-Setting-Verfahren	A posteriori, vorrangig anhand formaler Kriterien

- wenn Schulen durch externe Prüfungen zur Verantwortung gezogen werden, und insbesondere
- wenn externe Prüfungen und Schulautonomie kombiniert werden (vgl. Wößmann 2005, 2006).

Auch im ersten Ländervergleich zu sprachlichen Kompetenzen, der 2009 auf der Basis der Kompetenzmodelle der KMK-Bildungsstandards und des Gemeinsamen Europäischen Referenzrahmens für Sprachen durchgeführt wurde, zeigten sich massive Unterschiede in den Englischkompetenzen (Lese- und Hörverstehen) zwischen den alten und

neuen Bundesländern, die in der Größenordnung von 1–1,5 Jahrgangsstufenäquivalenten liegen (vgl. Leucht et al. 2010). Als Konsequenz haben die neuen Länder politische Maßnahmenkataloge und konkrete Interventionen im Bereich der Fort- und Weiterbildung für Englischlehrkräfte konzipiert und mit erheblichen Ressourcen versehen.

2 Nutzen auf Schul- und Unterrichtsebene

In dem Ende 2009 verabschiedeten Beschluss der KMK zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung (KMK 2010) wird bekräftigt, dass Leistungsrückmeldungen aus Vergleichsarbeiten als zentraler Bestandteil eines „datengestützten Entwicklungskreislaufs an einer Schule“ fungieren sollen (KMK 2010, S. 13).

Die zentralen pädagogischen Potenziale der Vergleichsarbeiten für Lehrkräfte und Schulen liegen, neben der durchgängigen Kompetenzorientierung von Testaufgaben, Ergebnissrückmeldungen und begleitenden didaktischen Kommentierungen, bei

- dem „Blick von außen“, d. h. den multiplen Vergleichsmöglichkeiten zum Lernstand der eigenen Klasse,
- dem Ausbau diagnostischer Kompetenzen bei Lehrkräften,
- der Begründung und Planung pädagogischer Interventionen und Fördermaßnahmen,
- der Nutzung der Leistungsrückmeldungen für kooperative Unterrichtsentwicklung im Kollegium.

2.1 Multiple Vergleichshorizonte

Landesweite Kompetenztests erlauben neben den sozialnormorientierten Vergleichen auch kriteriale Vergleiche. Die Leistungsrückmeldung kann dank der flächendeckenden Durchführung auf nahezu beliebiger Aggregatebene erfolgen. Dies ermöglicht beispielsweise Lehrkräften, den Lernstand ihrer Klasse mit dem von Parallelklassen, dem Schuldurchschnitt oder anderen regionalen Referenzgruppen zu vergleichen.

Jede Lehrkraft kann eine Standortbestimmung ihrer Klasse im Hinblick auf die vorab definierten Kompetenzziele der Bildungsstandards vornehmen. Je nach Kompetenzmodell können theoretisch fundierte Kompetenzprofile erstellt werden, die zum Beispiel auf Klassenebene eine spezifische Stärken-Schwächen-Analyse zulassen.

2.2 Ausbau und Optimierung diagnostischer Kompetenzen bei Lehrkräften

Kompetenztestungen erlauben es Lehrkräften, ihre eigene Diagnosefähigkeit zu prüfen und zu trainieren. Unter diagnostischer Kompetenz im engeren Sinne wird die Urteilsgenauigkeit verstanden, mit der Lehrkräfte Schülerleistungen einschätzen können (vgl. ausführlich Helmke 2009). Wird das Augenmerk nicht auf die Lernenden, sondern auf die Testaufgaben gelegt, dann zeigt sich hohe diagnostische Kompetenz in der korrekten Beurteilung der Schwierigkeit von Aufgaben, gemessen über die vorhergesagte beziehungsweise tatsächliche durchschnittliche Lösungshäufigkeit in der Klasse (vgl. Leucht et al. 2012).

2.3 Begründung pädagogischer Interventionen

Für die einzelne Lehrkraft eröffnet sich durch die Kenntnis des Stärken-Schwächen-Profiles ihrer Klasse die Möglichkeit, differenzierte, zugeschnittene Fördermaßnahmen zu begründen und zu initiieren. Darüber hinaus können Lehrkräfte die diagnostische Information aus Kompetenztests als ein zusätzliches, jenseits eigener Bewertungsmuster erstelltes Datum in Elterngesprächen nutzen. Um entsprechend geeignete Fördermöglichkeiten zur Verfügung zu stellen, können Kompetenzmodelle in der fachdidaktischen Forschung zur Überprüfung der Effektivität von Interventionen eingesetzt werden.

2.4 Nutzung für Schulentwicklungsprozesse

Zu den praxisbezogenen Vorschlägen auf Schulebene zählen die Stärkung von Kommunikation und Kooperation bei der Reflexion von Rückmeldeergebnissen und der Entwicklung angemessener Fördermaßnahmen im Rahmen professioneller Lerngemeinschaften (vgl. Fishman et al. 2003). Als besonders geeignete Ebene gelten jahrgangsbezogene Fachgruppen beziehungsweise jahrgangsübergreifende Fachkonferenzen. Die Diskussion im Kollegenkreis von einer ersten Einschätzung und Analyse über die Codierung und Bewertung der Kompetenztest-Aufgaben bis zu den später bereitgestellten Rückmeldungen bietet zahlreiche Austauschgelegenheiten.

Bewährt haben sich die Analyse der Aufgaben und Fehlermuster aus methodisch-didaktischer Sicht und die Diskussion möglicher Fehlerursachen. Als Folgerungen sind unter anderem denkbar, in Parallelklassen wechselseitig zu hospitieren, Unterrichtsprojekte gemeinsam vorzubereiten, Materialien arbeitsteilig zu erstellen und auszutauschen oder gezielte Fortbildungen zu organisieren.

Die empirischen Befunde zum Rezeptionsverhalten von Lehrkräften (vgl. z. B. Maier 2008) zeigen jedoch, dass insgesamt noch erhebliche Forschungslücken hinsichtlich der Determinanten des Nutzungsverhaltens auf Kollegiumsebene existieren.

3 Nutzen für individualdiagnostische Zwecke

Weder die internationalen und nationalen Schulleistungsstudien noch die an allen Schulen durchgeführten Vergleichsarbeiten streben von ihrer Konzeption her individualdiagnostische Aussagen an. Dennoch werden in der Praxis individuelle Lernstandsrückmeldungen aus den Kompetenztestungen sowohl von Lehrkräften als auch von Eltern und den Lernenden selbst erwartet und eingefordert.

Das zentrale Argument gegen eine individualdiagnostische Nutzung wird darin gesehen, dass Schulleistungsstudien auf individueller Ebene keine reliablen diagnostischen Informationen liefern (vgl. Leutner et al. 2007). Diese Ungenauigkeiten in den Messungen individueller Lernstände können bei der Zuordnung einzelner Schüler/-innen zu den jeweiligen Kompetenzstufen zu Fehlklassifikationen führen (vgl. Ercikan und Julian 2002).

Insgesamt müsste ein auf Individualebene regelmäßig in Schulen einsetzbares Instrument Folgendes ermöglichen:

- a. Individualdiagnostische Aussagen,
- b. formative Evaluation, d.h. reliable und änderungssensitive Erfassung der Kompetenzentwicklung von Schülerinnen und Schülern innerhalb eines Schuljahres,
- c. selbst initiierte, einfache Anwendung und
- d. weitgehende „Hoheit“ über die erzeugten Ergebnisse beziehungsweise Ergebnismeldungen (vgl. u. a. das in den Niederlanden existierende Modell für ein Schülermonitoring-System; Staphorsius und Krom 2008).

Im Bereich der (Fremd-)Sprachen existiert bereits für eine Reihe von Sprachen ein umfangreiches, auf Individualebene nutzbares System. Dieses besteht unter anderem aus

- der klaren Beschreibung der Kompetenzen im Europäischen Referenzrahmen für Sprachen (Council of Europe 2001),
- deren Integration in bildungspolitisch relevante Dokumente (z. B. KMK-Bildungsstandards für die erste Fremdsprache) und für den Lerner bzw. die Lernerin potenziell relevante gesellschaftliche Anforderungsprofile (z. B. Studienzugangsvoraussetzungen),
- Selbst- bzw. Fremdeinschätzungsbögen in unterschiedlichen Detailgraden (z. B. <http://europass.cedefop.europa.eu/LanguageSelfAssessmentGrid/en>),
- Portfoliovorlagen zur Dokumentation der Kompetenzentwicklung (z. B. <http://www.sprachenportfolio.ch>),
- Testinstrumenten (z. B. <http://diplomas.cervantes.es>),
- umfangreichen und abgestimmten Angeboten in Bezug auf Lernmaterialien (auch zum Selbststudium) und Kurse.

Gerade aufgrund der hohen Verfügbarkeit (z. B. Online-Versionen vergangener Prüfungen inkl. Lösungsschlüssel) ist dieses Modell damit nicht nur in Schulen, sondern vor allem auch von den Lernerinnen und Lernern selbst einsetzbar. Durch die mittlerweile hohe gesellschaftliche Anerkennung des Modells und der Testinstrumente wird gleichzeitig auch eine Einschätzungs- und Vergleichsmöglichkeit individueller Lerner/-innen geboten (z. B. im Rahmen von Job- oder Studienbewerbungen).

4 Fazit

Kompetenzmodelle weisen vielfältige Nutzungsmöglichkeiten auf. Ihr tatsächlicher Nutzen hängt jedoch von unterschiedlichen Faktoren ab, insbesondere von der ihnen zugewiesenen gesellschaftlichen bzw. bildungspolitischen Rolle und Anerkennung und der Verfügbarkeit und Zugänglichkeit von entsprechenden Testinstrumenten und Lernmaterialien.

Insgesamt lässt sich dabei immer noch eine große Heterogenität in Bezug auf den Entwicklungsstand von Kompetenzmodellen und den sie begleitenden Angeboten zwischen den einzelnen schulischen Fächern beobachten. Gerade im nicht-sprachlichen Bereich kann dabei das weitgehende Fehlen von passgenauen Unterstützungsangeboten die Motivation von Lehrkräften (und Lernenden) untergraben, sich mit Kompetenzmodellen und kompetenzorientierten Rückmeldungen aus Leistungstests auseinanderzusetzen. Gerade in Domänen mit bereits vorhandenen empirisch validierten Kompetenzmodellen sollten

zukünftige Strategien den Anwendungsnutzen von Kompetenztestungen dadurch erhöhen, dass sie begleitende Angebote ins Zentrum stellen.

Danksagung: Diese Veröffentlichung wurde ermöglicht durch Sachbeihilfen der Deutschen Forschungsgemeinschaft (Kennz.: PA 1532/2) im Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (SPP 1293).

Literatur

- Council of Europe. (2001). Common European framework of reference for languages. http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf. Zugegriffen: 28. Okt 2012.
- Ercikan, K., & Julian, M. (2002). Classification accuracy of assigning student performance to proficiency levels: Guidelines for assessment design. *Applied Measurement in Education, 15*, 269–294.
- Fishman, B. J., Marx, R. W., & Best, S. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education, 19*, 643–658.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett-Kallmeyer.
- KMK (2004) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004). *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss*. Neuwied: Wolters-Kluwer.
- KMK (2006) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. Neuwied: Wolters-Kluwer.
- KMK (2010) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2010). *Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung*. Köln: Wolters-Kluwer.
- Leucht, M., Frenzel, J., & Pöhlmann, C. (2010). Der Ländervergleich im Fach Englisch. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 97–104). Münster: Waxmann.
- Leucht, M., Tiffin-Richards, S., Vock, M., Pant, H. A., & Köller, O. (2012). Diagnostische Kompetenz von Englischlehrkräften. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 44*, 163–177.
- Leutner, D., Fleischer, J., Spoden, C., & Wirth, J. (2007). Landesweite Lernstandserhebungen zwischen Bildungsmonitoring und Individualdiagnostik. In I. Gogolin, M. Prenzel & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (8. Sonderheft der Zeitschrift für Erziehungswissenschaft, S. 149–167). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Maier, U. (2008). Vergleichsarbeiten im Vergleich – Akzeptanz und wahrgenommener Nutzen standardbasierter Leistungsmessungen in Baden-Württemberg und Thüringen. *Zeitschrift für Erziehungswissenschaft, 11*(3), 453–474.
- Pant, H. A., Emmrich, R., Harych, P., & Kuhl, P. (2011). Leistungsüberprüfungen durch Schulleistungsstudien und Vergleichsarbeiten. In W. Sacher & F. Winter (Hrsg.), *Diagnose und Beurteilung von Schülerleistungen – Grundlagen und Reformansätze* (S. 123–142). Baltmannsweiler: Schneider Hohengehren.
- Pant, H. A., Böhme, K., & Köller, O. (2012). Das Kompetenzkonzept der Bildungsstandards und die Entwicklung von Kompetenzstufenmodellen. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik* (S. 49–55). Münster: Waxmann.

- Staphorsius, G., & Krom, R. (2008). Das Schüler-Monitoring-System in den Niederlanden. In N. Berkemeyer, W. Bos, V. Manitius & K. Müthing (Hrsg.), *Unterrichtsentwicklung in Netzwerken. Konzeptionen, Befunde, Perspektiven* (S. 151–182). Münster: Waxmann.
- Wößmann, L. (2005). The effect heterogeneity of central exams: Evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics*, 13, 143–169.
- Wößmann, L. (2006). Bildungspolitische Lehren aus den internationalen Schülertests: Wettbewerb, Autonomie und externe Leistungsüberprüfung. *Perspektiven der Wirtschaftspolitik*, 7, 417–444.