

Wann kann Kompetenzdiagnostik negative Auswirkungen haben?

Nele McElvany · Camilla Rjosk

Zusammenfassung: Der vorliegende Beitrag setzt sich mit möglichen negativen Auswirkungen systematischer externer Kompetenzdiagnostik auf verschiedenen Ebenen des Bildungssystems auseinander. Behandelt werden Quellen fehlerhafter Interpretationen von Rückmeldungen aus Kompetenzdiagnostik-Maßnahmen, mögliche Folgen der Rückmeldungen und Maßnahmen für verschiedene Akteure/-innen im Bildungssystem (Lehrkräfte, Lernende und Bildungsadministration) sowie mögliche negative Auswirkungen systematischer Kompetenzdiagnostik für die Lehr- und Lernkultur. Der Beitrag schließt mit Überlegungen zu Ansatzpunkten zur Vermeidung beziehungsweise Verringerung der diskutierten negativen Auswirkungen.

Schlüsselwörter: Kompetenzdiagnostik · Negative Folgen · Rückmeldungen · Teaching to the test

When may competence assessment have negative effects?

Abstract: This paper addresses potential negative effects of systematic external competence assessment on different levels of the education system. We discuss possible sources of misinterpretations of competence assessment feedback, potential consequences of feedback and assessment for different education stakeholders (teachers, students, and educational administration) as well as negative effects of systematic competence assessment on the culture of teaching and learning. The paper concludes with some reflections on possibilities to prevent or reduce the discussed negative effects.

Keywords: Competence assessment · Feedback · Negative effects · Teaching to the test

© Springer Fachmedien Wiesbaden 2013

Prof. Dr. N. McElvany (✉)
Institut für Schulentwicklungsforschung (IFS), TU Dortmund,
Vogelpothsweg 78, 44227 Dortmund, Deutschland
E-Mail: mcelvany@ifs.tu-dortmund.de

Dipl.-Psy. C. Rjosk (✉)
Institut zur Qualitätsentwicklung im Bildungswesen, Humboldt-Universität zu Berlin,
Unter den Linden 6, 10099 Berlin, Deutschland
E-Mail: camilla.rjosk@iqb.hu-berlin.de

Insbesondere im Zuge der ersten PISA 2000-Befunde, die auf suboptimale Leistungsergebnisse von Schülerinnen und Schülern im deutschen Bildungssystem verwiesen (vgl. Deutsches PISA-Konsortium 2001), gewann im Bildungssektor das Bildungsmonitoring verstärkte Bedeutung. Die damit verbundene Output-Orientierung manifestiert sich unter anderem durch eine umfangreiche, an neu definierten Bildungsstandards orientierte systematische Kompetenzdiagnostik (auf Systemebene „Monitoring“), in deren Kontext gleichzeitig auch effektive Maßnahmen zur Verbesserung des deutschen Bildungswesens ermittelt werden sollen (vgl. z. B. Klieme et al. 2003). Mit der verstärkten und neu orientierten Kompetenzmessung bei Lernenden sind viele positiv bewertete Ziele verbunden (Systemmonitoring, Steuerungswissen, Unterrichtsentwicklung, Diagnostik, Evaluation; vgl. Berkemeyer und Bos 2009). Weniger bedacht und diskutiert wird bisher jedoch, dass es auch eine Reihe von Bedingungen gibt, unter denen eine solche extern gesteuerte Kompetenzdiagnostik negative Auswirkungen haben kann. Die Auswirkungen können sich dabei auf alle beteiligten Akteure des Bildungssystems beziehen: Lehrende, Lernende, Schulen und Bildungsadministration. Im Folgenden werden drei ausgewählte Themenfelder im Kontext von systematischer externer Kompetenzdiagnostik näher beleuchtet: 1) die Interpretation von Rückmeldungen aus Kompetenzdiagnostik-Maßnahmen, 2) die sich daran anschließenden Folgen beziehungsweise der Umgang mit Rückmeldungen auf den Ebenen der Lehrkräfte, Schüler/-innen und der Bildungsadministration und 3) die grundsätzlicheren Folgen für die Lehr- und Lernkultur.

1 Interpretation von Rückmeldungen aus Kompetenzdiagnostik-Maßnahmen

In Bezug auf die Rückmeldungen aus schulübergreifenden Maßnahmen der Kompetenzdiagnostik (vgl. Watermann et al. 2003) können sich fehlerhafte Interpretationen im öffentlichen Diskurs oder bei Schulen, Lehrenden und Lernenden ergeben. Derartige Interpretationen mit anzunehmenden unerwünschten Folgen entstehen beispielsweise durch die Überinterpretation von Rangunterschieden von Schulen oder Klassen gegenüber den tatsächlich gegebenen metrischen Abständen. Fehlerhafte Interpretationen können auch von Seiten der Durchführenden der Kompetenzdiagnostik-Maßnahmen und den zugehörigen Rückmeldungen für Schulen, Klassen oder Lehrkräfte verursacht oder begünstigt werden, wenn im Rahmen der Rückmeldungen die vorhandenen Ergebnisse überinterpretiert werden. Dies ist beispielsweise der Fall, wenn verzerrende Darstellungsformen gewählt werden, Signifikanzen ohne Effektstärken berichtet werden oder die Auswahl rückgemeldeter Daten einseitig ist.

Grundsätzliche Probleme bezüglich der Interpretation der Kompetenzdiagnostik-Ergebnisse können somit vor allem in einer mangelnden „Assessment Literacy“ (vgl. z. B. Stiggins 1991) und dem Experten-Laien-Problem (vgl. z. B. Bromme et al. 2004) in der Kommunikation von Testergebnissen als Quelle negativer Auswirkungen von Kompetenzdiagnostik gesehen werden.

Eine weitere Quelle von Fehlinterpretationen bezieht sich auf die Ebene von Leistungsergebnissen, über die durch die Kompetenzdiagnostik Aussagen möglich sind, beziehungsweise das diagnostische Ziel. Eine besondere Problematik liegt darin, Kompetenzdiagnostik im Rahmen von System- oder Schulmonitoring-Tests wie beispiels-

weise PISA oder IGLU fälschlicherweise für Individualdiagnostik nutzen zu wollen – ein diagnostisches Ziel, für welches diese Art von Tests weder intendiert noch konzipiert wurde (vgl. Bos und Voss 2008). Ebenso sind die Large-Scale-Untersuchungen im Rahmen internationaler Vergleichsstudien (z. B. PISA, TIMSS, IGLU/PIRLS) wie auch die flächendeckenden VERA- oder die bildungsstandardüberprüfenden Erhebungen darauf ausgerichtet, im Sinne von Assessment und Accountability Rückmeldungen zum Leistungsstand und zur Erreichung bestimmter Standards zu geben, können aber nicht systematische Untersuchungen zur Unterrichtsentwicklung ersetzen.

2 Folgen der Kompetenzdiagnostik und der Rückmeldungen auf den Ebenen von Lehrenden, Lernenden und Bildungsadministration

Negative Auswirkungen von externer systematischer Kompetenzdiagnostik können sich in dem Umgang mit den Befunden zeigen, der sich an die Interpretation von Rückmeldungen anschließt. Dies kann die affektive, kognitive und Verhaltensebene der Akteurinnen und Akteure des Bildungssystems betreffen.

Auf Seiten von Lehrkräften können beispielsweise Frustrationen entstehen, wenn trotz Engagements im schulischen Alltag im Rahmen der offiziellen Kompetenzdiagnostik keine oder kaum Kompetenzverbesserung auf Schülerseite feststellbar ist und diese Situation im Folgenden nicht weiter von außen konstruktiv begleitet wird. Außerdem können schul- oder allgemein-öffentliche Rankings aufgrund von Ergebnissen im Rahmen der Lernstanderhebungen Ängste bei Lehrerinnen und Lehrern hervorrufen, den an sie gesetzten Ansprüchen nicht gerecht zu werden. Dies kann unerwünschte Reaktionen der Lehrkräfte induzieren, wie einen einseitigen Fokus des Unterrichts auf die Testinhalte oder Manipulationen bei der Testvorbereitung, -durchführung oder -auswertung (zur Bewertung von Rückmeldungen durch Lehrkräfte vgl. z. B. Bonsen et al. 2006).

Allgemein ist anzunehmen, dass eine teilweise durchaus gegebene, übermäßige Testbelastung der Schulen infolge mangelnder Koordination unterschiedlicher Testprogramme zu einer Testmüdigkeit führen kann, die einem produktiven Umgang mit Maßnahmen der Kompetenzdiagnostik entgegensteht. Vielfache Testungen, aber insbesondere auch die subjektive und/oder teilweise kollegiums- oder schulintern geteilte Wahrnehmung mangelnder Nützlichkeit der Kompetenzdiagnostik für die schulische Alltagspraxis kann zu einer generellen Infragestellung dieser Maßnahmen und in der Folge zu einem suboptimalen Umgang mit ihnen führen.

Auf Ebene der Schüler/-innen kann zunächst festgehalten werden, dass die reine Rückmeldung eines Punktwertes zum individuellen Abschneiden nicht den Kriterien entspricht, die im Rahmen der Feedback-Forschung als dienlich identifiziert wurden (vgl. Kluger und DeNisi 1996). So könnte es vor allem dann unerwünschte Reaktionen geben, wenn Lernende negative Ergebnismrückmeldungen im Rahmen der Kompetenzdiagnostik nur als Scheitern und Bestätigung eigener Unfähigkeit wahrnehmen und so zukünftige Lern- und Testsituationen nachteilig beeinflusst werden. Eine derartige Wahrnehmung stellt eine grundsätzliche Problematik in Situationen von Leistungsrückmeldungen dar (zu Kausalattributionen vgl. zusammenfassend Stiensmeier-Pelster und Schwinger 2008). Pädagogische Begleitung bei der lerndienlichen Kausalattribution erreichter Testergebnisse und

entsprechende Förderangebote dürften nicht nur, aber gerade für leistungsschwächere Lernende ein wesentlicher Punkt sein, um möglichen negativen Auswirkungen auf die Lernmotivation und Performanz zu begegnen.

Der Umgang mit Ergebnismeldungen auf der Ebene der Bildungsadministration könnte unter anderem insbesondere dann problematisch sein, wenn die vergleichende Kompetenzdiagnostik als Evaluationsinstrument eingesetzt wird, in dessen Folge bei schlechten Schulergebnissen Ressourcen gekürzt werden, ohne sinnvolle Hilfestellungen zur Verbesserung zu gewähren (vgl. Fuhrman und Elmore 2004).

3 Folgen für die Lehr- und Lernkultur

Neben den im vorherigen Abschnitt aufgeführten möglichen konkreten Folgen systematischer Kompetenzdiagnostik für die einzelnen Akteurinnen und Akteure im Bildungssystem können auch weitreichende problematische Auswirkungen in der allgemeinen Lehr- und Lernkultur entstehen: Wenn der Fokus von Wissenschaft, Bildungsadministration und pädagogischer Praxis nahezu ausschließlich auf den Kompetenzerwerb im kognitiven Sinne eingeengt wird (u. a. Lesekompetenz, Mathematik, Naturwissenschaften), werden andere bedeutsame Gesichtspunkte der Entwicklung von Kindern und Jugendlichen im schulischen Kontext wie beispielsweise motivationale und psycho-soziale Aspekte vernachlässigt. Bezogen auf die Lehrkultur könnten komplexe pädagogische Ziele der Engführung von pädagogischen Zielsystemen weichen, wenn „gute Schule“ bzw. „guter Unterricht“ nur anhand von Kompetenzzuwächsen im Rahmen der Kompetenzdiagnostik festgemacht werden. Diese Problematik verweist auf das vor allem im US-amerikanischen Kontext viel diskutierte „Teaching [Learning] to the test“ – d. h. die Einschränkung auf diejenigen Lernbereiche, für die kompetenzdiagnostische Instrumente verfügbar sind und eingesetzt werden, während andere Lern- und Kompetenzbereiche nicht mehr adäquat gefördert werden (vgl. z. B. Cheng et al. 2004; Herman 2004). Eine analoge Folge von extern gesteuerter, an aktuellen Bildungsstandards orientierter Kompetenzdiagnostik kann gleichfalls für Ziele der Lernkultur angenommen werden: Die Bedeutung von Performance-Goals, also des Lernens, um einen bestimmten Wert in den vergleichenden Leistungstests zu erreichen, dürfte steigen, während Mastery-Goals, also das Lernen, um die eigene Kompetenz zu erhöhen, an Bedeutung im Lernprozess verlieren könnten. Dieses generelle Phänomen im Kontext von (high-stakes) Tests im Bildungssystem wird durch die Aufmerksamkeit und Bedeutung von übergreifender Kompetenzdiagnostik vermutlich verstärkt (zu Performance- und Mastery-Goals vgl. u. a. Pintrich 2000; Rawsthorne und Elliot 1999). Schließlich kann der Druck, evaluativ vergleichende Kompetenzdiagnostik durchzuführen, bei der pädagogischen Arbeit an Schulen zu einem Problem werden, wenn aufgrund vorschneller Kompetenzdiagnostik neuen pädagogischen Programmen nicht ausreichend Zeit gegeben wird, um ihre Entwicklung und Einflüsse zu entfalten. Dies betrifft die generell erwünschten Maßnahmen schulischer Innovation.

Aus den in diesem Beitrag diskutierten möglichen negativen Auswirkungen von Kompetenzdiagnostik können eine Reihe von Ansatzpunkten extrahiert werden, wie diesen Herausforderungen begegnet werden kann: Zur Vermeidung bzw. Verringerung negativer

Auswirkungen von Kompetenzdiagnostik können demnach unter anderem 1) die Verbesserung der Assessment-Literacy aller Beteiligten, 2) eine weiter optimierte Unterstützung von Lehrkräften sowie Schülerinnen und Schülern beim Umgang mit Rückmeldungen, 3) eine umfassende, kritische Diskussion über die Folgen der schulübergreifend vergleichenden Kompetenzdiagnostik für die Lehr- und Lernkultur an den Schulen sowie 4) ein verstärkter Fokus auf die effektive und konstruktive Nutzung der Ergebnisse aus der Kompetenzdiagnostik beitragen. Darauf aufbauend kann insbesondere dahingehend weiterer Forschungsbedarf identifiziert werden, wie der Umgang mit den diagnostischen Ergebnissen auf den verschiedenen Ebenen – Gesamtinformation über die Evaluation, individuelle Rückmeldungen, Schulrückmeldungen oder Rückmeldungen an die politisch systemische Ebene – mit ihren verschiedenen Interessen und Herausforderungen am besten zu gestalten ist.

Danksagung: Diese Veröffentlichung wurde ermöglicht durch Sachbeihilfen der Deutschen Forschungsgemeinschaft (Kennz.: BA 1461/7-1, BA 146/8-1, MC 67/7-2) im Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (SPP 1293).

Literatur

- Berkemeyer, N., & Bos, W. (2009). Professionalisierung im Spannungsfeld von interner und externer Evaluation. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrprofessionalität – Bedingungen, Genese, Wirkungen und ihre Messung* (S. 529–542). Weinheim: Beltz.
- Bonsen, M., Büchter, A., & Peek, R. (2006). Datengestützte Schul- und Unterrichtsentwicklung. Bewertungen der Lernstandserhebungen in NRW durch Lehrerinnen und Lehrer. *Jahrbuch der Schulentwicklung, 14*, 125–148.
- Bos, W., & Voss, A. (2008). Empirische Schulentwicklung auf Grundlage von Lernstandserhebung. Ein Plädoyer für einen reflektierten Umgang mit Ergebnissen aus Leistungstests. *Die Deutsche Schule, 100*, 449–458.
- Bromme, R., Jucks, R., & Rambow, R. (2004). Experten-Laien-Kommunikation im Wissensmanagement. In G. Reinmann & H. Mandl (Hrsg.), *Psychologie des Wissensmanagements. Perspektiven, Theorien und Methoden* (S. 176–188). Göttingen: Hogrefe.
- Cheng, L., Watanabe, Y., & Curtis, A. (2004). *Washback in language testing. Research contexts and methods*. Mahwah: Erlbaum.
- Deutsches PISA-Konsortium. (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Fuhrman, S. H., & Elmore, R. F. (2004). *Redesigning accountability systems for education*. New York: Teachers College Press.
- Herman, J. L. (2004). The effects of testing on instruction. In S. H. Fuhrman & R. F. Elmore (Hrsg.), *Redesigning accountability systems for education* (S. 141–166). New York: Teachers College Press.
- 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*. Berlin: BMBF.

- Kluger, A. N., & DeNisi, A. (1996). The Effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*, 254–284.
- Pintrich, P. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, *92*, 544–555.
- Rawsthorne, L., & Elliot, A. J. (1999). Achievement goals and intrinsic motivation: A meta-analytic review. *Personality and Social Psychology Review*, *3*, 326–344.
- Stiensmeier-Pelster, J., & Schwinger, M. (2008). Kausalattribution. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 74–83). Göttingen: Hogrefe.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, *72*, 534–539.
- Watermann, R., Stanat, P., Kunter, M., Klieme, E., & Baumert, J. (2003). Schulrückmeldungen im Rahmen von Schulleistungsuntersuchungen: Das Disseminationskonzept von PISA-2000. *Zeitschrift für Pädagogik*, *49*, 92–111.