Editorial

Felicitas Thiel · Jasmin Tarkian · Harm Kuper

Online publiziert: 25. Juli 2016
© Springer Fachmedien Wiesbaden 2016

Nahezu alle Bundesländer haben in den vergangenen 10–20 Jahren Maßnahmen externer Evaluation oder Schulinspektorate eingeführt – teilweise sogar bereits wieder abgeschafft (Döbert et al. 2008; Dederling 2015; Wurster 2016). Neben Vergleichsarbeiten und interner Evaluation zählt die deutsche Kultusministerkonferenz in ihrem überarbeiteten Strategiepapier zum Bildungsmonitoring aus dem Jahr 2015 (KMK 2015) externe Evaluationen zu den zentralen Instrumenten der Sicherung und Entwicklung von Bildungsqualität auf der Ebene der Schule.

Mit der Einführung von Schulinspektionen waren mehrere Ziele verbunden. Die Beurteilung der Lern- und Leistungsergebnisse in den Einzelschulen sowie die Beurteilung der Prozessqualität auf der Ebene von Schule und Unterricht durch die Inspektorate geschieht a) mit dem Ziel der Rechenschaftslegung anhand spezifizierter Qualitätsindikatoren, b) mit dem Ziel der Initiierung von Entwicklungsaktivitäten sowie c), ergänzend zu nationalen und internationalen Schulleistungsvergleichsstudien, mit dem Ziel des Systemmonitorings.

Insbesondere hinsichtlich der Entwicklungsfunktion war die Implementation der Schulinspektion mit hohen Erwartungen verbunden, ermöglicht doch der Blick auf die Prozesse in Schule und Unterricht grundsätzlich nicht nur die Feststellung von

Prof. Dr. F. Thiel (✉) · J. Tarkian
Arbeitsbereich Schulpädagogik/Schulentwicklungsforschung, Freie Universität Berlin,
Habelschwerdter Allee 45, 14195 Berlin, Deutschland
E-Mail: felicitas.thiel@fu-berlin.de

J. Tarkian
E-Mail: jasmin.tarkian@fu-berlin.de

Prof. Dr. H. Kuper
Arbeitsbereich Weiterbildung und Bildungsmanagement, Freie Universität Berlin,
Arnimallee 12, 14195 Berlin, Deutschland
E-Mail: harm.kuper@fu-berlin.de

Entwicklungsbedarfen, sondern auch die Identifikation von Ansatzpunkten zur Verbesserung der schulischen Arbeit.

Externe Evaluation bezieht sich auf eine Vielzahl von Aspekten der Qualität von Schule und Unterricht. In der Regel werden die Qualitätsdimensionen aus den Perspektiven unterschiedlicher Akteure und mit unterschiedlichen Instrumenten erfasst. Qualitätsdimensionen und -indikatoren werden aus sogenannten Qualitätsrahmen abgeleitet. Die Qualitätsrahmen der Bundesländer unterscheiden sich ebenso wie die Verfahren der Evaluation oder die Auswahl und Qualifizierung der Inspektorenteams teilweise deutlich voneinander (Dederling 2015). Auch hinsichtlich der Rückmeldeformate, der einzuleitenden Maßnahmen nach Feststellung von Qualitätsmängeln, der Zugänglichkeit der Befunde für unterschiedliche Akteursgruppen sowie der vorgehaltenen Unterstützungsangebote sind große Unterschiede zwischen den Bundesländern festzustellen (Dederling 2015).

Kuhn weist darauf hin, dass die Inspektionsverfahren in einem vergleichsweise kurzen Zeitraum flächendeckend implementiert wurden, „allerdings anders als bei der Entwicklung nationaler Bildungsstandards ohne explizite Anbindung an bildungswissenschaftliche Forschungen oder Institutionen der empirischen Bildungsforschung“ (Kuhn 2015, S. 81). Dies betrifft sowohl die Definition und Operationalisierung von Qualitätsdimensionen als auch die Entwicklung und Validierung von Evaluationsinstrumenten. Nur in wenigen Bundesländern, in denen Schulinspektionen in Qualitätsinstituten verortet wurden, die mit einem Forschungsauftrag ausgestattet sind, erfolgte in den vergangenen Jahren eine systematische Prüfung der Effekte sowie eine (nachholende) Validierung der Verfahren, die dem *state of the art* sozialwissenschaftlicher Methoden entspricht.

Dass die Entwicklung der Qualitätsrahmen sowie der Evaluationsinstrumente ohne maßgebliche Beteiligung der erziehungswissenschaftlichen Forschung stattgefunden hat, ist auch durch die lange Zurückhaltung weiter Teile der deutschen Schulpädagogik in der empirischen Forschung bedingt. Zugleich waren innerhalb der sich etablierenden Bildungsforschung Fragen der Unterrichts- und vor allem der Schuleffektivitätsforschung zunächst noch eher Randthemen. Die internationale Schuleffektivitätsforschung fand in den deutschsprachigen Ländern zum Zeitpunkt der Implementation der Schulinspektion jedenfalls wenig Resonanz und auch Studien zur Unterrichtsqualität waren damals im Vergleich zur Übergangsforschung oder der Forschung zur Entwicklung von Kompetenzmodellen und Kompetenztests eher unterrepräsentiert.

Die seit einigen Jahren beobachtbare (nachholende) Verwissenschaftlichung der Schulinspektion ist aus der Perspektive der Wissenschaftssoziologie betrachtet ein äußerst interessantes Phänomen, das im Unterschied zu den internationalen Schulleitungsstudien (Aljets 2015) bislang allerdings noch kaum Beachtung gefunden hat. Die These, dass sich im Fall der Schulinspektion ein anderes Muster der Verwissenschaftlichung eines Felds der Bildungspolitik zeigt als bei der Implementierung von Kompetenztests, kann angesichts der Berichte von Insidern (Maritzen 2006; Kuhn 2015) einige Plausibilität für sich in Anspruch nehmen.

Inzwischen liegen einige Studien zur Schulinspektion vor, die teilweise von Mitarbeiterinnen und Mitarbeitern der Landesinstitute, die Schulinspektionen verantworten, durchgeführt wurden. Der Forschungsstand zur Schulinspektion lässt sich

kursorisch wie folgt zusammenfassen (vgl. dazu ausführlich Altrichter und Kemethofer in diesem Heft).

- Zunächst beschäftigte sich die Forschung zur Schulinspektion in erster Linie mit Fragen der Einstellung zum und der Akzeptanz des Verfahrens.
- Das Forschungsinteresse galt außerdem bereits früh den durch Schulinspektion ausgelösten Entwicklungsaktivitäten von Schulen, die in der Regel über Selbstberichte der Schul- oder Fachkonferenzleitungen erfasst wurden.
- Auch wurden einige Untersuchungen zur Rekonstruktion von Nebenfolgen von Inspektionen durchgeführt.
- Fragen zu Effekten von Schulinspektionen auf Schülerleistungen wurden im deutschsprachigen Bereich eher selten bearbeitet. Erst in jüngster Zeit wurden dazu einige Studien vorgelegt. Die Befundlage ist hier allerdings nicht eindeutig.
- Außerdem liegen internationale und nationale Studien vor, die unterschiedliche Inspektionssysteme hinsichtlich ihrer spezifischen Signatur beschreiben und in explorativen Analysen differentielle Effekte dieser Systeminterventionen rekonstruieren. In der Regel wird hier zwischen high-stakes und low-stakes Systemen unterschieden.

Im Vergleich mit der internationalen Forschung zu Schulinspektionen weist die deutschsprachige Forschung einige Leerstellen auf. So wurden Untersuchungen zu Effekten der Schulinspektion lange Zeit vernachlässigt. Außerdem wurden nur in Ausnahmen Anstrengungen einer systematischen Validierung der in Inspektionen eingesetzten Instrumente unternommen. Ein großes Desiderat besteht auch hinsichtlich der theoretischen Unterfütterung der Qualitätsrahmen sowie Versuchen, diese teilweise wildwüchsig entstandenen Konstrukte mit den theoretisch begründeten Konstrukten der Unterrichts- und Schulforschung in Verbindung zu bringen. Was die Perspektive der Governanceforschung auf Schulinspektionen betrifft, besteht ein großes Interesse an der historischen Rekonstruktion und funktionalen Analyse von Inspektoraten im Zusammenspiel mit den traditionellen Aufsichtsstrukturen. Vollständig fehlen bildungsökonomische Studien, die die Effizienz von Inspektionssystemen erfassen.

Das vorliegende Schwerpunktheft adressiert einige der beschriebenen Desiderate und versteht sich als ein Beitrag zur nachholenden Verwissenschaftlichung eines zentralen Steuerungsinstrumentes der deutschen Bildungspolitik.

Herbert Altrichter und David Kemethofer zeichnen in ihrem Stichwortbeitrag zunächst die historischen Entwicklungslinien von Inspektionen in unterschiedlichen gesellschaftlichen Teilsystemen nach und charakterisieren, wie diese spezifische Form bürokratischer Kontrolle mit der Linienverwaltung, im Fall des Schulsystems also mit der traditionellen Schulaufsicht, in Zusammenhang steht. Die Autoren diskutieren außerdem unterschiedliche theoretische Modelle der Schulinspektionsforschung, an denen sich internationale Studien zur Schulinspektion orientieren. Was den Forschungsstand zur Schulinspektion betrifft, sehen Altrichter und Kemethofer sieben Forschungsfelder, die in unterschiedlicher Intensität bearbeitet werden: Einstellung zu und Akzeptanz von Schulinspektionen, Entscheidungsfindung in Schulinspektionen, schulische Reaktionen auf Schulinspektionen, Effekte von Schulinspektionen, Effekte unterschiedlicher Merkmale von Schulinspektionen, un-

erwünschte Nebeneffekte von Schulinspektionen sowie Prozesse und vermittelnde Mechanismen von Schulinspektionen. Bezüglich der Perspektiven der Forschung zur Schulinspektion identifizieren die Autoren einen anhaltend hohen Bedarf an Studien zu Effekten externer Evaluation, sehen vor allem aber Bedarf an einer Weiterentwicklung theoretischer Modelle zur Beschreibung der Prozesse und Mechanismen von Inspektionen.

Der Beitrag von Holger Gärtner unternimmt den Versuch, die Qualitätsindikatoren, die den Qualitätsrahmen und den Instrumenten der Schulinspektion zugrunde liegen, auf theoretisch und empirisch begründete Konstrukte der Schul- und Unterrichtsforschung zu beziehen und mittels Reliabilitätsanalysen sowie konfirmatorischer Faktorenanalysen zu prüfen. Auf dieser Grundlage kann Gärtner nicht nur zeigen, dass die Qualitätsindikatoren der Inspektorate auf die Konstrukte der Schul- und Unterrichtsforschung abbildbar sind, sondern auch, dass die Schulmerkmale in Zusammenhang mit Merkmalen auf der Unterrichtsebene stehen und dass die Qualität der Prozesse auf der Schulebene nicht nur die mittlere Qualität der Prozesse auf der Unterrichtsebene beeinflusst, sondern auch die Streuung zwischen den unterschiedlichen Klassen einer Schule. Qualitativ hochwertige Programme einer Schule scheinen einen homogenisierenden Einfluss auf die Unterrichtsqualität zu haben. Dieser Befund bestätigt organisationstheoretische Annahmen zur Koordinationsfunktion von Programmen. Neben der Frage der theoretischen Unterfütterung der Indikatoren und Instrumente von Inspektoraten adressiert Holger Gärtner in diesem Beitrag auch die Frage der Nutzbarkeit von Schulinspektionsdaten zur Bearbeitung von Fragen der Schul- und Unterrichtsforschung.

Marcus Pietsch, Markus Lücken, Franziska Thonke, Stefan Klitsche und Frank Musekamp präsentieren eine Studie, die auf der Grundlage eines breiten Verständnisses von Validität die Frage der „Interpretier- und Nutzbarkeit von Schulinspektionsergebnissen im Bereich Führung von Schulen“ untersucht. Hier steht nicht nur die Frage im Zentrum, ob bestimmte Instrumente der Schulinspektion in der Lage sind, die Qualität von Schule oder Unterricht zuverlässig zu bewerten, sondern auch die Frage, ob die entsprechenden Rückmeldungen der Daten für den intendierten Zweck effektiv genutzt werden. Pietsch und Kollegen untersuchen dazu beispielhaft den durch die Inspektion erfassten Bereich des Führungsverhaltens der Schulleitung, dem in der Praxis nicht nur eine wichtige Bedeutung für Schulentwicklungsprozesse zukommt, sondern zu dem auch eine lange Forschungstradition existiert, die eine Ableitung differenzierter Hypothesen erlaubt. Pietsch und Kollegen können zeigen, dass die entsprechenden Instrumente der Schulinspektion eine zuverlässige Identifikation von Führungsstilen ermöglichen. Eine zuverlässige Bewertung ihrer Wirksamkeit erweist sich allerdings deshalb als schwierig, weil offensichtlich ein Passungsverhältnis zwischen effektiver Führung und der spezifischen Situation der Schule besteht. Vor diesem Hintergrund kommen Pietsch und Kollegen zum Schluss, dass der zweite Aspekt der Validität, die Nutzbarkeit von Daten, nur eingeschränkt erfüllt ist.

Das Design der Studie von Sebastian Wurster, Tobias Feldhoff und Holger Gärtner folgt der Logik der Evaluation von Systeminterventionen. Zwei Inspektionssysteme, die sich hinsichtlich zentraler Aspekte unterscheiden und deren Ziele hinsichtlich der Frage Stimulierung von Entwicklung oder Kontrolle divergieren, werden bezüglich

der Akzeptanz der Verfahren sowie der Verwendung der Evaluationsergebnisse für Entwicklungsaktivitäten verglichen. Das längsschnittlich angelegte Design erlaubt dabei eine vergleichende Betrachtung von Aktivitäten im Vorfeld und im Nachgang der Inspektion. Zwar lassen sich unterschiedliche Muster von Akzeptanz und Verwendung feststellen, die auch plausibel auf die Unterschiede im Inspektionsverfahren zugerechnet werden können, allerdings werfen die Befunde doch die Frage auf, ob die theoretisch angenommene Unvereinbarkeit von Entwicklung und Kontrolle wirklich begründbar ist oder ob es nicht angemessener ist, davon auszugehen, dass die beiden Steuerungslogiken in unterschiedlicher Weise zusammenwirken.

Literatur

- Aljets, E. (2015). *Der Aufstieg der Empirischen Bildungsforschung. Ein Beitrag zur institutionalistischen Wissenschaftssoziologie*. Wiesbaden: Springer VS.
- Dedering, K. (2015). The same procedure as every time? School inspections and school development in Germany. *Improving Schools*, 18(2), 171–184.
- Döbert, H., Rürup, M., & Dedering, K. (2008). Externe Evaluation in Deutschland. In H. Döbert & K. Dedering (Hrsg.), *Externe Evaluation von Schulen* (S. 63–151). Münster: Waxmann.
- KMK (2015) = Kultusministerkonferenz der Länder. (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring* (Beschluss der KMK 06/2015). https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf. Zugriffen: 06. Juli 2016.
- Kuhn, H.-J. (2015). Schulinspektion in Deutschland – Erfolgsmodell oder neue Illusion? In C. G. Bühren (Hrsg.), *Handbuch Feedback in der Schule* (S. 79–107). Weinheim: Beltz.
- Maritzen, N. (2006). Eine Trendanalyse – Schulinspektion zwischen Aufsicht und Draufsicht. In H. G. Rolff, L. Horster & H. Buchen (Hrsg.), *Schulinspektion und Schulleitung* (S. 7–26). Stuttgart: Raabe.
- Wurster, S. (2016). *Evaluationsgestützte Schul- und Unterrichtsentwicklung. Eine multiperspektivische Analyse von Vergleichsarbeiten, zentralen Abschlussprüfungen, Schulinspektion und interner Evaluation* (Dissertation). Berlin: Humboldt-Universität zu Berlin.