

Nutzung und Effekte zentraler Abschlussprüfungen und standardbasierter Schulleistungstests als Instrumente der Neuen Steuerung

Felicitas Thiel · Bettina Hannover · Hans Anand Pant

Zentrale Abschlussprüfungen und standardbasierte Tests, wie z. B. Vergleichsarbeiten oder Lernstandserhebungen, stellen wichtige Instrumente des in den vergangenen 15 Jahren implementierten neuen Steuerungsmodells dar. Aus steuerungstheoretischer Sicht werden mit diesen Instrumenten die Voraussetzungen für die Rechenschaftslegung (Accountability) und für erweiterte Gestaltungsspielräume von Schulen (Autonomy) geschaffen. Auf der Grundlage der Konkretisierung und Operationalisierung erwarteter Lernergebnisse in Testaufgaben können Leistungen verglichen werden. Dabei geht es nicht nur um die Feststellung und Beurteilung von Schülerleistungen: Standardbasierte Tests erlauben, wenn die Ergebnisse in sogenannten „fairen Vergleichen“ aufbereitet werden, auch die Evaluation von Klassen und Schulen sowie von Schulamtsbezirken und Regionen.

Zentrale Abschlussprüfungen dienen in erster Linie der Zertifizierung einer Leistung in Kernfächern am Ende relevanter Phasen der Bildungsbiographie (Hauptschulabschluss, mittlerer Schulabschluss, Abitur). Sie definieren Anschlussoptionen im Bildungs- und Erwerbssystem und wirken aus diesem Grunde individuell selektiv. Im Unterschied dazu sollen Vergleichsarbeiten vor allem Prozesse der Schul- und Unterrichtsentwicklung durch Feedback über den Leistungsstand von Schulklassen unterstützen (KMK 2010, 2012). Auf der Grundlage der Leistungsergebnisse der Schülerinnen und Schüler wird im Idealfall Handlungsbedarf identifiziert und über die Implementation von Qualitätssiche-

Online publiziert: 08.02.2014

© Springer Fachmedien Wiesbaden 2014

Prof. Dr. F. Thiel (✉)

AB Schulpädagogik/Schulentwicklungsforschung, Freie Universität Berlin,

Habelschwerdter Allee 45, 14195 Berlin, Deutschland

E-Mail: felicitas.thiel@fu-berlin.de

Prof. Dr. B. Hannover

AB Schul- und Unterrichtsforschung, Freie Universität Berlin,

Habelschwerdter Allee 45, 14195 Berlin, Deutschland

E-Mail: bettina.hannover@fu-berlin.de

Prof. Dr. H. A. Pant

Institut für Qualitätsentwicklung im Bildungswesen, Humboldt-Universität zu Berlin,

Unter den Linden 6, 10099 Berlin, Deutschland

E-Mail: iqboffice@iqb.hu-berlin.de

rungs- und -entwicklungsmaßnahmen auf der Ebene der Schule, der Fachgruppe und des Unterrichts entschieden.

Während zentrale Abschlussprüfungen in einigen deutschen Bundesländern eine lange Tradition haben, wie z. B. das Zentralabitur in Bayern, betraten alle Bundesländer mit der Einführung von Bildungsstandards und standardbasierten Tests Neuland. Wichtige Impulse bei der Implementation der neuen Testverfahren kamen aus den angelsächsischen Ländern, die zum damaligen Zeitpunkt bereits über langjährige Erfahrungen mit diesen Formaten verfügten. Das in den angelsächsischen Ländern implementierte sogenannte High-Stakes-Modell, in dem Testdaten die Grundlage für durchaus drastische Entscheidungen bilden, wie etwa über den Fortbestand von Schulen oder die Weiterbeschäftigung von Lehrkräften, unterscheidet sich allerdings deutlich von dem deutschen Low-Stakes-Modell, in dem die Nutzung von Leistungsdaten in hohem Maße den Lehrkräften und Schulen selbst überantwortet ist (vgl. Gärtner 2013; Gärtner et al. 2009).

Die Frage, inwiefern zentrale Abschlussprüfungen und standardbasierte Tests tatsächlich zu einer Steigerung der Lernergebnisse der Schülerinnen und Schüler und zur Unterrichts- und Schulentwicklung beitragen, kann derzeit nicht empirisch gesichert beantwortet werden. So finden manche Studien positive Effekte standardisierter Tests und zentraler Abschlussprüfungen auf Schülerleistungen (Bishop 1997; Wößmann 2003; Jürges et al. 2005; Wößmann et al. 2009), andere hingegen nicht (vgl. im Überblick Maag Merki 2010). Zudem wird über eine ganze Reihe nicht-intendierter Effekte von Testsystemen berichtet (Amrein-Beardsley et al. 2010; Koretz 2011). Befunde aus den angelsächsischen Ländern können nicht direkt auf das deutsche System der Vergleichsarbeiten und zentralen Prüfungen übertragen werden, da standardbasierte Tests in High- und Low-Stakes-Systemen unterschiedliche Funktionen erfüllen (vgl. zusammenfassend Richter et al., im Druck).

Vergleichsweise gut entwickelt ist der Forschungsstand dagegen zur Frage der Rezeption und Nutzung der Ergebnisse zentraler Abschlussprüfungen und standardbasierter Tests für Entwicklungsmaßnahmen. So wurde in den letzten Jahren eine Reihe von Studien vorgelegt, die Effekte zentraler Prüfungen auf Unterrichtsprozesse (Maag Merki et al. 2008; Maag Merki 2010) und Bedingungen sowie Barrieren der Nutzung von Vergleichsarbeitsdaten für die Unterrichtsentwicklung untersuchen (z. B. Koch et al. 2006; Maier und Kuper 2012; Groß Ophoff 2013). Kaum Beachtung fand in diesen Studien bisher allerdings die Ebene der Schule, sprich die Nutzung von Daten für Schulentwicklung oder Personalentwicklung. Dasselbe gilt für die Untersuchung nachhaltiger Effekte zentraler Abschlussprüfungen auf die Beschäftigungssituation am Arbeitsmarkt.

Diese unterforschten Fragestellungen werden im vorliegenden Themenheft in den Blick genommen und in zwei Beiträgen zu zentralen Abschlussprüfungen und einem Beitrag zur Nutzung von Vergleichsarbeiten näher untersucht.

Im Aufsatz von Klein, Krüger, Kühn und van Ackeren wird ein systematischer Überblick über die internationale Befundlage zu zentralen Abschlussprüfungen gegeben. Erstmals werden Studien unterschiedlicher Forschungsrichtungen, wie Schuleffektivitätsforschung, Unterrichtsforschung und bildungsökonomische Forschung, im Zusammenhang diskutiert. Dabei wird der immer wieder angemahnten Differenzierung im Hinblick auf Kontextbedingungen und Untersuchungsperspektiven Rechnung getragen. Genauer ordnen Klein et al. die internationalen Befunde zu zentralen Prüfungen nach

den Ebenen Bildungspolitik/Bildungsadministration, Schule, Unterricht sowie individuelle Lernerträge. Sie diskutieren intendierte und nicht-intendierte Effekte dieses Steuerungsinstruments und kommen zu dem Schluss, dass zukünftige Untersuchungen länder-, fach- und schulformspezifische Unterschiede der Prüfungsformate stärker berücksichtigen sollten.

Piopiunik, Schwerdt und Wößmann beschäftigen sich in ihrer Studie mit der in der deutschsprachigen Bildungsforschung bislang vernachlässigten Frage, welche langfristigen Effekte zentrale Abschlussprüfungen auf den Arbeitsmarkterfolg haben. Auf der Basis agenturtheoretischer Überlegungen zur Signalwirkung von Abschlussprüfungen können sie zeigen, dass zentrale Prüfungen in Hauptschule und Gymnasium (Abitur) mit einem geringeren Risiko der Arbeitslosigkeit sowie einem höheren Einkommen einhergehen. Mit dem Nachweis distaler Effekte zentraler Abschlussprüfungen wird die Befundlage zu proximalen Wirkungen zentraler Abschlussprüfungen auf Schulleistungen um einen wichtigen Aspekt ergänzt.

Bach, Wurster, Thillmann, Pant und Thiel untersuchen in ihrem Beitrag die Nutzung von VERA-Daten auf der Schulebene im Rahmen eines Low-Stakes-Systems für Maßnahmen der Personalentwicklung. Sie nehmen dabei sowohl Personalentwicklungsaktivitäten von Schulleitungen als auch von Fachkonferenzen in den Blick. Am Beispiel eines Bundeslands zeigen Bach et al., dass die Nutzung von VERA-Daten für die Personalentwicklung durch Schulleitungen stark davon abhängt, ob diese mit dem Instrument eine hohe subjektive Nützlichkeit verbinden. Die Nutzung der VERA-Ergebnisse durch Fachkonferenzen steht dagegen im Zusammenhang mit einem datenfreundlichen Führungsstil der Schulleitung sowie einer kollaborativen Auswertung der Daten.

Literatur

- Amrein-Beardsley, A., Berliner, D. C., & Rideau, S. (2010). Cheating in the first, second, and third degree: Educators' responses to high-stakes testing. *Education Policy Analysis Archives, 18*, 1–31.
- Bishop, J. H. (1997). The effect of national standards and curriculum-based exams on achievement. *American Economic Review, 87*, 260–264.
- Gärtner, H. (2013). Zum Verhältnis von interner und externer Evaluation im Schulsystem. *Zeitschrift für Erziehungswissenschaft, 16*(4), 693–712.
- Gärtner, H., Hüsemann, D., & Pant, H. A. (2009). Wirkungen von Schulinspektion aus Sicht betroffener Schulleitungen. Die Brandenburger Schulleiterbefragung. *Empirische Pädagogik, 23*, 1–18.
- Groß Ophoff, J. (2013). *Lernstandserhebungen: Reflexion und Nutzung*. Münster: Waxmann.
- Jürges, H., Schneider, K., & Büchel, F. (2005). The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany. *Journal of the European Economic Association, 3*, 1134–1155.
- KMK (2010). *Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung*. Köln: Wolters Kluwer.
- KMK (2012). *Vereinbarung zur Weiterentwicklung von VERA*. Berlin: KMK.

- Koch, U., Groß Ophoff, J., Hosenfeld, I., & Helmke, A. (2006). Von der Evaluation zur Schul- und Unterrichtsentwicklung – Ergebnisse der Lehrerbefragungen zur Auseinandersetzung mit den VERA-Rückmeldungen. In F. Eder, A. Gastager & F. Hofmann (Hrsg.), *Qualität durch Standards? Beiträge zur 67. Tagung der Arbeitsgruppe der Empirischen Bildungsforschung (AEPF), Salzburg* (S. 187–199). Münster: Waxmann.
- Koretz, D. (2011). Lessons from test-based education reform in the U.S. In I. Gogolin, J. Baumert & A. Scheunpflug (Hrsg.), *Umbau des Bildungswesens. Bildungspolitische Großreformprojekte und ihre Effekte* (13. Beiheft der Zeitschrift für Erziehungswissenschaft, S. 9–24). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Maag Merki, K., Klieme, E., & Holmeier, M. (2008). Unterrichtsgestaltung unter den Bedingungen zentraler Abiturprüfungen. Eine Analyse auf Schulebene mittels Latent Class Analysen. *Zeitschrift für Pädagogik*, 54(6), 791–809.
- Maag Merki, K. (2010). Theoretische und empirische Analysen der Effektivität von Bildungsstandards, standardbezogenen Lernstandserhebungen und zentralen Abschlussprüfungen. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 145–170). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Maier, U., & Kuper, H. (2012). Vergleichsarbeiten als Instrumente der Qualitätsentwicklung an Schulen. Überblick zum Forschungsstand. *Die Deutsche Schule*, 104(1), 88–99.
- Richter, D., Böhme, K., Becker, M., Pant, H. A., & Stanat, P. (im Druck). Überzeugungen von Lehrkräften zu den Funktionen von Vergleichsarbeiten: Zusammenhänge zu Veränderungen im Unterricht und den Kompetenzen der Schülerinnen und Schüler. *Zeitschrift für Pädagogik*.
- Wößmann, L. (2003). Schooling resources, educational institutions and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics*, 65, 117–170.
- Wößmann, L., Luedemann, E., Schuetz, G., & West, M. R. (2009). *School accountability, autonomy and choice around the world*. Cheltenham, Northampton: Edward Elgar.