

Marcus Pietsch, Barbara Scholand,
Klaudia Schulte (Hrsg.)

Schulinspektion in Hamburg

Der erste Zyklus 2007–2013:
Grundlagen, Befunde,
Perspektiven



HANSE
Hamburger Schriften zur Qualität
im Bildungswesen

herausgegeben von
Detlef Fickermann und Knut Schwippert

Band 15

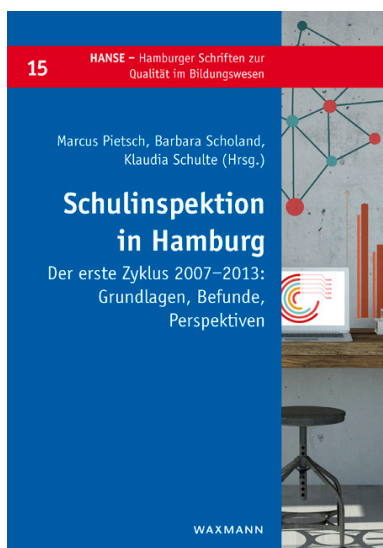


Waxmann 2015
Münster • New York

Marcus Pietsch, Barbara Scholand,
Klaudia Schulte (Hrsg.)

Schulinspektion in Hamburg

Der erste Zyklus 2007–2013:
Grundlagen, Befunde und Perspektiven



Das gedruckte Buch finden Sie [hier](#).



Waxmann 2015
Münster • New York

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten
sind im Internet über <http://dnb.d-nb.de> abrufbar.

HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Band 15

ISSN 1864-2225

ISBN 978-3-8309-3278-9

© Waxmann Verlag GmbH, 2015
Steinfurter Straße 555, 48159 Münster

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Pleßmann Design, Ascheberg
Umschlagfoto: © Robert Kneschke – Fotolia.de
Lektorat und Satz: Judith Zimmer, Hamburg
Druck: Mediaprint, Paderborn

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.
Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages
in irgendeiner Form reproduziert oder unter Verwendung elektronischer
Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Inhalt

Vorwort.	9
---------------	---

Teil 1: Die Grundlagen der Schulinspektion

Norbert Maritzen

Schulinspektion – Aspekte der Transformation von Governance-Strukturen . . .	13
--	----

Norbert Maritzen & Herbert Altrichter

Schulinspektion in Hamburg – ein Gespräch über Startbedingungen und Entwicklungen	37
--	----

Martina Diedrich

Aufbau und Rolle der Schulinspektion Hamburg	57
--	----

Thomas Brüsemeister

Erfahrungsbereiche der Hamburger Schulinspektion – Governance-analytische Programmtheorie	77
--	----

Susanne Ulrich

Das Qualitätsleitbild der Schulinspektion Hamburg	101
---	-----

Teil 2: Die Arbeit der Schulinspektion

Marcus Pietsch, Ann-Katrin van den Ham & Olaf Köller

Wirkungen von Schulinspektion: Ein Rahmen zur theoriegeleiteten Analyse von Schulinspektionseffekten	117
---	-----

Moritz G. Sowada

Expertenurteile – Achillesferse oder Trumpf der Schulinspektion?	137
--	-----

Knut Schwippert

Daten für die Schulentwicklung – auf die Perspektive kommt es an	157
--	-----

<i>Barbara Scholand unter Mitarbeit von Stephanie Graw</i>	
Wie Schulinspektor/innen fragen und Grundschüler/innen antworten	177

<i>Peter Schulze & Christiane von Schachtmeyer</i>	
Der Blick aus der Praxis – die Schulformexpertinnen und -experten	219

Teil 3: Die Befunde der Schulinspektion

<i>Melanie Ehren & Jaap Scheerens</i>	
Evidenzbasierte Referenzrahmen zur Schulqualität als Grundlage von Schulinspektion	233

<i>Holger Gärtner</i>	
Zusammenhang von Schul- und Unterrichtsqualität und schulischen Rahmenbedingungen	273

<i>Marcus Pietsch & Timo Ehmke</i>	
Die Qualität des Unterrichts an SINUS-Grundschulen. Eine Programmevaluation mithilfe von Daten der Schulinspektion Hamburg	295

<i>Klaudia Schulte & Markus Lücken</i>	
Der Einfluss schulischer Prozesse auf die Lernentwicklung der Schülerinnen und Schüler an weiterführenden Schulen in Hamburg	317

<i>Susanne Peters</i>	
Eltern als Stakeholder von Schule. Erkenntnisse über die Sichtweise von Eltern durch die Hamburger Schulinspektion.	341

Teil 4: Die Zukunft der Schulinspektion

<i>Sabine Müller</i>	
Zukunft der externen Evaluation – Trends und Perspektiven in Deutschland. .	369

<i>Hans-Jürgen Kuhn</i>	
Rechtliche, strukturelle und politische Rahmenbedingungen für externe Evaluation in Deutschland.	385

Martina Diedrich

Der zweite Zyklus der Schulinspektion Hamburg: ein Ausblick. 419

*Martin Spiewak, Angela Hellberg, Claudia Lemke, Gerrit Petrich,
Norbert Rosenboom, Mina Zulal*

Die Perspektive der Stakeholder: Ein Interview von Martin Spiewak 437

Autorinnen und Autoren 455

Vorwort

Vor rund zehn Jahren wurden die ersten Schulinspektorate in Deutschland eingeführt. Als Teil einer veränderten Steuerungsstrategie, die die Qualitätssicherung und -entwicklung von Schulen ins Zentrum setzt, sollen sie die Qualität schulischer Prozesse evaluieren, prozessuale Mindeststandards in Schule und Unterricht sichern und darüber hinaus Schulentwicklung stimulieren.

Die Hamburger Schulinspektion ist keine Inspektion der allerersten Stunde, wurde sie doch erst im Jahr 2007 offiziell eingerichtet. Zu diesem Zeitpunkt waren in vielen anderen Bundesländern schulische Evaluatorinnen und Evaluatoren bereits zwei bis drei Jahre im Feld und hatten die unterschiedlichsten Verfahren entwickelt und verschiedene Konzepte erprobt. Von ihren vielfältigen Erfahrungen mit diesem in Deutschland dahin unbekannten Instrument konnte die Hamburger Schulinspektion profitieren.

Gleichwohl hatte man in Hamburg bereits seit der Jahrtausendwende erste Konzepte zur externen Prozessevaluation an Einzelschulen in Hamburg diskutiert und entwickelt. Den Impuls hierfür gaben insbesondere die Ergebnisse der Längsschnittstudie LAU (Aspekte der Lernausgangslage und der Lernentwicklung), die seit Mitte der 1990er Jahre regelmäßig durchgeführt wurde. Es vergingen jedoch noch mehrere Jahre, bis die erste Hamburger Schule tatsächlich durch ein Inspektionssteam evaluiert wurde. Seitdem bewegt sich die Schulinspektion im Hamburger Schulsystem stets im Spannungsfeld von Administration, Praxis und Wissenschaft und muss dabei ihren eigenen sowie den Ansprüchen der verschiedenen schulischen und außerschulischen Stakeholder gerecht werden.

In ihrem ersten Zyklus, der im Jahr 2007 begann und bis zum Jahr 2013 andauerte, inspizierte die Schulinspektion 375 Hamburger Schulen. Während dieser sechs Jahre nahmen etwa 10.000 Lehrerinnen und Lehrer, gut 36.000 Eltern, rund 92.000 Schülerinnen und Schüler sowie etwa 2.000 Vertreterinnen und Vertreter Hamburger Unternehmen an den schriftlichen Befragungen der Inspektion teil. Darüber hinaus wurden fast 30.000 Unterrichtssequenzen durch die Inspektorinnen und Inspektoren beobachtet und etwa 2.000 Interviews mit Schulleitungen, Eltern- und Schülerräten sowie Lehrerkollegien und weiteren Mitarbeiterinnen und Mitarbeitern der inspizierten Schulen, aber auch mit Vertreterinnen und Vertretern von Ausbildungsbetrieben geführt.

Die Stärke des vorliegenden Bandes ist es, so meinen wir, dass er der besonderen Stellung der Schulinspektion im Gesamtgefüge des Hamburger Bildungssystems Rechnung trägt, indem er Akteure aus den verschiedenen Bereichen und mit unterschiedlichen Perspektiven zu Wort kommen lässt. Ein Teil der Beiträge ist empirisch ausgerichtet und analysiert die Arbeit der Inspektion oder nutzt die im Rahmen der Schulinspektion Hamburg erhobenen Daten für weiterführende

Analysen zum Hamburger Schulsystem. Andere Beiträge wiederum setzen sich mit den Rahmenbedingungen, Grundlagen und Folgen von Schulinspektion auseinander oder widmen sich den Erfahrungen und Erwartungen einzelner Akteure und Akteursgruppen.

Wir bedanken uns herzlich bei allen Autorinnen und Autoren dieses Bandes für die kooperative, kollegiale und anregende Zusammenarbeit. Insbesondere haben wir uns darüber gefreut, dass die vielen Kolleginnen und Kollegen aus diversen Kontexten und Arbeitsfeldern umstandslos bereit waren, ihre Expertise in Form fundierter und kritischer Beiträge sowie Gutachten beizusteuern. Sehr glücklich sind wir vor allem auch darüber, dass sich die Vertreterinnen und Vertreter der schulischen Stakeholder aktiv in diese Publikation eingebracht haben.

In der Zusammenarbeit mit all diesen verschiedenen Personen ist uns deutlich geworden, dass es institutionen- und akteursübergreifend ein großes Interesse daran gibt, über gute Bildung in und für Hamburg ins Gespräch zu kommen und sie gemeinsam zu gestalten. Dafür bieten – nicht nur, aber auch – die Schulinspektion und ihre Ergebnisse konstruktive Gesprächsanlässe. In diesem Sinne wünschen wir Ihnen eine ertragreiche Lektüre!

Hamburg im Juni 2015

Marcus Pietsch, Barbara Scholand & Klaudia Schulte

Teil 1

Die Grundlagen der Schulinspektion

Schulinspektion – Aspekte der Transformation von Governance-Strukturen¹

Norbert Maritzen

Zusammenfassung

Der Beitrag begreift die Einführung von Schulinspektionen in den deutschsprachigen Ländern im Allgemeinen und in Hamburg im Besonderen zum einen als Element des Umbaus von Steuerungsstrukturen im Schulwesen. Bildungspolitik und Bildungsverwaltung versuchen mit datengestützten externen Schulevaluationen eine Steuerungsfähigkeit zurückzugewinnen, die die traditionelle Schulaufsicht allein nicht mehr verbürgen kann. Zum anderen ist mit den Verfahren der Schulinspektion ein Anspruch an empirische Fundierung verknüpft. Dieser wird erklärt als Antwort auf Erfordernisse der Evidenzbasierung von politisch-administrativer Entscheidungsfindung, wobei zugleich aktuelle Inkonsistenzen gefundener Systemlösungen, Ungleichzeitigkeiten und noch ungelöste Entwicklungserfordernisse deutlich werden.

Die Einrichtung von Schulinspektoraten hat seit gut zehn Jahren einen besonderen Platz auf der Reformagenda der Schulministerien in den deutschsprachigen Ländern. Dokumentationen laufender Vorhaben oder institutionalisierter Inspektorate lassen sich in fast allen Internetauftritten der Landesministerien oder der ihnen nachgeordneten Landesinstitute bzw. Qualitätsagenturen² finden. Die Entwicklungen sind im Fluss, hier und da infolge politischer Wechselfälle auch rückläufig. Bestandsaufnahmen über Entwicklungsstände, die Überblickswissen und Orientierung vermitteln, bedürften eigentlich kontinuierlich der Aktualisierung (siehe exemplarisch Maritzen 2006; Dederling & Müller 2008; Döbert & Dederling 2008; Müller et al. 2011; von Saldern 2011; Böttcher et al. 2013 sowie den Beitrag von Kuhn in diesem Band). Während die kontroverse Diskussion über die Einrichtung von Schulinspektionen einige Jahre bestimmt wurde durch eher programmatische Verlautbarungen der politisch verantwortlichen Administrationen oder durch Projektberichte der für die Implementierung zuständigen Akteure, liegt inzwischen auch ein erster, durchaus kritisch zu diskutierender Korpus von Befunden empirischer Begleitforschung oder empirischer Wirksamkeitsforschung vor. Vor allem im Rahmen der Förder-

1 Überarbeiteter und ergänzter Nachdruck von Maritzen 2008.

2 Eine Übersicht über die Qualitätsagenturen bietet Fickermann 2014.

linie des BMBF „Steuerung im Bildungssystem“ (SteBis)³ wurden entsprechende Untersuchungen vorgelegt. Auch international vergleichende empirische Ansätze, die für eine Beschreibung und Einschätzung der deutschen Spezifika aufschlussreich sind, werden seit einiger Zeit vorangetrieben (vgl. Ehren et al. 2013). Wichtige Impulse erhalten diese Forschungen von theoretischen Modellen zur *Educational Governance* (siehe grundlegend Altrichter et al. 2007; Heinrich 2007; Kussau & Brüsemeister 2007; Maag Merki et al. 2014; Abs et al. 2015), die es ermöglichen, die Einführung und Wirkungsbedingungen von Schulinspektionen als Ergebnis der Intervention in die Handlungskoordination zwischen unterschiedlichen Ebenen des Schulsystems zu beschreiben, als systemische Veränderung des Beziehungsverhältnisses von Akteuren (z. B. Lehrkräfte, Eltern, Schulleitungen, Schulaufsicht, Politik, Öffentlichkeit), die auf und zwischen diesen Ebenen (z. B. Klasse, Schule, Region, Land) agieren. Während zu Beginn der Einführung von Schulinspektionen noch festgestellt werden konnte, dass mit erheblichem finanziellen Aufwand Steuerungsstrukturen des Schulwesens umgebaut wurden, ohne zugleich empirisch gesättigte Wirksamkeitsannahmen zu haben, lässt sich nun sagen: Das Wissen über Wirkungen, Nebenwirkungen und Folgen wächst beträchtlich, aber auch das verfängt kaum in der politisch induzierten Reform oder auch der Re-Reform der Steuerungspraxis, die weiter geprägt bleibt durch die Persistenz ungleichzeitiger paradigmatischer Regime von Aufsicht vs. Evaluation, von normativen Ansprüchen vs. empirischen Möglichkeiten (siehe kritisch bereits Lange 2003; Füssel 2008). Vermutlich sickert der empirische Erkenntnisgewinn am ehesten noch in die Verfahrensentwicklung der Inspektorate selbst ein, die zum Teil in engem Diskurs mit der Forschung stehen.

Die scheinbar breite Entwicklungskonvergenz in den deutschsprachigen Ländern täuscht gleichwohl über deutliche Unterschiede der jeweiligen Konzepte hinweg. Im Folgenden soll deshalb der Versuch unternommen werden, einige übergreifende Grundmuster der Entwicklung zu skizzieren, um darin die Hamburger Entwicklung zu verorten. Dabei geht es weniger um eine vergleichende Darstellung der Entwicklungen als um die Herausarbeitung einiger durchgängiger Fragestellungen oder Problemlagen. Insbesondere interessieren die mehr oder weniger ausgewiesenen Bezüge der Einführung von Inspektionen zu strategischen Vorhaben des Systemumbaus im Bereich der Steuerung.

Von besonderem Interesse für eine Einschätzung der aktuellen Situation sind in diesem Beitrag

- die jeweils ausgewiesenen Ziele und politischen Rahmenbedingungen der Implementierung von Inspektionsverfahren,

3 Siehe die Literaturdatenbank des Förderprogramms: http://www.literatur-rahmenprogramm-ebf.de/ebf_lit.html?t=Suchen&fl=Forschungsschwerpunkt&i1=Steuerung_im_Bildungssystem&pl=2000&

- die Veränderung der Konfigurationen von Steuerung: von Schulaufsicht zu Evaluation,
- die Frage der Institutionalisierung von Inspektionsverfahren inkl. rechtlicher Aspekte,
- die Ansprüche an Inspektionsverfahren im Kontext eines umfassenden Bildungsmonitorings.

1. Ziele und politische Rahmenbedingungen der Implementierung von Inspektionsverfahren

Sichtet man die offiziellen Verlautbarungen der Ministerien oder Schulverwaltungen, fällt eine gewisse Unterbestimmtheit auf. Die mit der Einführung dieses Instruments verbundenen Zielsetzungen bedurften entweder keiner besonderen Nennung oder sie changierten zwischen allgemeinen Hinweisen auf die gestiegene Ergebnisverantwortung und Rechenschaftspflicht der Einzelschule einerseits und Verweisen auf abstrakte Funktionen andererseits. Solche Verweise offenbarten eine Vielfalt von Funktionszuweisungen.⁴ Inspektionen haben

- eine Spiegel- oder Feedbackfunktion aus externer Sicht,
- eine Qualitätssicherungsfunktion im Rahmen staatlicher Gewährleistungsverantwortung,
- eine Unterstützungsfunktion insbesondere für Schulleitungen und Lehrkräfte in einzelnen Schulen,
- eine Impuls- oder Katalysatorfunktion für die Schul- und Unterrichtsentwicklung sowie
- eine Erkenntnisfunktion hinsichtlich der Wirkungen schulischer Arbeit.

Dieser in den Selbstbeschreibungen deutlich werdende Funktionenmix kann als Hinweis darauf gewertet werden, dass in der Frühphase hinsichtlich der mit den Schulinspektionen verbundenen Zielsetzungen erhebliche Unsicherheiten bestanden. Es schien schwerzufallen, Ziele, Funktionen und Leistungen von Schulinspektionen differenziert zu beschreiben, begrifflich gegeneinander abzusetzen und im Konzert weiterer Maßnahmen zu spezifizieren.

Diese „Unschärfe“ hängt vermutlich auch zusammen mit einer gewissen „Unterbestimmtheit“, die in den damaligen politisch-administrativen Kontextbedingungen festzustellen ist. Diese waren einerseits durch bestimmte wiederkehrende Grundmuster gekennzeichnet. Einen wesentlichen Bezugspunkt von Inspek-

4 Systematisch aufgearbeitet hat die spannungsreiche Funktionszuschreibung Landwehr 2011.

tionsverfahren bilden – in variierender Begrifflichkeit – Konzepte der erweiterten Verantwortung der Einzelschule. Vor diesem Hintergrund organisierten Länder, die Schulinspektionen einführten, Systematik und Verfahren von Aufsicht und Qualitätssicherung so um, dass der Schulinspektion eine besondere, aber selten trennscharf explizierte Aufgabe im Kontext der bereits etablierten Qualitätssicherungsmaßnahmen zukommen sollte. Schulinspektion bedeutete den Blick in oder auf Einzelschulen auf der Grundlage einer Zusammenschau vorhandener bzw. intern und/oder extern gewonnener Daten. In der Regel diente ein verbindliches Qualitätskonzept, d. h. ein System von Qualitätsbereichen und Qualitätsindikatoren, als Referenz- oder Orientierungsrahmen für die interne Evaluation und die Schulinspektion. Inspektionen sollten auf der Grundlage von Verfahren mit wiederkehrenden Standardelementen erfolgen, die meist in öffentlich zugänglichen Handbüchern niedergelegt worden sind. An dieser im ersten Zugriff nur allgemein beschriebenen Situation hat sich seither nichts Substanzielles geändert, auch in Hamburg nicht.

Trotz scheinbar konvergenter Entwicklungen sind gleichwohl deutliche Unterschiede bei der Implementation von Schulinspektionen in den verschiedenen Ländern vorzufinden, ohne dass immer sichtbar würde, welche strategischen oder inhaltlich-programmatischen Prämissen in die Entscheidung für bzw. gegen bestimmte Realisierungsoptionen eingegangen sind. Diese Unterschiede betreffen

- den Stand der Implementation systematischer Schulentwicklungsverfahren (Schulprogramm und interne Evaluation obligatorisch/nicht obligatorisch) und zentraler Verfahren der Lernstandserhebung und Ergebnismeldung;
- den Stand und die Nutzung des routinisierten Monitorings von Schulqualitätsdaten (rudimentäre bis systematische Verfügbarkeit von Schulqualitätsdaten auf allen Systemebenen);
- damit verbunden Art, Herkunft, Dichte, Qualität und Tiefe des „Datenkranzes“, der den Inspektionen für die zu inspizierende Einzelschule vorab oder ex-post zur Verfügung steht;
- die Rekrutierung und Zusammensetzung der Inspektionsteams (mit/ohne Beteiligung von Schulaufsicht, Peers, Eltern, Vertretern von Betrieben usw.);
- die Schnittstelle zur jeweils für die inspizierte Schule zuständigen Schulaufsicht (ins Inspektionsverfahren selbst involviert/nicht involviert);
- den Grad der Institutionalisierung der Inspektion (von temporären Netzwerken, in denen Kommissionen mit wechselnden Zusammensetzungen gebildet werden, bis zu ausgebauten Geschäftsstellen bzw. Inspektoraten) sowie
- schließlich den Grad der Verbindlichkeit der Verfahren, der von einem für die Einzelschule fakultativen Angebotscharakter bis hin zu obligatorischen „Zwangsveranstaltungen“ in behördlicher Verantwortung variieren kann.

Für die in diesem Kontext relevanten strategischen Entscheidungen Hamburgs kann man zusammenfassend feststellen:

- Die Schulinspektion ist von Anfang an als Instrument staatlicher Gewährleistungsverantwortung rechtlich normiert worden.
- Sie ist institutionalisiert worden mit spezifisch für die Aufgabe professionalisiertem Personal.
- Sie ist institutionell eingebunden worden in ein sukzessiv ausgebautes Bildungsmonitoring.
- Sie ist in ihrer Funktion deutlich getrennt worden von der Schulaufsicht und
- sie ist schließlich gekennzeichnet durch ein hohes Maß an Obligatorik für die betroffenen Akteure (Schulen, Schulleitungen, Lehrkräfte, Eltern).

Diese Grundentscheidungen prägen bis heute die Verfahren und Handlungsbedingungen in ihren Möglichkeiten und Grenzen.

2. Schulinspektion – von Schulaufsicht zu Evaluation

Man kann die Einführung von Schulinspektionen als Antwort auf eine Krise der Schulaufsicht interpretieren. In den Ländern der Bundesrepublik Deutschland und darüber hinaus ist seit den 1990er Jahren immer lautstärker Kritik an der Schulaufsicht geübt worden.⁵ Wenn auf diesen Umstand hingewiesen werden muss, ist weniger das Rechtsgut „Schulaufsicht“ gemeint, wie es z. B. im Grundgesetz der Bundesrepublik Deutschland und – darauf aufbauend – in den Verfassungen der Bundesländer fixiert ist. Wenn von einer Krise der Schulaufsicht die Rede ist, ist der institutionelle Organisationsrahmen gemeint, in dem das Rechtsgut „Aufsicht“ praktisch ausgeübt wird. Gemeint sind die unterschiedlichen Aufsichtsebenen der Schulverwaltungen und Ministerien. Die Krise legt, wenn sie ernsthaft analysiert wird, die Schlussfolgerung nahe, dass *erstens* schulaufsichtliches Handeln in entscheidenden Qualitätsfragen von „durchschlagender Wirkungslosigkeit“⁶ gekennzeichnet ist und *zweitens* die Dimension der inhaltlichen Qualitätssicherung von Unterricht und Schule nicht ausreichend in den Horizont der Schulaufsicht rückt.

Bildungspolitik und Schulverwaltungen reagierten auf diese Problemlage in den Ländern in unterschiedlicher Weise:

5 Zu Rolle und Entwicklungsproblematik der klassischen Schulaufsicht in den 1990er Jahren siehe insbesondere die Arbeiten von Rosenbusch 1994 und 1997; Burkard & Rolff 1994; Rosenbusch & Wissinger 1995 und Rosenbusch & Schlemmer 1997; Schratz 1996; Rolff 1998.

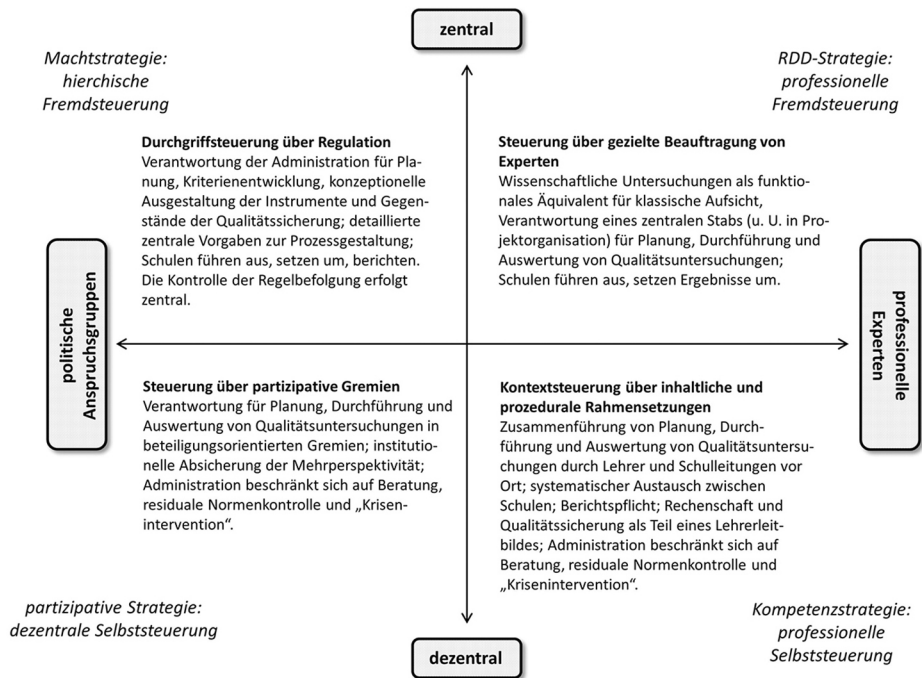
6 Max Frisch bescheinigte 1964 Bertolt Brecht „die durchschlagende Wirkungslosigkeit eines Klassikers“.

- Sie verschlankten, besser reduzierten die mehrstufigen Aufsichtsebenen, indem sie die untere oder mittlere Ebene der klassischerweise dreistufig organisierten Aufsicht einfach kappten. Dies ging einher mit Deregulierungsmaßnahmen, indem der nächst unteren Ebene (vor allem Schulleitungen) Aufgaben übertragen wurden, die bisher beanspruchtes Terrain der oberen Hierarchien ausmachten.
- Sie gingen daran, die im weitesten Sinne aufsichtlichen Instrumentarien zu modernisieren: Schülerleistungsuntersuchungen, Schulprogrammentwicklung, Evaluation, Ziel- und Leistungsvereinbarungen, Berichtswesen, Controlling, Assessment-Center, Beratung und manches mehr bereichern zunehmend das Interventions- und Steuerungsrepertoire.
- In den seltensten Fällen kam es zu einer institutionellen Trennung von Beratungs-, Aufsichts- und Evaluationsfunktionen, d. h. zum Versuch, in Fragen der Qualitätssicherung von zentralistischer Vorgabenregulierung und Eingriffsverwaltung hin zu einem Steuerungsmodell zu wechseln, das einen Regelkreis von kontraktuellen Rahmensetzungen in dezentralen, beteiligungsorientierten Verfahren einerseits und ziel- und ergebnisorientierter Evaluation andererseits miteinander verbindet.

Damit sind nur einige Stichworte einer Modernisierungsdebatte genannt, die insgesamt dadurch gekennzeichnet war, dass es an grundsätzlichen Vorstellungen und schlüssigen Begründungen dafür mangelte, wie Schulentwicklung in Zukunft überhaupt gesteuert werden sollte. Als Instrument zur kriterienorientierten Strukturierung dieser Positionsfindung sei deshalb eine Matrix (Abb. 1) vorgeschlagen, die dabei behilflich sein könnte, die existierenden Positionen zur Steuerung von Qualitätssicherungs- und Qualitätsentwicklungsfragen idealtypisch zu charakterisieren und zu strategischen Optionen zu kommen.

Im Bereich von „Aufsicht“ als einem Qualitätssicherungssystem lassen sich nämlich die in der Matrix skizzierten Hintergrundannahmen als steuerungsstrategische Orientierungen feststellen. Diese organisieren sich um zwei wesentliche Leitdifferenzen: einerseits die Fixierung auf zentrale Entscheidungshoheit von Politik und Administration bzw. umgekehrt die mehr oder weniger ausgeprägten Bemühungen um eine Dezentralisierung von Kompetenzen; andererseits die Konzeptionierung von Qualitätsfragen als Prozess politischer bzw. umgekehrt professioneller Definition. Man muss diese Polaritäten als Spannungsverhältnisse deuten, die nicht sich wechselseitig ausschließende Positionen, sondern jeweils unterschiedliche Ausprägungen auf einem Kontinuum andeuten. Über die Frage der Legitimation von Positionen und Konzepten ist damit noch nichts ausgesagt.

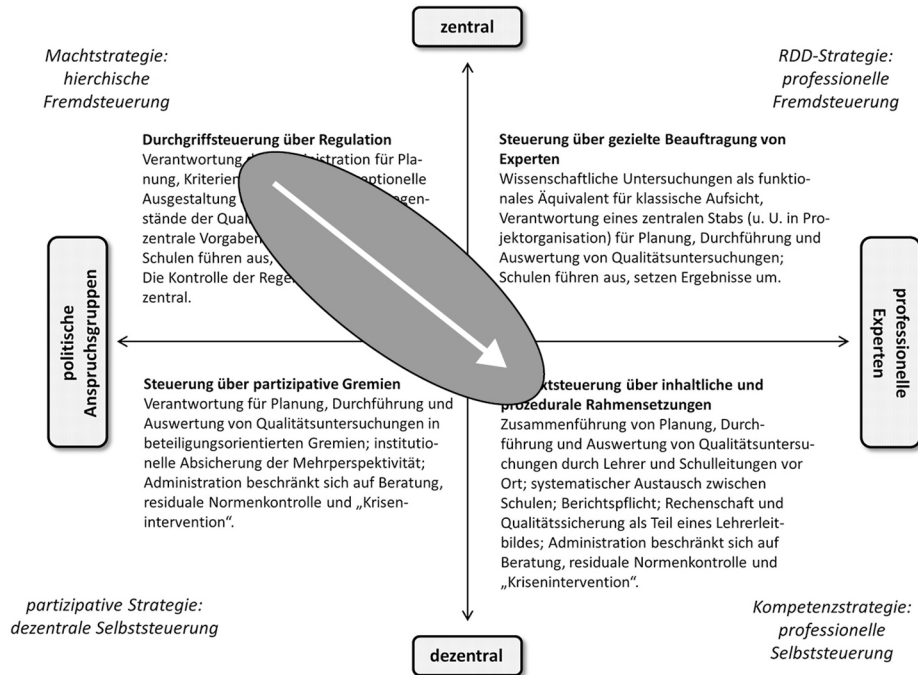
Abbildung 1: Steuerungsmatrix zur Qualitätssicherung



Quelle: Maritzen 1999

Kreuzt man diese Leitdimensionen erhält man vier idealtypische strategische Felder, die für alle Akteure im Bereich schulischer Qualitätsentwicklung und -sicherung sehr unterschiedliche Rollen und Aufgaben bereithalten. Angewandt auf Entscheidungen, die zu den Konzepten moderner Instrumentarien (z. B. Qualitätsmanagement und Evaluation) nötig werden, lässt sich unschwer erkennen, dass die gleichen Instrumente jeweils sehr unterschiedliche Ausprägungen erhalten, je nachdem ob – grob vereinfachend gesprochen – ihre Konzeptionierung eher zentralen Machtstrategien folgt oder der Logik dezentraler Selbstorganisation. Deutlich wird auch, dass alle Felder ihre spezifische Funktion besitzen. Es kommt wesentlich darauf an, auf der Grundlage eines explizit gemachten Konsenses in der Gesamtstrategie die Akteure, Handlungsebenen und Verfahren/Prozesse in einem stimmigen und legitimierbaren Gesamtzusammenhang zu vernetzen. Versucht man, die traditionelle Steuerung im Bereich der Qualitätssicherung in der Matrix zu skizzieren, ergibt sich grob folgendes Bild, wobei die Ellipse sehr grob die strategischen Prioritäten verdeutlicht:

Abbildung 2: Traditionelle Steuerung und ihre Entwicklung



Es lässt sich nun feststellen, dass Programme der Systemreform, die unter dem Schlagwort Schulautonomie rubrizieren, unterschiedlich weit, manche auch nur rhetorisch, in folgende Richtung gehen: in Richtung Dezentralisierung einerseits und Professionalisierung verknüpft mit Ausschöpfung partizipativer Möglichkeiten in den Schulen andererseits, d. h. – um im Bild zu bleiben – die Ellipse verschiebt sich unterschiedlich weit nach „Süd/Südost“. Diese Entwicklung tendierte dahin,

- in Fragen der Qualitätsentwicklung und -sicherung unter dem Stichwort Qualitätsmanagement auf eine Professionalisierung und Stärkung derjenigen zu setzen, die in der Schule die pädagogische Arbeit gestalten und verantworten müssen, d. h. Lehrerinnen und Lehrer und insbesondere Schulleitungen;
- die Professionellen in den Schulen gegenüber den direkten schulischen Anspruchsgruppen (Eltern, Schülerinnen und Schüler, Betriebe) verstärkt rechenschaftspflichtig zu machen und letztere systematischer und konsequenter auch in Qualitätsfragen zu beteiligen;
- zentrale Vorgaben nicht zu dicht zu gestalten zugunsten stärkerer und wirksamer Überprüfung der Einhaltung von Ergebnis- und Verfahrensstandards, der Überprüfung der Effizienz und Effektivität von Prozessen und Strukturen (z. B. durch externe Evaluation der Einzelschule/Schulinspektionen zurückzufahren);

- notwendige übergreifende Evaluationsmaßnahmen als Systemmonitoring so auszugestalten, dass die Verfahren komplementär und nicht konkurrierend zu dezentralen Evaluationen angelegt sind, d. h., z. B. die einzelnen Schulen zu Ergebnisevaluationen und zum Abgleich mit übergreifend gewonnenen Daten und Standards zu verpflichten;
- zentral gewonnene aggregierte Daten über Stärken und Schwächen des Gesamtsystems, über neu auftretende Probleme und Bedürfnisse, über die Wirksamkeit bestimmter Reformmaßnahmen, über schulrelevante gesellschaftliche Entwicklungen, über Entwicklungen allgemeiner Standarderreichungen systematischer in die zentralen Entscheidungsprozesse einzuspeisen und in den Schulbereich insgesamt rückzukoppeln.

Schulinspektion als spezifische Form externer Evaluation gewann also ihre aktuelle Bedeutung aufgrund von Verschiebungen im Gesamtsystem der Steuerung des Schulwesens. Damit ist noch nichts darüber gesagt, inwiefern diese Verschiebungen vielleicht Resultat der Auswirkungen sozialer Entwicklungsdynamiken auf die Steuerungsfähigkeit von Schulpolitik oder die Steuerbarkeit des Schulsystems sind.

Die Rahmenbedingungen für Veränderungen der Steuerungsstrukturen in einem Stadtstaat wie Hamburg hatten noch einmal besondere Spezifika, die sich von denen in Flächenländern tendenziell unterscheiden. Hier wird bis heute der Schulaufsicht auch das Geschäft des kommunalen Schulträgers aufgebürdet, was zur Folge hat, dass die Aufsicht noch stärker im operativen Tagesgeschäft gefangen bleibt. So kann es ihr kaum gelingen, den qualitativen Kern des Schulehaltens, d. h. Unterricht und Schulleben in den Blick zu nehmen. Soll der Blick doch darauf gerichtet werden, verschließen sich Schulen in der Regel. Wie anders ist zu erklären, dass der für die Stadtstaaten besonders heftige „PISA-Schock“ Anfang der 2000er Jahre die Schulaufsicht offensichtlich weitgehend unvorbereitet traf? Bei Einführung der Schulinspektion in Hamburg musste diesem Spezifikum in mehrfacher Hinsicht Rechnung getragen werden:

- Hamburg hat eine einstufige Schulaufsicht, die zugleich ministerielle und kommunale Zuständigkeiten wahrnimmt. Dies hat Folgen für die Strukturierung von Interventionsgelegenheiten und -maßnahmen der Schulaufsicht und umgekehrt für die Wahrnehmung der Schulaufsicht durch die Schulen. Die Zuschreibung von Wahrnehmungskonstrukten wie z. B. „Kontrolldichte“ oder „Steuerungszentralismus“ erfolgt in einstufigen Aufsichtssystemen notwendigerweise anders als in mehrfach gestuften Systemen.
- Anders als Schulsysteme mit bereits seit langem eingeführten Inspektionen (insbesondere die Niederlande oder England) verfügten Hamburger Schulen kaum über ein ausgebautes, professionalisiertes und nach innen weiter gestuftes

Schulmanagement. Dies hatte zur Folge, dass Aufsichtsfunktionen und operative kommunale Zuständigkeiten (Personal- und Ressourcenmanagement) erst in einem längerfristigen Prozess von der Ebene der Behörde auf die Schulebene delegiert werden konnten.

- Die institutionelle bzw. personelle „Dichte“ und Verflechtung der Akteure im Bereich der Systemgestaltung (Aufsicht, Schulleitungen, Landesinstitut, Politik) und der schulpolitischen Öffentlichkeit sowie die Tradition des personalisiert verteilten Systemwissens waren und sind ebenfalls durch die Spezifika eines Stadtstaates geprägt. Sie bestimmten (Vor-)Einstellungen zu Qualitätsaspekten des Schulbereichs, und zwar auf das Schulwesen insgesamt wie auf Einzelschulen bezogen. Der öffentliche wie der administrative Diskurs über Schulqualität erfolgt weniger grundsätzlich und systematisch als vielmehr konkretistisch und fallbezogen. Der Zugang der Öffentlichkeit zum Wissen über die Qualität der Einzelschule wurde nicht strategisch transparent gestaltet, indem ein solches Wissen seitens der Behörde breit zur Verfügung gestellt wurde. Dem entsprach, dass die einzelne Schule einen hohen Schutzanspruch hinsichtlich der Preisgabe von internem Wissen geltend machen konnte.
- Hamburg schaute allerdings auf eine langjährige Geschichte der Implementierung systematischer Schulentwicklung (Schulprogramm/interne Evaluation) mit hoher Obligatorik für alle Schulen und dem Ausbau einer unterstützenden und kontrollierenden Infrastruktur zurück.
- Hamburg verfügte ferner seit vielen Jahren über ein System zur Erhebung und Auswertung einer Vielzahl von Daten im Bereich der Strukturqualität (Lehrerdaten, Abschlüsse, Unterrichtsausfall, Klassenwiederholungen, Schüler-Lehrer-Relationen, Unterrichtsstundeneinsatz usw.). Die daraus entwickelten längsschnittlichen Zahlenreihen beleuchteten sehr unterschiedliche Bereiche des Schulsystems, ohne dass damit bereits der konsequente Schritt zu einem institutionalisierten Monitoring zu Aspekten der Schulqualität geleistet worden wäre.
- Gleichwohl waren die Voraussetzungen für ein Schulqualitätsmonitoring in Hamburg vergleichsweise günstig. Anders als andere Länder verfügte Hamburg bereits flächendeckend über „Daten“ zu Unterrichtsergebnissen aus unterschiedlichen Quellen: aus schulbezogenen/klassenbezogenen Ergebnissen unterschiedlicher Schulleistungsuntersuchungen (Lernausgangslagenuntersuchung: LAU, Kompetenzen und Einstellungen von Schülerinnen und Schülern: KESS u. a.), aus Rückkopplungs- und Beratungsverfahren für Schulleistungsergebnisse, aus Ergebnissen zentraler Prüfungen und Vergleichsarbeiten, aus den Ergebnissen verbindlicher Schulprogrammentwicklung, schulinterner Evaluationen und der Berichterstattung an die Schulaufsicht, aus Berichten zur Implementierung der Bildungspläne, aus formalisierten, thematisch fokussierten Hospitationsrunden der Schulaufsicht und Schulberatung.

Zu diesen noch weitgehend unverbundenen und unsystematisch verteilten Wissensbeständen mussten die Verfahren der Schulinspektion in ein explizites Verhältnis gesetzt werden. Das bleibt im Übrigen eine bis heute uneingelöste Entwicklungsaufgabe.

3. Institutionalierungsmodelle und ihre Konsequenzen

Schulinspektionsverfahren sind in vier idealtypisch zu verstehenden Modellvarianten implementiert worden:

1. *Netzwerkmodell*: fallbezogene Kooperation von systeminternen Evaluatoren, die im Hauptamt andere Funktionen im Schulsystem haben (z. B. als Schulleitungsmitglied, Ausbilder, Schulaufsicht); Staat oder Schulträger als Auftraggeber
2. *Expertenmodell*: fallbezogene Kooperation von systemexternen Evaluatoren, die auf Werkvertragsbasis direkt beauftragt werden oder indirekt über Auftragsvergabe an wissenschaftliche Institute, die als Auftragnehmer fungieren; Staat oder Schulträger als Auftraggeber
3. *Inspektoratsmodell*: fester Stamm von hauptamtlichen Inspektoren in einer eigenen staatlichen Einrichtung, die i. d. R. aus dem Organisationsverband der Schulbehörde ausgegliedert; Staat als Auftraggeber
4. *Zertifizierungsmodell*: Qualitätscheck und -bescheinigung durch eine „Agentur“ (wiss. Institut, Verein, Konsortium), durch zertifizierte Evaluatoren bzw. Auditoren; Schule als Auftraggeber

Die strategische Entscheidung für bestimmte Modellvarianten hatte immer Folgen für den Grad der Institutionalisierung der Inspektion, für Möglichkeiten und Grenzen der Professionalisierung, für Rollen und Verfügungsrechte der beteiligten Akteure, für die Steuerung des Inspektionsverfahrens oder die Verbindlichkeit von Folgen. Ob die Schulinspektion überhaupt eine Institutionalisierung im Sinne der Ansiedlung der Aufgabe in einer Einrichtung erfährt, wurde unterschiedlich beantwortet; ebenso die Frage, welchen Status die Organisationseinheit mit welchem Freiheits- bzw. Abhängigkeitsgrad hat.

Hinsichtlich der systemischen Anbindung der Schulinspektion sind folgende allgemeine Optionen denkbar:

Abbildung 3: Optionen der institutionellen Anbindung der Schulinspektion

	Anbindung innerhalb des Geschäftsbereichs des Schulministeriums	Anbindung außerhalb des Geschäftsbereichs des Schulministeriums
administrativ integriert	ministerieller Geschäftsbereich	Geschäftsbereich des Ministerpräsidenten
	mit folgenden Optionen: <ul style="list-style-type: none"> • in eine Behördenstruktur integriert vs. nicht integriert • nachgeordnete Dienststelle vs. selbstständiger Landesbetrieb oder andere Rechtsform (z. B. Anstalt öffentlichen Rechts) 	
Beauftragungsverfahren	wissenschaftliche Einrichtungen als Auftragnehmer	
politisch angebunden	z. B. in Stadtstaaten die Deputation der Schulbehörde	Parlament oder Parlamentsausschuss
dem Gesetz unterstellt	Inspektor/inn/en persönlich dem Minister unterstellt	Institution aus der Verwaltungsstruktur ausgegliedert (wie z. B. der Rechnungshof)

Die gegenwärtig in Deutschland implementierten Inspektionsmodelle, auch in Hamburg, zeichnen sich eher durch konservative Lösungen aus, d. h. sie belassen Inspektionen im Verantwortungsbereich des Schulministeriums. Unterschiedlich weit gehen sie in der Frage der Institutionalisierung: Entweder wird Schulinspektion als neues, gleichsam professionelleres Verfahren einer Schulaufsicht konzipiert, die zwar zu neuen kooperativen Vernetzungen mit anderen Akteuren führt, als solche aber in ihrer Zuständigkeit (Fach-, Dienst- und Rechtsaufsicht mit Kontroll- und Beratungsaspekten) nahezu unverändert bleibt (z. B. das vor einigen Jahren bereits abgeschaffte EVIT-Modell in Schleswig-Holstein). In anderen Fällen kommt es zu neuen institutionellen „Kernen“, die organisatorisch und personell getrennt von der Aufsicht agieren (z. B. in Niedersachsen, Brandenburg, Hessen oder Hamburg), obwohl sich ihr Personal zum Teil auch aus der Schulaufsicht rekrutiert oder die Anbindung an ministerielle Aufsichtsabteilungen erfolgt. Die Gründe für solche Lösungen sind vielschichtig. Je nach Kontext haben – nicht in Hamburg – zwei nachvollziehbare, aber sachfremde Strategien in der Gestaltungsfrage nicht selten den Ausschlag gegeben:

- zum einen der Versuch einer Modernisierung der unter verstärktem Legitimationszwang und operativem Druck stehenden Schulaufsicht mit den Mitteln der Inspektion,
- zum anderen die „Versorgung“ von freigesetztem Aufsichtspersonal mit neuen Aufgaben in Inspektoratseinrichtungen.

Teil der Institutionalisierung sind Entscheidungen zur rechtlichen Rahmung des Inspektionsverfahrens, da mit rechtlichen Normen zentrale Handlungs- und Wirkungsbedingungen festgelegt werden (vgl. den Beitrag von Kuhn in diesem Band). Die Länder verfahren in dieser Hinsicht sehr unterschiedlich unter Nutzung verschiedenster rechtlicher Ordnungsmittel (Gesetze, Erlasse, Verordnungen, Verfügungen, Dienstanweisungen). Zu den im Kontext der Schulinspektion zu klärenden Rechtsfragen sind mindestens folgende Aspekte als relevant zu erachten:

- Kodifizierung des Auftrags der Schulinspektion im Schul- bzw. Schulverwaltungsgesetz oder in amtlichen Erlassen und Rechtsverordnungen
- Absicherung des verbindlichen oder zwingend nötigen Umgangs mit Inspektionsberichten auf Gesetzes- oder Erlassebene bzw. in Dienstanweisungen
- Umgang mit schul- und personenbezogenen Daten, ggf. Änderungsbedarf bei einschlägigen datenschutzrechtlichen Bestimmungen (Datenverarbeitung zu Zwecken der Evaluation)
- Zusammenhang mit modifizierten rechtlichen Bestimmungen für die stärkere Eigenständigkeit der Schule
- Zusammenhang mit rechtlichen Bestimmungen für Schulen in freier Trägerschaft
- ggf. Errichtungsgesetze im Falle der Gründung eines eigenständigen Inspektors als Anstalt öffentlichen Rechts

Sehr grundsätzlich sind schließlich die Implikationen einer rechtlichen Absicherung der Qualitätskonzepte (Orientierungsrahmen), die in der Regel in den Ländern (so auch in Hamburg) konzeptioneller Referenzrahmen für Inspektionen sind. Qualitätskonzepte konkretisieren das Verständnis von Schul- und Unterrichtsqualität, also den Gegenstand von Inspektion, durch Angabe von Qualitätsbereichen, Qualitätsmerkmalen und Indikatoren. Hier sind mögliche Implikationen für das Verständnis von Aufsicht bisher kaum geprüft. Wenn in solchen Qualitätskonzepten bestimmte Merkmalbereiche und dort Ausprägungsgrade von Schulqualität normativ verbindlich gemacht werden, kann man von einem Schritt hin zu einem operationalisierten Aufsichtsverständnis sprechen. Dies hat möglicherweise restriktiv wirkende Konsequenzen

- für eine aufsichtliche Interventionspflicht bei Unterschreiten dieser operationalisierten Mindestnormen,
- für Ansprüche von Dritten (z. B. Eltern) auf Auskunft über das Erreichen dieser Normen,
- schließlich für die eben erst zaghaft erweiterten Handlungsspielräume von Schulen, die sich leicht über normativ ausgelegte Prozessstandards wieder verengen lassen.

Zu einer gesetzlichen Absicherung des Auftrags der Schulinspektion hat in Hamburg bei Einführung die Auffassung überwogen, dass die Schulinspektion mit ihrem Kernauftrag, wenn überhaupt, eher aus symbolischen Gründen im Schulgesetz genannt werden sollte, um der Einrichtung ein besonderes Gewicht zu verleihen. Dabei wurde Wert darauf gelegt, dass die Schulinspektion in der Ausübung ihrer Tätigkeit weisungsungebunden ist und anders als die Schulaufsicht gegenüber den Schulen kein Weisungsrecht hat. Rechtlich wird die Schulinspektion als Organ der staatlichen Gewährleistungsverantwortung („Aufsicht“ im weiteren Sinne) konzipiert, die gesetzlich hinreichend geregelt ist. Wichtig schien aus rechtlicher Sicht insbesondere, durch prozedurale Regelungen auch unterhalb des Schulgesetzes (z. B. Dienstanweisungen) sicherzustellen, dass der Umgang mit Berichten der Schulinspektion ein hohes Maß an Verbindlichkeit erhält. Dies gilt für einzelne Schulberichte ebenso wie für den vorgesehenen Jahresbericht.

4. Ansprüche an Inspektionsverfahren im Kontext eines umfassenden Bildungsmonitorings

4.1 Schulinspektion als Teil des Monitorings

Die steuerungsrelevanten Aufgaben im Bereich der Qualitätsentwicklung und Standardsicherung in Schulen haben sich in den letzten Jahren erheblich verändert, weitere Aufgaben sind hinzugekommen bzw. werden hinzukommen. Hinsichtlich ihrer funktionalen Unterscheidung sind sie zwar auf Paradigmen der Input- oder Output-Steuerung bezogen. Gleichwohl ist diese Unterscheidung nicht hinreichend, um die wachzunehmenden Aufgaben inhaltlich zu gruppieren und Schlussfolgerungen für deren institutionelle Verankerung zu ziehen. Die in diesem Feld virulenten Aufgaben werden seit einiger Zeit unter dem Begriff Bildungsmonitoring (Böttcher et al. 2008; Fickermann & Maritzen 2014; Döbert & Weishaupt 2015) zusammengefasst.

Zum einem System des Bildungsmonitorings gehören drei wesentliche Grundfunktionen, die sich im Übrigen auch in Inspektionsverfahren zeigen (siehe grundlegend Scheerens et al. 2003):

- *Akkreditierung/Zertifizierung*: Hiermit sind nicht im engeren Sinne Verfahrensmodelle der Qualitätsüberprüfung mit diesem Namen (z. B. nach ISO-Norm) gemeint, sondern im weiteren Sinne *alle* Verfahren der offiziellen Qualitätsbescheinigung anhand vorgegebener, formal festgelegter Normen, und zwar auf individueller Ebene (z. B. zentrale Prüfungen) oder auf Organisationsebene (z. B. externe Evaluation von Schulen u. a.).

- *Rechenschaftslegung*: Bildungsmonitoring hat wesentlich zum Ziel, den Zugang zu qualitätsrelevanten Informationen nach transparenten Regeln zu öffnen, und zwar auf der Ebene der Einzelinstitutionen ebenso wie auf Systemebene.
- *Diagnostik für systemisches Lernen*: Bildungsmonitoring hat diagnostische Informationen mit Gebrauchswert für unterschiedliche Systemebenen zu liefern, d. h., dass Produkte des Monitorings unter der Frage zu erstellen sind: Was leistet das Produkt zur Optimierung von „Zuständen“?

Bildungsmonitoring ist die systematische und auf Dauer angelegte Beschaffung und Aufbereitung von Informationen über ein Bildungssystem und dessen Umfeld. Absicht ist eine Verbesserung der Planung und Steuerung durch

- Beobachtung des Systems (Schulwesen, Einzelschule, Kontexte),
- Deskription, Vergleiche und Analysen,
- Untersuchung von Spezialfragen,
- Bestimmung von Handlungsbedarf,
- Feedback der Erkenntnisse.

Dies geschieht auf der Grundlage von „Daten“, die gewonnen werden aus verschiedenen amtlichen/nichtamtlichen Statistiken, einem auf Dauer gestellten und gepflegten System von Indikatoren, Maßnahmen der Qualitäts- und Standardsicherung und Forschungsdokumentationen. Bildungsmonitoring dient als Grundlage für evidenzbasierte Bildungsplanung und bildungspolitische Entscheide, auch für wissensbasierte Planungsentscheide einzelner Institutionen des Bildungssystems. Es steht grundsätzlich im Spannungsfeld zwischen der Bereitstellung von Steuerungswissen für Institutionen einerseits (Prinzip der Verantwortung) und der Gewährleistung öffentlicher Zugänglichkeit zu Informationen andererseits (Prinzip der Transparenz). Die beiden Prinzipien können durchaus miteinander in Konflikt geraten. Für einzelne Schulen wie für Ministerien gilt: Nicht alles, was gewusst werden kann, sollte zugleich auch öffentlich werden.

Bildungsmonitoring unter einer systemischen Perspektive einzuführen und zu gestalten bedeutet,

- dass die Verfahren nicht singulär, etwa für die Bewertung spezifischer Maßnahmen und Programme, eingesetzt werden, sondern dass sie als Bestandteil des Steuerungssystems institutionell eingeführt und laufend genutzt werden;
- dass Evaluation und Information Entwicklungs- und Lernprozesse in Gang setzen und am Leben halten, die dem Prinzip des feedbackgesteuerten Lernens genügen;

- dass verschiedene Arten der empirischen Erhebung (Testeinsätze, amtliche Daten u. a. m.), die je spezifischen Zwecken dienen, dennoch gemeinsam betrachtet und Synergien genutzt werden;
- und dass schließlich zur Systematisierung und Interpretation von Befunden ein theoretisches Rahmenmodell verwendet wird, welches Input-, Prozess- und Outputelemente zueinander in Beziehung setzt (vgl. Scheerens et al. 2003, S. 15).

Das Konzept des Bildungsmonitorings folgt in diesem Sinne dem Modell empiriegestützter Schul- und Systementwicklung. Entscheidend für das Modell ist, dass Ergebnis- und Prozessqualität gleichermaßen zum Gegenstand des Feedbacks gemacht werden. Insbesondere sollen möglichst alle wichtigen Prozessfaktoren abgedeckt werden, die gemäß den Ergebnissen der Schulwirkungsforschung für erfolgreiche Schulen kennzeichnend sind. Ein zentrales Argument für die Berücksichtigung von Ergebnis- *und* Prozessaspekten ergibt sich aus der Frage, wie *Qualitätsentwicklung* realisiert werden soll. Eine bloße Feststellung von Lernergebnissen mag für die Rechenschaftslegung nach außen oder (in einem noch stärker deregulierten System) für eine Akkreditierung von Institutionen genügen. Eine Schule als „lernende Organisation“ muss aber auch wissen, an welchen Prozessfaktoren sie ansetzen kann, um ihre Ergebnisse zu sichern bzw. zu verbessern (vgl. auch Klieme et al. 2005). Evaluation ohne Prozessanalyse bedient bestenfalls legitimatorische Zwecke und Kontrollwünsche, ermöglicht jedoch keine Weiterentwicklung. Hier kommt der Schulinspektion eine besondere Funktion zu.

4.2 Evidenzbasierung und wissenschaftliche Standards

Als Teil des oben skizzierten Bildungsmonitorings hat sich die Schulinspektion Anforderungen an Datenunterfütterung und an Verfahren zu stellen, die sich deutlich von administrativ geprägten Vorgaben unterscheiden, die etwa für Verfahren der Schulaufsicht gelten. Die Devise muss lauten: vom Augenschein zur Empirie, vom Vor-Gewussten zum Nachgewiesenen, vom Bescheidwissen zum Draufschaun, von kasuistischen Vorgehensweisen zu standardisierten Prozessen. Dies ist auf operativer Ebene zunächst auch eine Frage der Verfügbarkeit, Generierbarkeit und Nutzbarkeit von inspektionsrelevanten Daten und Informationen. Angedeutet seien hier keineswegs triviale Aspekte wie beispielsweise

- die Zusammenführung von Daten aus Beständen unterschiedlicher Datenhalter (diverse Akteure bzw. Zuständigkeiten in Schulen, bei Schulträgern, in Ministerien),

- die Qualität der Daten und die Performanz der Datenflüsse z. B. von der Schule zum Schulträger,
- die Zugriffsrechte und der Datenschutz,
- die Organisation eines arbeitsteiligen und dezentralen Daten- und Wissensmanagements,
- die technischen Infrastrukturen der Vernetzung von Daten und zur Organisation von Datenflüssen inkl. Such-, Auswertungs- und Berichtsfunktionalitäten,
- praktikable Instrumentarien für Datenerhebungen (webbasiert),
- eine verwendungsgerechte Aufbereitung von Daten für die Inspektion,
- die Verfügbarkeit von Referenzwerten für Schulqualitätsdaten.

Festzustellen ist, dass die Voraussetzungen hinsichtlich der genannten Aspekte des Datenmanagements in den Ländern, die Schulinspektionen eingeführt haben, höchst unterschiedlich sind. Nicht selten wird den Schulen vor Inspektionen eine äußerst kritisch zu beurteilende Bringschuld für Informationen und Daten aufgebürdet, gelegentlich auch für solche, die in den Schulbehörden eigentlich vorhanden sind.

Eine große qualitative Bandbreite zeichnet – bei aller festzustellenden Konvergenz der Verfahrensschritte⁷ – auch die empirische Fundierung des Inspektionsvorgehens aus. Wichtig zu wissen ist, dass mit den Verfahren alle Probleme, Klippen und Tücken vermacht sind, die man aus der empirischen Sozialforschung hinlänglich kennt. Man mag einwenden, dass Anforderungen an Zuverlässigkeit oder Objektivität der Verfahren nicht neu sind, also prinzipiell auch für Verfahren der Schulaufsicht gelten. Gleichwohl tritt die Schulinspektion mit dem Anspruch an, herkömmliches Aufsichtshandeln zur Qualitätssicherung professionell zu überbieten; insofern steht sie hinsichtlich des Geltungsanspruchs ihrer Verfahren und Urteile unter ungleich schärferem Legitimationsdruck.

Dies gilt insbesondere für ein Kernstück der Inspektionsverfahren, die Begutachtung von Unterrichtsqualität durch Unterrichtshospitationen. Hier ist die Kontrolle der Messungen hinsichtlich ihrer Objektivität, Reliabilität und Validität von besonderer Bedeutung, um Fehlerquellen bei der Bewertung von Unterrichtsqualität an Einzelschulen auf ein vertretbares Maß zu reduzieren. Eine besondere Herausforderung in diesem Zusammenhang stellt die Tatsache dar, dass sich die Einsichtnahmen in Unterricht in der Regel nur auf einen stichprobenartigen Zeitpunkt im Schuljahr konzentrieren und darüber hinaus die Stichprobengrößen, in denen Unterrichtssequenzen eingesehen werden können, prozess- und ressourcenbedingt eher klein sind, was das Risiko hoher Stichprobenfehler birgt. Stichprobenbedingte Messfehlerprobleme sind im Verfahren insofern kritisch

7 Zu den Inspektionsverfahren geben die Handbücher Auskunft, die üblicherweise von den Schulinspektionen bzw. den Qualitätsagenturen der Länder im Internet bereitgestellt werden.

abzuschätzen, als aufgrund einer Auswahl von Unterrichtssegmenten auf die Unterrichtsqualität einer Schule geschlossen werden soll.

Bisher ist in diesen Fragen das adäquate methodische Repertoire zum Umgang mit solchen Rahmenbedingungen erst in Ansätzen ausgearbeitet. Wie Beobachungskriterien operationalisiert sein müssen, welche Voraussetzungen durch Beobachtertraining oder Kalibrierung von Ratings durch Doppelbeobachtung geschaffen werden müssen, welche Anforderungen an eine Zufallsstichprobe von Unterrichtssegmenten aus der Menge aller an den Inspektionstagen erteilten Unterrichtsstunden zu stellen sind, wie die Zufallszuordnung der Beurteiler zu den Unterrichtssegmenten zu erfolgen hat, wie man Unterrichtseinschätzungen aggregiert und die Validität von Beobachtungen auch im Hinblick auf Beobachterstrenge oder -milde überprüft ... Dies sind Fragen, die dringend beantwortet werden müssen, und zwar nicht nur aus innerwissenschaftlichen Gründen, sondern zur Legitimation der Verfahren und damit zusammenhängender praktischer Entscheidungen. Die Sicherung der Verfahrensqualität ist zentrale Voraussetzung z. B. für

- Gewichtungen der Einzelbeurteilungen und Bestimmung von Normwerten für Beurteilungen,
- die Definition von Cut-offs zur Bestimmung von Sanktionsschwellen bei Unterschreitung von Normwerten,
- die Gewährleistung zwischenschulischer Vergleichbarkeit der Verfahren und Berichterstattung,
- die Nachvollziehbarkeit und Überprüfbarkeit der Urteile,
- Transparenz, Fairness, Revidierbarkeit von Urteilen,
- das Vertrauen der involvierten Akteure und die „Autorität“ der Inspektion.

Anders als in manchen sozialwissenschaftlichen Forschungsprojekten ist die Beantwortung solcher Fragen also keine erkenntnistheoretische Übung, sondern unmittelbar praxisrelevant. Schulinspektionen sind als komplex strukturierte soziale Veranstaltungen zu begreifen mit unter Umständen einschneidenden Folgen oder Nebenwirkungen für Individuen. Eine kritische Analyse des Inspektionsinstrumentariums muss deshalb von Beginn an Aufgabe der Schulinspektion selbst sein. Hamburg hat sich bei Gründung der Schulinspektion deshalb entschieden, in das hauptamtlich arbeitende Inspektorat Wissenschaftlerinnen und Wissenschaftler mit einschlägigen Kenntnissen aufzunehmen.

4.3 Vertrauen und Wissenwollen – eine spannungsreiche Beziehung

Die Einführung von Schulinspektionen geschieht – wie gezeigt auch in Hamburg – im Kontext eines weitreichenden Systemumbaus. Die Geschwindigkeit und Konsequenz, mit der die Reform in den letzten Jahren vorangetrieben worden ist, ist nicht zuletzt Ausdruck einer tiefen Vertrauenskrise. Nicht nur der Glaube an die Leistungsfähigkeit des deutschen Schulsystems ist im Gefolge von PISA erschüttert worden. Vom Ausmaß der Problemlagen nur höchst unzureichend gewusst zu haben, ist für Steuerungsinstanzen eine mindestens ebenso unverzeihliche Fehlleistung. Verlorenes Vertrauen in die Steuerfähigkeit zurückzugewinnen, darauf sind nicht wenige Maßnahmen des Systemumbaus – auch die Einführung von Schulinspektionen – angelegt. Dies soll geschehen durch Verfahren, die im Wesentlichen drei Anforderungen genügen: Nutzung wissenschaftlicher Verfahren der Erkenntnisgewinnung, Evidenzbasierung politischer und administrativer Entscheidungen, Transparenz über Rahmenbedingungen, Prozesse und Ergebnisse gegenüber Akteuren in Schule und Öffentlichkeit.

Vertrauen schaffen über Wissenwollen und Vorzeigen ist aber eine höchst ambivalente Angelegenheit, insofern Vertrauen zwischen Individuen und Institutionen eigentlich vom Verzicht darauf lebt, alles wissen zu wollen. Luhmann hat Vertrauen einen „Mechanismus der Reduktion sozialer Komplexität“ (Luhmann 2000a) genannt. Schulinspektionen stellen Einzelschulen unter Beobachtung; sie steigern potenziell die Komplexität der Wahrnehmung und des „Bildes“ vom Evaluationsgegenstand. Sie tun dies unter Rekurs auf einen ausdifferenzierten Beobachtungsapparat, dessen Solidität durch wissenschaftliche Expertise gesichert werden soll. Je größer die Investition in die Rationalität der Verfahren ist, umso deutlicher kann aber die Misstrauensbotschaft gleichsam als Kehrseite des Verfahrens zutage treten: Wir schauen genau, weil ihr es nicht wollt oder könnt. In der Perspektive der beobachteten Institutionen und Akteure ist ein Kontrollverdacht schnell bei der Hand, mindestens aber eine abwartende Reserve gegenüber Wissensangeboten, die mit dem Anspruch des Neuen, des von den Akteuren vor Ort so nicht Gewussten daher kommen. Hier in den Verfahren eine Beobachtungsökonomie zur Geltung zu bringen, die nicht als chronischer Misstrauensantrag perzipiert werden muss, bedarf eines schlüssigen strategischen Gesamtkonzeptes, das die Anwendung von Ansprüchen an Evidenz und Transparenz auf die Steuerungsinstanzen einschließen muss.

Letzteres ist mindestens ebenso ambivalent wie die Konstituierung von Beobachter-Objekt-Beziehungen im Verhältnis zu Schulen. Im Hinblick auf die Steuerungsdimension seien abschließend deshalb einige Dilemmata verdeutlicht:

- Mit Schulinspektionen kommt im Schulsektor ein Formenwandel des Regierens zum Ausdruck. Indem Schulpolitik und -verwaltung steuerungsrelevante

Wissensbestände systematischer in Entscheidungsprozesse integrieren und konsequenter auf das ohnehin nur lose gekoppelte Gesamtsystem (Verwaltung, Einzelschulen, Unterstützungssysteme, Nutzer) verteilen, geraten tradierte Formen der Generierung von Systemwissen (z. B. über Schulaufsicht) tendenziell in einen Modernitätsrückstand, so sie ihre traditionellen Funktionen nicht neu ausschärfen. Andererseits erhöht das wissensbasierte Steuerungssystem die Komplexität der Akteursbeziehungen, statt sie transparenter zu machen, da Wissen potenziell jedem zusteht, die Systemverantwortung aber ungleich verteilt bleibt. Verpflichtungskapazität und Problemlösepotenzial treten in Konkurrenz zueinander: Wissen, auf dessen Reflexion sich unterschiedliche Akteure leicht gemeinsam verpflichten lassen, löst nur selten Probleme; auf Entscheidungen angesichts immer differenzierter durchdrungener Problemlagen lassen sich eigenständiger werdende Akteure aber immer schwerer diskursiv, d. h. nicht-hierarchisch verpflichten.

- Das klassische Motiv des Einsatzes von Wissenschaften ist das der Absorption von Unsicherheit (vgl. ausführlich Luhmann 2000b, S. 183 ff.). Unsicherheit erwächst aus faktischem oder vermutetem Nichtwissen. Personen oder soziale Systeme wollen sicherer werden, wenn sie in ihrem Kontext wissenschaftlichen Wissenserwerb systematisch vorantreiben. In dem Maße, wie ihnen dies gelingt, schwindet aber auch die Illusion, Handlungs- und Entscheidungssicherheit wüchsen mit dem Wissen. Im Gegenteil, vermutlich wächst die Zahl konkreter Handlungsoptionen und damit die Unsicherheit, je genauer man wissenschaftlichen Aufschluss über Problembereiche erhält. Sicherheit/Unsicherheit und Wissen/Nichtwissen stehen gleichsam in einem rekursiven Verhältnis, sie regenerieren sich wechselseitig neu. Dies bekommt man in Steuerungsinstanzen deutlich zu spüren, wenn empirisch fundierte „Diagnosen“ zur radikalen Ausdifferenzierung von „Therapien“ zwingen, was einem die Wissenschaft keineswegs abnimmt. Wissen ist nicht ablösbar von dem Kontext, in dem es generiert wird. Weder in Politik, Verwaltung oder Schule lässt sich unter den jeweiligen komplexen Handlungsbedingungen wissenschaftliches Wissen einfach anwenden, um zu Entscheidungsoptionen zu kommen. Es erfordert vielmehr eine forschende Haltung zu den Fragestellungen des eigenen Handlungsbereichs, die Integration des Wissens in ein Konglomerat von bewussten und unbewussten Theorien und die Ankoppelung an häufig fallweise erworbene Lösungsalgorithmen. Nicht selten müssen dabei schmerzhaft Ablösungsprozesse von Konzepten in Kauf genommen werden, die sich bei manchen Akteuren über Jahre axiomatisch verfestigt haben. Die Bereitstellung von wissenschaftlichem Wissen erhöht zwar potenziell die Problemlösekapazität zur Systemsteuerung, als solches bleibt es aber inaktiv, wenn es nicht gelingt, es gezielt in relevanten konkurrierenden oder nichtwissenschaftlichen Wissensbeständen gleichsam im Hegel'schen Sinne aufzuheben. Davon weiß ein Lied zu singen, wer versucht,

Inspektionsergebnisse in Schule und Schulverwaltung rückzukoppeln, und sich über eine gewisse Reaktionsträgheit des Systems wundert.

- Wissen bzw. Nichtwissen sind auch unter Legitimationsgesichtspunkten durchaus ambivalent. Öffentlich Gewusstes (z.B. die konkrete, kleinräumig beschreibbare Verteilung unfairer Mechanismen sozialer Selektion im eigenen Schulsystem oder die Nichteinlösung von programmatischen Reformzielen) zwingt umso stärker zum Handeln, je differenzierter es vorliegt und verbreitet wird. Wenn also das instrumentelle Gestaltungsrepertoire und die vernetzte Problemlöseaktivität von politischen, administrativen und schulischen Akteuren nicht adäquat Schritt hält mit dem Wissen, dann besteht entweder die Versuchung, die entstehende Legitimationslücke durch intentionales Nichtwissen zu schließen („Es muss ja nicht jeder alles wissen“), oder es schleicht sich ein chronischer Legitimationsnotstand mit Konsequenzen für das Vertrauen in die Überlebensfähigkeit öffentlicher Verantwortungsbereiche ein oder einzelne Bereiche richten das Wissen gegen den, der im Gesamtkonzert die Verantwortung für Zukunftsentscheidungen trägt. In mehrfacher Hinsicht lässt sich also von einem Risiko des Wissenwollens sprechen. Es wundert nicht, dass angesichts solcher Legitimationsrisiken manche die empirische Wende bereits wieder am Ausgang ihrer Epoche sehen.

Mit der letzten Antinomie ist die Frage der Verantwortung angesprochen. Max Weber unterscheidet bekanntlich zwischen Gesinnungs- und Verantwortungsethik (vgl. Weber 1992). Erstere misst den „sittlichen Wert“ von Handlungsprogrammen – als solche könnte man auch bildungspolitische oder pädagogische Vorhaben bezeichnen – ausschließlich an Intentionen und Inhalten. Hält man an einem öffentlichen Auftrag für das Schulwesen, d. h. an einem Auftrag, der sich demokratisch legitimiert und eine staatliche verfasste Rahmung finden muss, fest, stellt sich für alle Akteure im Schulbereich unabweisbar die Frage, wie für Bedingungen und Folgen in unserem Bereich in Zukunft wirksam und nachhaltig Verantwortung zu übernehmen ist. Wissenschaftliche Erkenntnisse über Bedingungen und Wirkungen unseres Tuns machen, je differenzierter und tief gehender sie sind, das Leben keineswegs leichter, im Gegenteil. Eine Verantwortungsethik gebietet die kritische Nutzung der Wissenschaft, sie verbietet umgekehrt die bequeme Flucht durch die Tapetentür selbstverordneter Ignoranz.

Literatur

Abs, H.-J., Brüsemeister, Th., Schemmann, M. & Wissinger, J. (Hrsg.) (2015). *Governance im Bildungssystem – Analysen zur Mehrebenenperspektive, Steuerung und Koordination*. Wiesbaden: Springer VS.

- Altrichter, H., Brüsemeister, Th. & Wissinger, J. (Hrsg.) (2007). *Educational Governance. Handlungskoordination und Steuerung im Bildungssystem*. Wiesbaden: Springer VS.
- Böttcher, W., Bos, W., Döbert, H. & Holtappels, H. G. (Hrsg.) (2008). *Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive*. Münster: Waxmann.
- Böttcher, W., Hense, J. & Keune, M. (2013). Schulinspektion als eine Form externer Evaluation – ein Forschungsüberblick. In J. Hense, S. Rädiker, W. Böttcher & T. Widmer (Hrsg.), *Forschung über Evaluation. Bedingungen, Prozess und Wirkungen* (S. 231–250). Münster: Waxmann.
- Burkard, C. & Rolff, H.-G. (1994). Steuerleute auf neuem Kurs? Funktionen und Perspektiven der Schulaufsicht für die Schulentwicklung. In H.-G. Rolff, K.-O. Bauer, K. Klemm, H. Pfeiffer, & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung* (Bd. 8, S. 205–266). Weinheim, München: Juventa.
- Dederling, K. & Müller, S. (2008). Schulinspektion in Deutschland. In W. Böttcher, W. Bos, H. Döbert & H. G. Holtappels (Hrsg.), *Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive* (S. 241–252). Münster: Waxmann.
- Döbert, H. & Dederling, K. (Hrsg.) (2008). *Externe Evaluation von Schulen*. Münster: Waxmann.
- Döbert, H. & Weishaupt, H. (Hrsg.) (2015). *Bildungsmonitoring, Bildungsmanagement und Bildungssteuerung in Kommunen. Ein Handbuch*. Münster: Waxmann.
- Ehren, M. C. M., Altrichter, H., McNamara, G. & O'Hara, J. (2013). Impact of School Inspections on Improvement of Schools – Describing Assumptions on Casual Mechanisms in Six European Countries. *Educational Assessment, Evaluation and Accountability*, 25 (1), 3–43.
- Fickermann, D. (2014). Einrichtungen zur Qualitätssicherung und -entwicklung als „nachgeordnete Dienststelle besonderer Art“. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine datengestützte Schulentwicklung – Konzeption und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)* (S. 291–298). Münster: Waxmann.
- Fickermann, D. & Maritzen, N. (Hrsg.) (2014). *Grundlagen für eine daten- und theoriegestützte Schulentwicklung. Konzeption und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)*. Münster: Waxmann.
- Füssel, H.-P. (2008). Schulinspektion und Schulaufsicht, Schulinspektion oder Schulaufsicht, Schulinspektion versus Schulaufsicht, Schulinspektion als Schulaufsicht? In H. Döbert & K. Dederling (Hrsg.), *Externe Evaluation von Schulen* (S. 153–163). Münster: Waxmann.
- Heinrich, M. (2007). *Governance in der Schulentwicklung. Von der Autonomie zur evaluationsbasierten Steuerung*. Wiesbaden: Springer VS.
- Klieme, E., Steinert, B., Ciompa, R. & Gerecht, M. (2005). Auf dem Weg zu einem integrierten System der Qualitätssicherung für Schulen. In H. Avenarius, K. Klemm, E. Klieme & J. Roitsch (Hrsg.), *Bildung gestalten – erforschen – erleben* (S. 68–91). Neuwied: Luchterhand.

- Kussau, J. & Brüsemeister, Th. (2007). *Governance, Schule und Politik. Zwischen Antagonismus und Kooperation*. Wiesbaden: Springer VS.
- Landwehr, N. (2011). Thesen zur Wirkung und Wirksamkeit der externen Evaluation. In C. Quesel, V. Husfeldt, N. Landwehr & P. Steiner (Hrsg.), *Wirkungen und Wirksamkeit der externen Schulevaluation* (S. 39–43). Bern: hep-Verlag.
- Lange, H. (2003). Schulaufsicht zwischen normativen Anforderungen und faktischen Wirkungsmöglichkeiten. In H.-P. Füssel, P.M. Roeder, Peter M. (Hrsg.), *Recht – Erziehung – Staat. Zur Genese einer Problemkonstellation und zur Programmatik ihrer zukünftigen Entwicklung* (Zeitschrift für Pädagogik: Beiheft 47, S. 137–155). Weinheim: Beltz.
- Luhmann, N. (2000a). *Vertrauen. Ein Mechanismus der Reduktion sozialer Komplexität*. Stuttgart: UTB.
- Luhmann, N. (2000b). *Organisation und Entscheidung*. Opladen: Westdeutscher Verlag.
- Maag Merki, K., Langer, R. & Altrichter, H. (Hrsg.) (2014). *Educational Governance als Forschungsperspektive. Strategien. Methoden. Ansätze*. Wiesbaden: Springer VS.
- Maritzen, N. (1999). Schulaufsicht: Auf dem Weg von der Regulation zum Vertrag? Risiken und Chancen von Vertragsmodellen in der Schulentwicklung. In M. Beucke-Galm, G. Fatzer & R. Rutrecht (Hrsg.), *Trias-Kompass EHP-Organisation. Bd. 2: Schulentwicklung als Organisationsentwicklung* (S. 125–140). Köln: Edition Humanistische Psychologie.
- Maritzen, N. (2006). Schulinspektion zwischen Aufsicht und Draufsicht – Eine Trendanalyse. In H. Buchen, L. Horster & H.-G. Rolff (Hrsg.), *Schulinspektion und Schulleitung* (S. 7–26). Stuttgart: Raabe, Fachverl. für Bildungsmanagement.
- Maritzen, N. (2008). Schulinspektionen – Einige Aspekte der Transformation von Governance-Strukturen im Schulwesen. *Die Deutsche Schule*, 100 (1), 85–96.
- Müller, S., Pietsch M. & Bos, W. (Hrsg.) (2011). *Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht*. Münster: Waxmann.
- Rolff, H.-G. (1998). Schulaufsicht und Administration in Entwicklung. In H. Altrichter, W. Schley & M. Schratz (Hrsg.), *Handbuch der Schulentwicklung* (S. 190–217). Innsbruck: StudienVerlag.
- Rosenbusch, H. S. (1994). *Lehrer und Schulräte. Ein strukturell gestörtes Verhältnis. Berichte und organisationspädagogische Alternativen zur traditionellen Schulaufsicht*. Bad Heilbrunn: Klinkhardt.
- Rosenbusch, H.S. (1997). Organisationspädagogische Perspektiven einer Reform der Schulorganisation. Ein zielorientierter Denkansatz zur Bestimmung von Handlungsstrukturen und -vollzügen. *SchulVerwaltung – Bayern*, 10, 329–334.
- Rosenbusch, H. S. & Schlemmer, E. (1997). Die Rolle der Schulaufsicht bei der Pädagogischen Entwicklung von Einzelschulen. *Schulmanagement*, 28 (6), 9–17.
- Rosenbusch, H. S. & Wissinger, J. (Hrsg.) (1995). *Schule und Schulaufsicht – Wege zur Reform*. Braunschweig: SL Verlag.

- Scheerens, J., Glas, C. & Thomas, S. M. (2003). *Educational Evaluation, Assessment and Monitoring. A systematic Approach*. Lisse: Swets & Zeitlinger.
- Schratz, M. (1996). *Die Rolle der Schulaufsicht in der autonomen Schulentwicklung*. Innsbruck/Wien: StudienVerlag.
- von Saldern, M. (Hrsg.) (2011). *Schulinspektion: Fluch und Segen externer Evaluation*. Norderstedt: Books on Demand.
- Weber, M. (1992). *Politik als Beruf*. Stuttgart. Reclam.

Schulinspektion in Hamburg – ein Gespräch über Startbedingungen und Entwicklungen

Norbert Maritzen & Herbert Altrichter

Zusammenfassung

Im Jahr 2007 wurde die Schulinspektion in Hamburg offiziell eingerichtet. Bis es soweit war, mussten jedoch die Voraussetzungen hierfür geschaffen und verschiedene Entscheidungen auf administrativer Ebene getroffen werden. Darüber hinaus mussten, gerade zu Beginn der Inspektionsarbeit, aber auch im weiteren Entwicklungsverlauf, verschiedenste systemische Spannungsverhältnisse austariert und die Rollen einzelner Akteursgruppen im Hamburger Schulsystem und innerhalb der damals neu gegründeten Schulinspektion geklärt und definiert werden. Der Governance-Experte Herbert Altrichter, der u. a. im Rahmen eines EU-Projektes zu den Gelingensbedingungen erfolgreicher Schulinspektionen forscht, spricht hierüber mit dem Leiter des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung, Norbert Maritzen, der seitens der Hamburger Bildungsadministration bereits um die Jahrtausendwende mit der Einführung einer Schulinspektion in Hamburg betraut war und der deren Entwicklung bis heute begleitet.

Herbert Altrichter: Lieber Norbert, ich freue mich, mit dir über die Entwicklung der Hamburger Schulinspektion zu reden. Ich nehme an, 2006 war ein entscheidendes Jahr. Wenn du an diesen Schnittpunkt denkst: Was ist eigentlich zuvor passiert, was waren Entwicklungen, die ermöglicht haben, dass eine solche Schulinspektion entsteht?

Norbert Maritzen: Es hat in Hamburg schon weit vor 2006 Überlegungen in der Schulbehörde gegeben, sich mit der externen Evaluation von einzelnen Schulen – das wurde noch nicht unbedingt Schulinspektion genannt – konzeptionell zu beschäftigen. Da war schon 1997 die Schulgesetzgebung ausschlaggebend gewesen, die die Schulprogrammentwicklung und -evaluation gesetzlich als Auftrag für alle Schulen festgeschrieben hat. In den darauf aufsetzenden Planungen der Folgejahre hat es auch gelegentlich Versuche gegeben, Projektskizzen zu entwerfen für die Einrichtung von Schulinspektionen. Allerdings sind mehrere Anläufe gescheitert. Es gab auch öffentlichen Druck. Ausgezeichnet hat sich besonders die Hamburger Handelskammer, deren Geschäftsführer öffentlich verlautbart hat, die

Schulinspektion sei ein Desiderat in Hamburg, und eine international besetzte Tagung zu dem Thema organisiert hat.

Herbert Altrichter: In welcher Rolle warst du zu dem Zeitpunkt?

Norbert Maritzen: Ich war in meiner Funktion als Unterabteilungsleiter im Amt für Schule zugleich als Projektleiter mit der Einführung der Schulprogrammentwicklung und Evaluation in Hamburger Schulen beauftragt. Dazu leitete ich eine Projektgruppe, die eine ganze Weile der Motor solcher Überlegungen und Anstrengungen gewesen ist. In meiner Abteilung wurde dann auch 2002 eine Fachtagung zum Thema „Externe Evaluation der Einzelschule“ organisiert. Dort haben Kollegen aus Nordrhein-Westfalen, Niedersachsen und Schleswig-Holstein sowie wissenschaftliche Experten über die Einführung von externen Evaluationen berichtet. Auch die Handelskammer Hamburg war beteiligt. Das Vorhaben ist dann aber erst mal einige Jahre im Sande verlaufen. Erst mit dem Regierungswechsel von der SPD zu einem CDU-regierten Senat und im Schulbereich unter der Leitung von der Senatorin Dinges-Dierig sind die Überlegungen wieder aufgegriffen und systematisiert worden. Zu dem Zeitpunkt wurde eine neue Projektgruppe zur Einführung der Schulinspektion installiert.

Herbert Altrichter: Setzte sich diese Projektgruppe aus Personen aus dem Schulsystem zusammen oder gab es auch externe Mitglieder?

Norbert Maritzen: Die Gruppe setzte sich unter meiner Leitung zusammen aus einer Kollegin aus meiner Abteilung, zwei Vertreterinnen des Landesinstitutes für Lehrerbildung und Schulentwicklung, je einem Vertreter der Schulaufsicht der unterschiedlichen Schulformen, dem Vertreter der Grundsatzabteilung der Schulbehörde und je einer Schulleitung pro Schulform. Also gab es Personen aus der Administration, aber auch aus den Schulen und dem Unterstützungssystem.

Herbert Altrichter: Das klingt nach einer recht großen Gruppe.

Norbert Maritzen: Ja. Und dort ist die Arbeit relativ zügig aufgenommen worden. Ich bin damals davon ausgegangen, dass man diesen Prozess relativ zügig vorantreiben muss, wenn man eine Chance haben will, eine Institutionalisierung innerhalb einer Legislaturperiode abzusichern. Damals waren das noch vier Jahre. Ich habe eine Vorsteuerung der Gruppe versucht, indem ich mit konzeptionellen Vorüberlegungen in die Sitzungen reingegangen bin. Am Ende der Sitzung war immer ein anderes Produkt entstanden. Aber wesentlich war, dass wir so relativ schnell, nach zwei bis drei Monaten, Eckpunkte bzw. Fragen zu grundsätzlichen Entscheidungen an die Behördenleitungen geben konnten. Wir haben so bereits

zu Beginn klar beschrieben, dass die Einführung der Schulinspektion Implikationen für das Gesamtsystem, für die Rolle von Schulleitungen, für die Rolle von Schulaufsicht und für Unterstützungssysteme hat. Das heißt, für den Erfolg der Einführung war es wichtig, den Gesamtzusammenhang zu sehen. Darum mussten sehr grundsätzliche Entscheidungen gleich zu Beginn stehen.

Herbert Altrichter: Die Sorge um Qualität in einem Bereich ist eigentlich die Aufgabe der Linienvorgesetzten, also in diesem Fall der Schulaufsicht. Daher frage ich mich: Wie kam man auf die Idee, dass man eine spezielle Position im System braucht, die das tut, was die Linienvorgesetzten auch machen könnten? Was gewinnt man dadurch? Welche Probleme handelt man sich dadurch ein?

Norbert Maritzen: Das ist in der Tat richtig und hat die Diskurssituation zu Beginn des Projektes maßgeblich beeinflusst. Diese Diskussion wurde auch in der Projektgruppe geführt, weil dort auch die Vertreter der aufsichtführenden Schulformabteilungen vertreten waren. Sehr deutlich wurde, dass der qualitätssichernde Auftrag hinsichtlich bestimmter Aspekte der Qualität von Schule, insbesondere von Unterricht, angesichts des sonstigen Aufgabenportfolios der Schulaufsicht in einem Stadtstaat nicht erbracht werden kann. Denn Hamburg hat eine einstufige Schulaufsicht. Die Aufsichten übernehmen daher nicht nur die qualitätssichernde Funktion, sondern waren zum damaligen Zeitpunkt, als noch viele Aufgaben nicht an die Schulleitung delegiert waren, sehr beansprucht von dem, wie wir es nennen, Trägersgeschäft. Das heißt, den Aufgaben des Schulträgers hinsichtlich Gebäude, Personal, hinsichtlich der Notwendigkeit, Elternbeschwerden nachzugehen, und, und, und ... So hatten die Kolleginnen und Kollegen eigentlich gar keine Zeit, den Blick auf die Primärprozesse des Unterrichts zu richten. Einzelne Schulaufsichten haben das trotzdem getan und darüber z.B. Berichte geschrieben. Dies war aber sehr unsystematisch und folgte keinen Standards. Außerdem wurde in den ersten Wochen der Arbeit in der Projektgruppe sehr deutlich, dass es auch an der Infrastruktur fehlte, systematisch einzelschulbezogene Informationen in nachvollziehbarer und damit bewertbarer Weise aufzubereiten. Das hat für die Akzeptanz des Gedankens, zu diesem Zweck eine Einrichtung zu schaffen, gesorgt. Schwierig und sehr konflikthaft war dann aber auch, dass die Einrichtung der Inspektion unter dem Vorzeichen stand, dass die Finanzierung dieser Einrichtung aus dem Stellenbestand der Schulbehörde zu erfolgen hatte. Das heißt, frei werdende Stellen, und das waren häufig auch Stellen der Schulaufsicht, mussten vakant gehalten werden, um sie für die Gründung der Schulinspektion zur Verfügung zu haben. Damit ging natürlich bei einigen die Perspektive einher, dass die Schulinspektion zum Totengräber der Schulaufsicht werde. Durchgesetzt werden konnte diese Strategie im Prinzip nur durch den erklärten Willen der Behördenleitung, so vorzugehen. Deshalb haben wir uns in diesem Punkt sehr früh darum bemüht, entsprechende

Linienentscheidungen der Behördenleitung herbeizuführen. Aus einem Projekt, das neben der Linie organisiert ist, solche Entwicklungen anzustoßen, erschien mir nicht aussichtsreich.

Herbert Altrichter: Und die Schulaufsicht musste in die Schulinspektionen dann wiederum über Ziel- und Leistungsvereinbarungen im Anschluss an die Inspektionen eingebunden werden?

Norbert Maritzen: Richtig. Das war damals wichtig. In der Implementierung hat es allerdings lange gedauert, bis dieses Instrument als solches genutzt und erkannt worden ist und auch so weiter entwickelt wurde, dass es im Anschluss an Inspektionen praktikabel war. Das barg die Möglichkeit, sehr explizit den Punkt der Staffelübergabe an die Schulaufsicht fest zu institutionalisieren. Das hat auch den Konflikt an der Stelle entschärft, weil die Aufsicht dann eine Aufgabe erkannt hat, bei der sie wieder den Hut auf hat.

Herbert Altrichter: Es gab also in der Behörde eben eine erweiterte Projektgruppe, in der verschiedene Interessen vertreten waren. Du hattest außerdem gesagt, dass eine bestimmte politische Konstellation wichtig war. Welche weiteren – förderlichen und skeptischen – Akteure waren damals wichtig, um es bis 2006 zu der Institutionalisierung der Schulinspektion zu bringen?

Norbert Maritzen: Die Schulleitungen haben das zunächst distanziert betrachtet. Die Lehrerverbände, die Personalräte und auch die Gewerkschaften haben das sehr kritisch bis ablehnend begleitet. Ich kann mich an Einladungen erinnern, die ich dann natürlich als Projektleiter von Personalräten und auch von den Gewerkschaften erhielt. Diese Veranstaltungen waren nicht vergnügungssteuerverpflichtig. Da gab es eine große Skepsis, was den Zugriff auf Kenntnisse über das innere Funktionieren von Schule angeht, große Skepsis hinsichtlich der Verwendung der Daten, die Befürchtung, dass die Berichte veröffentlicht würden. Es gibt entsprechende Verlautbarungen der damaligen Senatorin, dass das nicht geschehen werde – inzwischen, einige Jahre später, im zweiten Durchgang der Schulinspektion, geschieht dies. Vorbehalte gab es auch bei nicht wenigen Schulaufsichtskollegen. In die Projektgruppenarbeit waren eher die Kolleginnen und Kollegen involviert, die selbst ein hohes Interesse an diesem Auftrag hatten. Sie hatten aber ihrerseits in der Rückkopplung in ihre Abteilung hinein doch erhebliche Schwierigkeiten unterschiedlicher Art. Das war auch zum Teil schulformabhängig. Dies war also eine Schwierigkeit in der Vermittlung des Auftrags. Eltern und Elternverbände haben sich relativ früh zu Wort gemeldet und darum gebeten, involviert zu werden. Sie haben das als wichtig erachtet. Von dort kam also eher Unterstützung. Allerdings wurden dort wiederum Erwartungen formuliert, in Inspektionen einge-

bunden zu sein oder Informationen aus Inspektionen zur Verfügung gestellt zu bekommen – Erwartungen, die nicht erfüllt werden konnten. Da hat es auch manche Enttäuschungen bei Eltern gegeben.

Herbert Altrichter: Also Erwartungen in Richtung Veröffentlichung der Ergebnisse einzelner Schulen?

Norbert Maritzen: Ja. Oder eine Funktionszuschreibung, dass die Schulinspektion Adressat sei für Beschwerden aller Art, die Eltern immer hinsichtlich einzelner Lehrkräfte, hinsichtlich der Schulleitung oder der Schule insgesamt haben können. Schulinspektion als diejenige, die längst bekannte, aber immer wieder unter den Teppich gekehrte Skandale aufdecken soll. Die Haltung der Elternseite war also eigentlich positiv, es gab aber häufig auch eine unrealistische Erwartungshaltung an die Schulinspektion. Das war nicht einfach, in diesem Punkt realistische Erwartungen zu erreichen. Im berufsbildenden Bereich haben wir relativ früh auch Betriebsvertreter involviert. Es war aber auch hier nicht einfach, Vertreter aus Betrieben oder aus Verbänden wie dem Unternehmensverband Nord oder der Handelskammer einzubeziehen. Dort gab es eine ganze Reihe von Aktivitäten wie z. B. Gespräche der Schulaufsichten aus dem Berufsschulbereich mit Wirtschaftsvertretern. Der Ansatz wurde von dieser Seite aber eher positiv betrachtet.

Herbert Altrichter: War das Thema Schulinspektion etwas, zu dem sich die politischen Parteien unterschiedlich verhalten haben?

Norbert Maritzen: Eigentlich war das Thema Schulinspektion jenseits der Parteien angesiedelt. Ich kann mich nicht erinnern, dass es einen richtigen Konflikt zwischen der damaligen Opposition, der SPD, und dem CDU-Senat gegeben hätte. Die eingeführten Inspektionen in der Bundesrepublik in anderen Ländern wurden auch von Parteien ganz unterschiedlicher Couleur in ihrem Parteiprogramm geführt. Ich kann mich an Sitzungen des Schulausschusses der Bürgerschaft erinnern, in dem man die entsprechenden Planungen vorlegen musste. Das lief immer relativ entspannt ab. Kritik gab es eher bei Umsetzungsdetails oder in der Frage, gleich zu Beginn schon für eine Veröffentlichung der Berichte zu sorgen oder erst mal die Inspektion ein paar Jahre Erfahrungen sammeln zu lassen, bevor man sich an so einen Schritt traut. Das war damals auch innerhalb der CDU nicht ganz unumstritten.

Herbert Altrichter: Was waren eigentlich kritische Punkte in dieser Entwicklung, Dinge, über die oft geredet wurde, über die unterschiedliche Meinungen vorhanden waren?

Norbert Maritzen: Kritische Punkte waren das lange ungeklärte Verhältnis der Institution der Inspektion auf der einen Seite und der Schulaufsicht auf der anderen. Da hat die Auftragsbeschreibung und die gesetzliche Grundlage der Schulinspektion nicht gereicht, um zum Beispiel aus Schulleitungs- oder aus Lehrkräftesicht klar zu haben: Das ist Aufgabe der Aufsicht und das ist Aufgabe der Inspektion. Lange waren die Übergabepunkte nicht klar, auch in der Perspektive.

Herbert Altrichter: Also: Was passiert mit dem Ergebnis und wann verlässt die Inspektion den Campus und überlässt das Feld der Vorgesetztenlinie?

Norbert Maritzen: Kritisch war innerhalb der Behörde vor allem auch die Nachwirkung der Stellenverlagerungen. Das ist organisationspolitisch von den Aufsichten als extreme Beschneidung betrachtet worden, die einherging mit einem Overload an Aufgaben. Da hat gerade die erste Generation der Inspektoren auch einiges abgekriegt, wofür sie nichts konnte.

Herbert Altrichter: Hat es in dieser Entwicklungsphase Modelle gegeben, an denen man sich orientiert oder von denen man sich abgegrenzt hat? Hat es Bausteine, Elemente gegeben, bei denen man gesagt hat, das brauchen wir unbedingt oder das wollen wir auf keinen Fall?

Norbert Maritzen: Wir haben uns bereits im Vorfeld der Gründung der Projektgruppe in meiner Abteilung intensiv damit beschäftigt, welche Inspektionsmodelle in einigen Bundesländern zu entstehen begannen und wie die Situation in Skandinavien, in den Niederlanden, in England aussah. Mit diesen Modellen haben wir uns intensiv auseinandergesetzt und schnell gesehen, dass man diese Modelle nicht kopieren kann, sondern dass es für die Bedingungen eines Stadtstaats eine passgenaue Lösung braucht. Die Bedingungen für Schulqualität in diesem engen städtischen Feld, in der eine Vielzahl von Schulen – damals waren es noch über 400 – existiert, sind andere als in einem Flächenstaat. Dort mögen sie in einzelnen Großstädten auch spezifisch sein, aber für diesen Stadtstaat ist kennzeichnend gewesen, dass die Wissensverteilung über das, was man glaubt, was Schule leistet, informell in verschiedenen Netzwerken organisiert ist. Damit ist das Wissen über Schulqualität eher fall- und ereignisbezogen organisiert und im Prinzip abkoppelt von dem, was man eigentlich an empirischer Wirklichkeit mit anderen Instrumenten feststellen kann. Dem muss man Rechnung tragen, wenn man mit so einer Einrichtung ins Feld gehen will. Wir haben daraus zum Beispiel für unser Modell in Hamburg die Schlussfolgerung gezogen, dass wir für die Schulinspektion Vollzeitinspektoren rekrutieren, diese ausbilden, und dass zum Team der Schulinspektoren damals zwei Wissenschaftlerstellen gehörten. Die Wissenschaftler sollten die Instrumentenentwicklung nach den Regeln der Kunst vorantreiben, die Ins-

pektoren entsprechend schulen, die Unterrichtsbeobachtungen als Kernstück der Einschätzung der Schulen entsprechend aufsetzen, da wir einen dezidiert empirischen Anspruch hatten. Als weitere „Inspektoren“ sollten dann Lehrkräfte oder vor allen Dingen Funktionsträger aus Schulen hinzugezogen werden, damit sie ihre Schulformexpertise bzw. als Betriebsvertreter ihre Berufsfeldexpertise einbringen. Aber die Verantwortung sollten von Anfang an Vollzeitinspektoren haben, die für ihre Arbeit ausgebildet sind.

Herbert Altrichter: Die speziellen Bedingungen des Stadtstaates brauchen professionelle Inspektoren und die Einbeziehung von Wissenschaftlerinnen und Wissenschaftlern zur autoritativen und methodisch-empirischen Absicherung, um so der Inspektion auch ihren Status zu geben. Ist das die Botschaft?

Norbert Maritzen: Nein, ich glaube, das ist nicht an den Stadtstaat geknüpft. Aber ich glaube, für diesen Stadtstaat, in dem das Informationsnetz über Schulen sehr dicht und die Annahme anzutreffen ist: „Ich weiß sowieso alles über die Schulen meiner Nachbarschaft“ – ich spitze das jetzt zu –, war von der Anlage des Projektes her eine Musterunterbrechung wichtig. Das heißt, einen deutlich anderen Zugriff auf beobachtbare Wirklichkeit zu nehmen als der, für den man viele Routinen hatte. Und diese Musterunterbrechung hat zu Beginn sehr verstörend gewirkt, auch weil wir gefragt wurden: Warum braucht ihr diese Wissenschaftler? Diese Diskussion haben wir im Übrigen auch in dem Startteam und eigentlich bis heute in dem Team der Schulinspektion und in der Abteilung, jetzt im Institut, in der die Schulinspektion organisiert ist, geführt. Das Spannungsverhältnis musste immer wieder austariert werden zwischen den Ansprüchen, die man an eine wissenschaftsbasierte Beobachtung von schulischer Wirklichkeit zu stellen hat, und dem, was man als Person, aus dem pädagogischen Geschäft kommend, ja auch an Routinen und Erfahrungen mobilisieren kann, um einen Blick auf die beobachtete Schule zu nehmen. Das lässt sich nicht nach der einen oder anderen Seite auflösen. Aber wir haben zunächst, um den Bruch zu kennzeichnen, stark auf standardisierte Verfahren gesetzt.

Herbert Altrichter: Du hast mir „multiprofessionelle Teams“ als wichtigen Teil des Konzepts erklärt. Ich verstehe „multiprofessionell“ jetzt so: Es gibt hauptamtliche Schulinspektorinnen und -inspektoren mit Back-up durch Wissenschaftlerinnen. Und es gibt schulische Funktionsträger als Schulformexperten, die hinzugezogen werden. Im berufsbildenden Bereich gibt es zusätzlich Wirtschaftsvertreter.

Norbert Maritzen: Wir haben auch innerhalb der hauptamtlichen Schulinspektoren im Laufe der Jahre nicht nur ehemalige Lehrkräfte gehabt, darunter ehemalige Schulleitungen oder eine Schulaufsichtsbeamtin, sondern darunter waren auch ein Erwachsenenbildner und ein Soziologe. Eine Erziehungswissenschaftlerin und ein

Erziehungswissenschaftler sind immer noch dabei. Das heißt, auch bei den hauptamtlichen Schulinspektoren gibt es eine Mischung an Professionen.

Herbert Altrichter: Und diese Mischung wurde gezielt gesucht?

Norbert Maritzen: Ja, und es wurde damals auch gezielt so ausgeschrieben.

Herbert Altrichter: Wenn ein ausländischer Beobachter die Entwicklung von Schulinspektionen in den deutschen Bundesländern beobachtet, hätte er wohl den Eindruck, Impulse aus den Niederlanden wurden in Niedersachsen aufgegriffen und dann hat jedes Bundesland versucht, ein ähnliches System mit gewissen Variationen zu produzieren. Was sind spezielle Akzente des Hamburger Systems? Einen habe ich eben aufgenommen, diese multiprofessionellen Teams, die du beschrieben hast. Gibt es noch andere Merkmale, die speziell für Hamburg sind?

Norbert Maritzen: Die intensive Einbindung der Inspektion in zwei übergreifende Aspekte scheint mir ein weiteres Spezifikum des Hamburger Systems zu sein. Zum einen wird unser Inspektorat dadurch, dass Empiriker Teil des erweiterten Teams sind und die zuständige Abteilungsleitung aus der Forschung kommt, laufend mit Entwicklungen konfrontiert, die der inspektoratsrelevante Forschungsdiskurs nimmt. Das gilt für die einzelne Inspektorin/den einzelnen Inspektor sicherlich in unterschiedlicher Weise. Aber aktuelle Forschungsansätze und -fragen sind immer wieder Thema; jedenfalls wünsche ich mir das so. Zum anderen ist die Inspektion in Hamburg eingebunden in eine Monitoring-Einrichtung (IfBQ), die weitere Aufgaben im Bereich der Qualitätsentwicklung und -sicherung für Hamburger Schulen hat. Daraus ergeben sich Ansprüche und Möglichkeiten, aber auch Herausforderungen, Schulqualität datenbasiert als komplexes, vieldimensionales Bündel von Merkmalen zu beschreiben, zu denen die Inspektion einen prozessbezogenen Teilbeitrag leistet.

Herbert Altrichter: Im Falle schulischen Qualitätsmanagements sind Grenzziehungen oft wichtig. Eine wichtige Frage ist oft: Soll Inspektion ein Teil der Behörde sein oder eine unabhängige Institution? In Hamburg hat man sich eher für eine behördennahe Lösung entschieden. Was sind für dich wichtige und unwichtige Grenzen?

Norbert Maritzen: Die Nähe zur beauftragenden Behörde ist im Entwicklungsprozess ganz unterschiedlich zu bewerten. Meines Wissens ist, jedenfalls in den deutschsprachigen Ländern, immer eine möglichst ministeriumsnahe Lösung gefunden worden. Die Haltekraft unserer Ministerien ist da ziemlich groß. Die Ambiguitätstoleranz, mit vielleicht unabhängigen Ergebnissen einer Evaluationsinstanz

umzugehen, ist hierzulande eher wenig ausgeprägt. Da spielen aber auch im internationalen Vergleich die Gründe einer unterschiedlichen Verfasstheit von Schule und Steuerungsinstanzen eine Rolle. In einem Land, in dem keine Aufsicht existiert, wie man sie in Deutschland kennt, braucht man andere Evaluationsinstanzen mit viel stärkerem „Biss“, als das in deutschen Ländern möglich und nötig ist. Ich habe im Laufe der Jahre nur wenige Beispiele erlebt, wo die vorgesetzte Behörde direkt eingegriffen und gesagt hat: Diese Schule inspiziert ihr jetzt nicht. Oder wo umgekehrt wir das Gefühl hatten, wir sind auf einen Zustand gestoßen, der, bevor er dann in einem Bericht niedergelegt wird, der Information der zuständigen Schulaufsicht bedarf. Also die Grenzüberschreitung kann ja in beide Richtungen erfolgen. Das war nur in wenigen Ausnahmefällen der Fall. Ansonsten muss man sagen, wir werden, was die Schulinspektion angeht, vonseiten der Schulbehörde mit viel Respekt und Unterstützung bedacht. Die Restriktionen ergeben sich eher aus Ressourcenfragen. Da sind wir dann im Rahmen der Haushaltsnot eine Institution unter vielen, die entsprechende Einschränkungen zu erleben hatten. Und das hat weniger mit unserem Auftrag zu tun als mit einer generellen Haushaltsenge, die die Handlungsmöglichkeiten auch beschränkt.

Herbert Altrichter: In der Broschüre über die Hamburger Schulinspektion steht: Wir beraten, welche Maßnahmen Schulen ergreifen sollen. Andere Inspektionssysteme arbeiten mit einer strikten Grenze zwischen Diagnose und Beratung. Wie wird das gesehen?

Norbert Maritzen: Sie beraten nicht im Sinne von schriftlichen konkreten Hinweisen, z. B. welche Unterrichtsmaßnahmen wie angelegt werden sollen. Was Inspektoren machen, ist, den Schulen Hinweise zu geben, wo sie die schwierigste Befundlage sehen und wo am dringlichsten eingegriffen werden sollte. Sie geben natürlich auch den Hinweis, an wen die Schulen sich wenden können. Das Landesinstitut in Hamburg hat eine entsprechende Schulbegleitagentur. An diese Agentur, die in Hamburg bekannt ist, wenden sich Schulen. Von dort bekommen sie Berater vermittelt, entweder fachspezifische oder Schulentwicklungsberater oder Coachings für Schulleitungen oder Ähnliches.

Herbert Altrichter: Schulinspektion will ja auch durch Berichte wirken. Die Hamburger Schulinspektion veröffentlicht als Institution einen jährlichen Bericht. Daneben gibt es die individuellen Schulberichte. Wer liest die? Werden die, auch in der Stadt, zur Kenntnis genommen? Lässt sich verfolgen, wie Konsequenzen zustande kommen?

Norbert Maritzen: Die schulbezogenen Inspektionsberichte werden zunächst in einer vorläufigen Fassung der Schulleitung zur Rückkopplung gegeben. Dann wird

entsprechend eine Endfassung an die Schulleitung gegeben mit der Maßgabe, dass dieser Bericht den schulischen Gremien zur Kenntnis zu geben ist. Das heißt, dass die Ergebnisse kommuniziert werden müssen. Zeitgleich erhält die zuständige Schulaufsicht den Bericht. Das wissen die Schulen. Es gibt ein Response-Verfahren, in dem festgelegt wird, in welchen Zeiträumen und in welcher Schrittigkeit die Schulaufsicht auf die Schulen zugeht und nachfragt, wie die Schule mit den Ergebnissen umgeht und welche Konsequenzen geplant sind. Der Bericht muss und wird also schulintern gelesen. Seit dem zweiten Durchlauf der Inspektion wird die Zusammenfassung des Berichts ins Internet gestellt und lässt sich auf einem Schulportal der Hamburger Schulbehörde abrufen. Um diese Entscheidung gab es heftige Diskussionen, weil es auch ein Rückfall hinter eine von Vorgängern gegebene Zusage war. Unsere Wahrnehmung ist die, dass das ziemlich geräuschlos passiert ist. Mir ist bis jetzt – die Veröffentlichungen laufen knapp zwei Jahre – kein Fall bekannt, in dem ein Skandal entstanden ist, nur weil die Information nun öffentlich abrufbar ist. Der systemische Bericht, der Teil unseres Systems der Bildungsberichterstattungen ist, wird allen Schulen, den Aufsichten der Behörden, allen Bildungsinstitutionen und auch der breiteren Öffentlichkeit zur Verfügung gestellt. Er war bisher auch immer Gegenstand der Beratung im Schulausschuss der Bürgerschaft.

Herbert Altrichter: Die Berichte werden also von der Politik wahrgenommen. Gibt es auch eine mediale Öffentlichkeit dafür?

Norbert Maritzen: Ja, wir haben den Bericht bislang immer in einer Landespressekonferenz vorgestellt. Da passiert dann das, was immer nach solchen Pressekonferenzen passiert. Aus der Presseerklärung des Senats bedient sich die Presse sehr selektiv und greift bestimmte Einzelaspekte heraus. So war bei dem vorletzten systemischen Bericht die Headline, dass die von uns mit unseren Instrumenten gemessene Unterrichtsqualität, wenn man den sozialen Hintergrund berücksichtigt, keinen Zusammenhang aufweist zur Schulform und die innerschulischen Disparitäten die größten Unterschiede ausmachen. Zwei Tage später geht die Presse zum nächsten Thema über. In wirklich wenigen Einzelfällen hat es ein Presseecho gegeben. So hat in Folge eines sehr negativ ausgefallenen Berichts eine Schulleiterin um Entbindung von ihrer Funktion gebeten.

Herbert Altrichter: Wir haben vor einiger Zeit eine Analyse verschiedener Inspektionssysteme in Europa durchgeführt und dabei Repräsentanten der Inspektorate nach den Wirkungsmodellen hinter ihrer Arbeit gefragt. Wie wirkt nun die Hamburger Schulinspektion? Was sind deiner Meinung nach die Wirkungsmechanismen für die Stimulation von Entwicklung im System?

Norbert Maritzen: Die Schulinspektion erfolgt aus Schulsicht alle vier Jahre. Angesichts unserer personalen Not ist das Intervall momentan sogar noch deutlich länger. Das ist für die Einzelschule ein exzeptionelles Ereignis. Aber mit Blick auf Transparenz hat das keine lange Wirkungskraft. Was wichtig ist, ist, dass die Schulinspektion die Instanz ist, die gleichsam die „normative Bugwelle“ vor sich herschiebt, die in unserem Orientierungsrahmen als Schulqualität formuliert ist, und dafür sorgt, dass dieser Referenzrahmen auch in der Aufmerksamkeit der Schulen mindestens eine temporäre Relevanz hat. Insofern ist das, was Norbert Landwehr „normensetzende Funktion der Inspektionen“ nennt, denke ich, sehr stark. Dies allerdings in einem Kontext mit weiteren Instrumenten des Monitorings, die wir in Hamburg auch haben. Z.B. KERMIT, also nicht nur VERA 3 und VERA 8 wie in den Bundesländern, sondern sechs Vollerhebungen jedes Jahr in allen Klassen in mindestens drei Domänen mit Längsschnitten und den entsprechenden datengestützten Feedbacks, die die Schulen erhalten. Außerdem die Rückmeldungen zum Sprachfördermonitoring und die Auswertungen zu zentralen Prüfungen und anderes mehr. Schulen erhalten also einen ganzen Kranz von Informationen, in die sich sozusagen in einem anderen Intervall und mit einem anderen Fokus die Information über die Inspektionen einordnen lassen. Wir überlegen jetzt, was etwa Schulleistungsinformationen für eine Einschätzung der Qualität an der Schule bedeuten. Und wir beginnen, systematisch die Längsschnitte in der Schulleistungsentwicklung in den einzelnen Schulen zu modellieren und zurückzumelden. Damit kann sozusagen der *Added Value*, für den die Schule verantwortlich ist, mindestens ansatzweise in den Blick genommen werden. Erst dies ermöglicht der Schulinspektion, Schulleistungsergebnisse zu nutzen, weil sie der Schule als Arbeitsergebnis zurechenbar sind.

Herbert Altrichter: Das ist für mich systemisch eine interessante Frage. Hamburg ist ja sehr reich ausgestattet mit empirischen Informationen über das Schulsystem. In welchem Verhältnis steht Schulinspektion zu den verschiedenen anderen laufenden Leistungs- und Qualitätserhebungen? Ist es auch eine Funktion der Schulinspektion, die verschiedenen Daten im Hinblick auf einzelne Standorte zu integrieren?

Norbert Maritzen: Historisch ist erst im zweiten oder dritten Schritt versucht worden, diese unterschiedlichen Instrumente der Systembeobachtung auf den unterschiedlichen Ebenen auch institutionell zu integrieren, eigentlich erst mit der Gründung des IfBQ vor zwei Jahren. In dieser Hinsicht stehen wir vielleicht vor der größten Entwicklungsaufgabe, die das Institut hat, nämlich diese Integrationsleistung zu vollziehen. Ich glaube nicht, dass das Aufgabe der Schulinspektion ist, die differentiellen empirischen Befundlagen zu einer einzelnen Schule zu integrieren. Wir haben deshalb auf diese herausfordernde Fragestellung ein abteilungs-

übergreifendes Projekt angesetzt, in dem aus unterschiedlicher Perspektive, auch aus Schulperspektive, überhaupt erst mal analysiert werden soll: Was bedeutet das eigentlich, wenn Schulen Daten ganz unterschiedlicher Qualität zu unterschiedlichen Zeitpunkten in unterschiedlichen Intervallen zur Verfügung gestellt bekommen? Wie ist eigentlich die Erlebnisperspektive einer Schulleitung, die das auf den Schreibtisch gepackt kriegt? Welche Möglichkeiten gibt es überhaupt, eine integrale Sicht vorzunehmen? Ist dieser Anspruch vielleicht in bestimmten Teilaspekten einlösbar, in anderen aber völlig illusorisch, auch konzeptionell nicht schlüssig? Das sind Fragen, mit denen wir uns gerade intensiv zu beschäftigen beginnen unter der Perspektive, dass wir Schulen damit nicht alleine lassen dürfen. Die brauchen Unterstützung. Und wir haben als IfBQ im Prinzip eine Bringschuld, zu einer Fokussierung und Konzentration der einzelschulischen Datenlage zu kommen. Da bietet die institutionelle Verfasstheit des Instituts, also dass diese Instrumente alle „aus einer Firma“ kommen, eine ganz gute Chance.

Herbert Altrichter: Ein wichtiger Baustein von Inspektionssystemen sind Orientierungsrahmen. Diese sind in einer gewissen Weise ziemlich ähnlich über die Bundesländer und Systeme hinweg, aber dann auch wieder speziell in einzelnen Aspekten. Wie ist es in dem Fall von Hamburg, wie ist der Orientierungsrahmen entstanden und hat er irgendwelche speziellen Akzente oder Botschaften?

Norbert Maritzen: Der erste Orientierungsrahmen ist von der eingangs erwähnten Projektgruppe Schulinspektion entwickelt worden, weil uns sehr schnell klar war, dass wir einen Referenzrahmen brauchen, mit dem wir den Blick der Inspektion auf die Schule strukturieren. So konnten wir bestimmte Aspekte aus pragmatischen Gründen auswählen, und andere, die auch relevant sind, aber nicht Teil des Inspektionsverfahrens sein können, nicht. Das war konzeptionell zu Beginn die schwierigste Arbeit. Dieses Erfordernis haben wir damals der Ausformulierung eines Inspektionsverfahrens vorgeschaltet. Wir haben uns die damals existierenden Beispiele angeschaut. Es war nicht so wichtig, dass wir etwas Neues erfinden. Wichtig war der Diskussionsprozess, in den auch Schulen involviert waren. Ich bin damals mit der Entwurfsfassung gemeinsam mit einem Kollegen monatelang in einer „Roadshow“ durch die Schulen und Schulleiterkonferenzen gezogen. Dieser Prozess war wichtig. Wir haben dann bestimmte konzeptionelle Entscheidungen getroffen, die andere Länder anders getroffen haben. Z.B haben wir den ganzen Bereich der Kontextqualität rausgelassen.

Herbert Altrichter: Warum?

Norbert Maritzen: Unsere Perspektive war, dass wir uns auf das fokussieren, worauf die Schule Einfluss hat. Also das, wo in der Schule Voraussetzungen geschaffen

werden können, um auf die Schulqualität, Prozessqualität und Ergebnisqualität zu wirken. Das heißt, wir haben uns entschlossen, erst gar nicht die Möglichkeit der externalisierenden Zuschreibung von Defiziten zu schaffen. Also, nach dem Motto: „Eigentlich wären wir ganz gut, aber in diesem Kontext geht das gar nicht.“ Diese Diskussion haben wir abgeschnitten. Das bedeutet keine Ignoranz gegenüber den Rahmenbedingungen, unter denen Schule arbeiten muss, sondern eine Fokussierung der Funktionen eines Orientierungsrahmens.

Herbert Altrichter: Wobei man auch sagen könnte, dass diese Kontextinformationen stärker den Schulträger und die Behörde prüfen, die den Rahmen zur Verfügung stellt.

Norbert Maritzen: Richtig. Mit dieser Diskussion hatten wir auch zu tun, gerade mit den Schulleiterverbänden, weil die sagten: Ihr nehmt gar nicht in den Blick, wie uns die Behörde am ausgestreckten Arm verhungern lässt. Ich glaube, das wäre für die Impulsgebung in die schulinterne Diskussion, um das „was können wir tun“ hinderlich gewesen, wenn wir dieses Diskussionsfeld aufgemacht hätten. Uns haben auch die Ergebnisse der schon damals vorliegenden empirischen Untersuchungen im Schulbereich gezeigt, welche innerschulischen und zwischen-schulischen Qualitätsunterschiede angesichts gleicher oder hoch vergleichbarer Ausgangsbedingungen existieren. Die Variabilität zwischen Schulen ist erheblich. Wir haben ja seit Anfang der 1990er Jahre intensiv die Schule beforscht in vielen Auftragsforschungsprojekten. So begannen wir seit den 1990er Jahren, Sozialindizes zu entwickeln. Damit konnten wir Schulen mit ähnlichen sozialen Belastungen miteinander vergleichen. Insofern hat dann auch die Diskussion darüber, dass wir diesen Bereich aus dem Orientierungsrahmen ausblenden, gar nicht lange gedauert. Das war relativ schnell vom Tisch, weil wir immer sagen konnten: Wir melden euch Ergebnisse zurück und wir beachten sozusagen die Rahmenbedingungen, unter denen ihr arbeiten müsst.

Herbert Altrichter: Du hast gesagt, du hast die Entwürfe zum Orientierungsrahmen bei Konferenzen vorgestellt. Wie waren die Rückmeldungen von Schulleitungen? Sind da neue Akzente genannt worden oder war der Entwurf etwas, womit Schulleitungen sich anfreunden konnten?

Norbert Maritzen: Deutlich ist damals auf jeden Fall geworden, dass viele Formulierungen in diesem Prozess realistischer wurden, insbesondere, weil wir viel mit Schulleitungen gesprochen haben, was die Erwartungen an Schulleitungshandeln angeht. Es ist auch differenzierter geworden bei uns. Also hat es noch mal einen Überarbeitungsprozess gegeben. Die Grundstruktur ist aber nicht verändert worden.

Herbert Altrichter: Ich hatte einmal den Auftrag, mit den österreichischen Auslandsschulen einen Prozess zur Entwicklung eines Referenzrahmens in Gang zu bringen. Wir hatten die Idee, den hessischen Referenzrahmen vorzulegen. Wir baten die Projektgruppen dieser Auslandsschulen, die aus Schulleitungen und zwei Lehrpersonen bestanden, diese Vorlage, die uns als zu lang, zu differenziert erschien, für ihre Situation umzuschreiben. Zu meinem Erstaunen traten eigentlich wenige Wünsche nach Veränderungen und Kürzungen auf und wenn, dann betrafen sie die speziellen, z. B. kulturellen Bedingungen, die Auslandsschulen haben. So kamen wir eigentlich überraschend schnell zu einer Übereinstimmung, was wichtige Aspekte der Schulqualität sind.

Norbert Maritzen: Das war auch unsere Erfahrung. Wir haben bei diesen Rückkopplungsschleifen bei Schulleitungen, innerhalb der Behörde und auch im Landesinstitut festgestellt: Der Rahmen bleibt. Es gibt dann allerdings Wünsche einzelner Bereiche. Dann fehlt den Schulleitungen aus dem Kiez z. B. die ganze Drogenprävention. Da mussten wir eine mittlere Abstraktionsebene finden, bei der die Kolleginnen und Kollegen mit unterschiedlichen Interessen sagen konnten: Ach so, unter dieser Dimension kann ich verankern, dass an meiner Schule Drogenprävention ganz wichtig ist. Das war keine leichte Kunst, solche Formulierungen zu finden. Wir haben ja vor zwei Jahren eine veränderte, überarbeitete Fassung des Orientierungsrahmens vorgelegt. Da haben wir wieder diese Schleifen gemacht, nun vor dem Hintergrund der Erfahrung des Umgangs mit der ersten Fassung. Da haben wir auch noch mal auf die Struktur geschaut und versucht, ein Organisationsprinzip zu finden, auch in Anlehnung an Ergebnisse der Schulforschung, um das zu fokussieren, was nah am Unterricht ist, und dem einen hohen Stellenwert zu geben. Die schulischen Ergebnisse, nicht nur Leistungsergebnisse, in die Mitte eines Kreismodells zu setzen und das, was nah am Unterricht ist, dort ran zu setzen und sozusagen, je weniger nah am Unterricht die Merkmale sind, sie auch konzentrisch weiter außen anzuordnen. Das ist noch mal ein neuer Gedanke gewesen.

Herbert Altrichter: Als ich jetzt den Hamburger Orientierungsrahmen gelesen habe, war mein Eindruck, dass er, im Gegensatz zu vielen anderen Orientierungsrahmen, etwas anders organisiert ist. Viele andere sind stark nach dem Input-Prozess-Output-Modell angeordnet. Im Hamburger Rahmen war mein Eindruck eine stärker prozesshafte Sprache, von Führungsmerkmalen und den Unterrichtsaspekten her organisiert. Wollte man näher an der schulischen Tätigkeit sein?

Norbert Maritzen: Genau. Und das war auch die Fokussierung der Schulinspektion auf das Leitungshandeln, Führungshandeln und auf Unterricht selbst und die Schaffung der Voraussetzung für die Unterrichtsentwicklung in der einzelnen Schule. Das sind die beiden zentralen Merkmale des Fokus der Schulinspektion.

Herbert Altrichter: Hat es seit der Einführung der Schulinspektion 2006 wichtige Umsteuerungen gegeben? Und was waren die Gründe dafür?

Norbert Maritzen: Eine große Veränderung hat es mit der Novellierung des Orientierungsrahmens gegeben. Das hat für die Schulinspektion bedeutet, dass das ganze Instrumentarium angepasst werden musste, die Gewichtung der einzelnen Bereiche verändert werden musste. Das war eine deutliche konzeptionelle Änderung. Damit einher ging für den jetzt laufenden zweiten Inspektionszyklus eine deutliche Haltungsänderung. Mit der inzwischen erworbenen Expertise referieren die Inspektorinnen und Inspektoren nicht nur die erhobenen Daten, sondern setzen sie auch expliziter miteinander in Beziehung und deuten sie. Die Ergebnismeldung ist deshalb dialogisch angelegt, sie gibt den Schulbeteiligten und den Inspizierenden die Möglichkeit zum Austausch der eigenen Wahrnehmung des Untersuchungsgegenstandes. Inspektionsergebnisse erhalten auf diese Weise den Status eines Deutungsangebots, das nur im Dialog mit der Schule „zum Sprechen gebracht“ werden kann.

Ein zweiter, schon früherer Einschnitt war, die normative Festlegung, dass, wenn Schulen bestimmte Erwartungen unterschreiten, ein spezielles Verfahren unter Einschluss der Schulaufsicht in Gang gesetzt wird. Also die ganze Diskussion um „*Failing Schools*“, die wir da auch hatten, war ein wichtiger Schritt, weil dieser Schritt auch in so einem Stadtstaat nicht geheim bleiben kann. Wenn die Inspektion der Aufsicht signalisiert, hier ist Interventionsbedarf, dann findet das tendenziell in der Öffentlichkeit statt. Das ist bisher nie in der Stadt skandalisiert worden, aber das hat mit Blick auf die Innenwirkung in die Inspektion herein auch noch mal für eine Veränderung der Aufmerksamkeit der Inspektoren gesorgt. Sie standen unter größerem Druck, die Bewertungen vertreten zu können. Und manchmal hatte ich das Gefühl, es gab eine Tendenz zur Mitte und eine Scheu, die damit verbunden ist, den *Cut-off* zur Feststellung des Nicht-Tolerablen zu überschreiten. Es gab dann unweigerlich einen hohen Rechtfertigungsdruck.

Herbert Altrichter: Aber das heißt, dass die Schulinspektion die Verantwortung hat, einen solchen Prozess anzustoßen.

Norbert Maritzen: Ja. Die Schulinspektion war Auslöser dieses Prozesses.

Herbert Altrichter: Ihre Funktion ist also nicht bloß, einen Bericht an die Schulaufsicht zu schreiben?

Norbert Maritzen: Genau. Unser Vorgehen war ein festgelegtes Verfahren mit Kriterien, ab wann eine Schule als „*Failing School*“ galt. Das wussten alle Schulen. Das sich an so eine Bewertung anschließende Verfahren war formal geregelt und

fand unter der Leitung des Leiters des Amtes für Bildung statt, das heißt dem direkten Vorgesetzten der Leitungen der Schulaufsicht. Es war also hoch angesiedelt mit einer hohen Aufmerksamkeit. Inzwischen ist das Verfahren stärker präventiv orientiert. Fallkonferenzen werden in einem gestuften Verfahren unter Beteiligung der betroffenen Schulleitungen bei kritischen Befunden schon früher angesetzt, um die Lage zu analysieren und den spezifischen Unterstützungs- und Beratungsbedarf zu sondieren. Nach ersten Erfahrungen betrifft das etwa 20 Prozent der inspizierten Schulen.

Herbert Altrichter: Wir waren bei der Frage: Hat es Änderungen, Umsteuerungen gegeben?

Norbert Maritzen: Die Veröffentlichung war auch eine Änderung, die im Vorfeld in der Bürgerschaft zwischen den Parteien politisch kontrovers diskutiert wurde, aber auch in der Inspektion und im Institut hinsichtlich der Implikationen. Umso überraschender ist, dass die Veröffentlichung jetzt relativ geräuschlos vonstattengeht.

Herbert Altrichter: Wie erklärst du dir das?

Norbert Maritzen: Ich sage immer, wo viel Information ist, ist auch viel Vergessen. Wir wissen bisher nicht, inwiefern die Inspektionsberichte, insbesondere kritische Inspektionsberichte, einen Einfluss auf das Schulwahlverhalten der Eltern haben, ob das darüber überhaupt beeinflusst wird, ob da nicht ganz andere Kriterien eine Rolle spielen. Manche Eltern ziehen mit dem Zirkel einen Kreis um ihren Wohnort und gucken dann, welche Schulen dort sind, weil sie ihre Kinder nicht durch die Stadt zu einer vermeintlich besseren Schule karren können. Aber in einem System freier Schulwahl wird oft vermutet, dass auch Inspektionsberichte einen entsprechenden Einfluss haben. Wir wissen das nicht genau. Über eine anekdotische Evidenz hinaus, dass man vielleicht mal einen Elternteil getroffen haben kann, der sagt: „Ich habe den Inspektionsbericht gelesen und dann gedacht, nein, lieber nicht“, wissen wir nichts Systematisches.

Herbert Altrichter: Hamburg hat in dieser Zeit überraschend viele Veränderungen seiner politischen Verhältnisse erlebt. Hatte das, die unterschiedlichen Stadtregierungen, einen Einfluss auf die Entwicklung der Schulinspektion?

Norbert Maritzen: Ich glaube, es gibt mit Blick auf die Inspektion in der einen oder anderen Variante einen parteiübergreifenden Konsens. Bisher jedenfalls habe ich bei den Linken als einziger Partei immer wieder gehört: Das braucht man eigentlich nicht, das ist überflüssig. Unsere Einladung an die schulpolitischen Sprecher der Linken in der Bürgerschaft: „Geht mal mit, schaut euch das an, kommt zu uns“

hat nicht dazu geführt, sich durch eine größere Kenntnis der Sachlage von den eigenen Vorurteilen zu entfernen. In den anderen Parteien ist das anders. Hilfreich ist, glaube ich, gewesen, dass die Strategie vor 2010 gewesen ist, die Schulinspektion außerhalb der eigentlichen Schulbehörde zu institutionalisieren und dort in dieser Einrichtung auch weitere Aufgaben des Bildungsmonitorings zu institutionalisieren und ein stringentes Konzept zu verfolgen. Das hat in so einem kleinen Stadtstaat, in dem die Inspektion keine große Behörde ist wie in großen Flächenländern, die Möglichkeit geschaffen, für einen größeren professionellen Austausch und für eine institutionelle Absicherung zu sorgen. Damit wurde ein Potenzial für eine größere Wirksamkeit geschaffen, indem man für eine integrale Sicht auf Schulqualität sorgt.

Herbert Altrichter: Man kann Schulinspektion als eines der Elemente der neuen Steuerung ansehen, der evidenzbasierten Steuerung des Bildungswesens, die ihrerseits stark mit dem PISA-Schock verbunden ist. Mein Eindruck ist, dass in vielen deutschen Bundesländern nach einer Phase medialer Aufregung eine entschiedene Entwicklungsphase stattgefunden hat. Und aktuell kommt mir vor, dass eine erste Müdigkeitsphase da ist. Wie betrifft das Schulinspektion? Wie siehst du die weitere Entwicklung der Schulinspektion?

Norbert Maritzen: Ich habe damals jedenfalls die PISA-Aufregung gar nicht so sehr bei der Gründung der Schulinspektion als Begleitmusik in der konkreten Projektplanung wahrgenommen. Das hängt aber auch damit zusammen, dass wirklich sehr kritische bis desaströse Schulleistungsergebnisse in Hamburg längst vor PISA bekannt waren, weil Hamburg selbst längsschnittlich angelegte, flächendeckende Schulleistungsuntersuchungen in Auftrag gegeben hatte. Und da sind uns unsere schulformbezogenen sozialen Disparitäten sehr deutlich vor Augen geführt worden. Die Diskussion hat nicht erst mit der Veröffentlichung von PISA 2001, sondern bereits 1997/1998 begonnen. Wir waren durch die ersten Schrecken, wenn man genau auf Ergebnisse schaut, schon durch. Das war für die Gründung der Schulinspektion insofern hilfreich, als die Inspektion in Hamburg nicht mit dem Vorzeichen gestartet ist, wir schneiden so schlecht in unseren Schulen ab, jetzt müssen wir mal genauer hingucken. Sie ist vielmehr aus dem Kontext entstanden: Was wissen wir schon über die Hamburger Ergebnisse? Und da gab es noch ganz andere Schwierigkeiten. Wir wussten z. B. sehr genau, wie hoch die Schulabbrecherquoten pro Standort sind. Das heißt, die Gründung der Inspektion war von der Zielsetzung in einen breiter angelegten Qualitätsdiskurs eingebettet. Und da war ganz günstig, dass wir die Diskussion um den Orientierungsrahmen davor gestalten haben, weil dadurch eine andere Fokussierung möglich war. Die Ausgangslage in Hamburg war also anders als in manch anderen Ländern, die in der Folge von PISA schnell Inspektionen eingeführt haben, so als sei die Schulaufsicht beim Nichtwissen erwischt worden. Jetzt mit Blick auf die weitere Entwicklung ist es in der Tat ja

so, dass in manchen Bundesländern so etwas wie eine Müdigkeit einkehrt. In Hamburg gibt es das politisch auch, also eine Diskussion unter der Leitfrage: „Müssen wir das eigentlich so genau wissen?“ Der Senat wird im Laufe einer Legislaturperiode auch im Schulbereich mit vielen Kleinen und Großen Anfragen traktiert, von denen im Schulbereich nicht wenige auf der Basis von einzelschulbezogenen Daten beantwortet werden müssen. Bisher hat sich das aber nie auf Ergebnisse der Schulinspektion bezogen. Das haben wir in anderen Bereichen schon tun müssen. Wir sind beispielsweise gezwungen worden, Schulleistungsergebnisse schulscharf zu veröffentlichen.

Herbert Altrichter: Was sind eigentlich kritische Entwicklungsfragen für die Schulinspektion?

Norbert Maritzen: Die kritischen Entwicklungsfragen betreffen zum einen die Daueraufgabe, eine Balance multipler Zielerreichung herzustellen. Schulinspektion ist und bleibt gleichsam eingezwängt in Funktionszuschreibungen, die sich zum Teil widersprechen oder miteinander konkurrieren: Kontrollfunktion, Impulsfunktion für die Schulentwicklung und Erkenntnisfunktion für die Systemsteuerung vertragen sich nur bedingt. Das Menü ist aber nicht einfach zu bereinigen, weshalb von Inspektorinnen und Inspektoren gewissermaßen eine Ambivalenz-Kompetenz verlangt wird.

Zum anderen ist der Aspekt der Nutzung der Ergebnisse zentral, und zwar im Kontext der Nutzung weiterer schulbezogener Daten. Das ist ein kritischer, zentraler Schritt, finde ich. Wenn die Schulinspektion vonseiten der Schulen als nützlicher Feedbackgeber betrachtet werden soll, muss sie sich auch in den Formen der Rückmeldung, in den Formaten, so entwickeln, dass in den Schulen das mögliche Potenzial für Entwicklungsimpulse erkannt werden kann. Das können wir nur in Kooperation mit anderen Akteuren. Das kann die Schulinspektion nicht alleine tun. Hier haben wir in den letzten Jahren ganz viel Arbeit geleistet, z. B. um die Schnittstelle zur Schulaufsicht glatt zu kriegen, das heißt, auch die Folgeprozesse mit in den Blick zu nehmen. Das war zentral, weil uns deutlich geworden ist, dass man einen Legitimationsverlust für die Inspektion erleiden wird, wenn einem nicht nur an der Nutzbarkeit, sondern auch an der Nutzung der Ergebnisse gelegen ist. Wir haben auch in den Jahren seit der Einführung zunächst in der Inspektion den Kurs gefahren: Wir schauen uns die Wirklichkeit an, schreiben einen Bericht, geben die Daten inklusive Bewertung zurück und sind dann wir raus aus dem Geschäft. Also eine klare Trennung zwischen Inspektion als Diagnoseinstrument auf der einen und Schulleitungen, Schulaufsichten, Unterstützungseinrichtungen als „Therapeuten“ auf der anderen Seite. Ich glaube, das war für die Rollenklärung zu Beginn notwendig. Aber mit zunehmender Erfahrung wird deutlich: Wir sind auch für die Nutzung unserer Ergebnisse der Inspektion und anderer datengestützter

Rückmeldungen in den Schulen verantwortlich und müssen dafür Voraussetzungen schaffen.

Schließlich ist die einzelschulspezifische, qualitätsbezogene Informationsrückkopplung alle vier bis sechs Jahre wenig sinnvoll, wenn sie nicht eingebettet ist in ein differenziertes Programm verschiedener Informationsrückkopplungen. Das ist die größte Herausforderung – für das Institut übrigens auch, in dem diese Rückkopplungsformate unterschiedliche Qualitätsaspekte der Schule betreffend generiert werden. Ich will das nicht dramatisieren, aber das ist eine strategische Überlebensfrage. Das muss uns gelingen, wenn wir keinen Legitimationsverlust erleiden wollen. Wir können die Einzelmeldungen für die Inspektionsberichte oder für Schulleistungsergebnisse oder für Sprachfördermaßnahmen noch weiter ausfeilen, noch differenzierter, noch umfangreicher machen. Die Verarbeitungskapazität der Schulen wird dem nicht so einfach folgen können. Und auch die Verarbeitungskapazität eines Ministeriums, einer Schulbehörde wird der Differenziertheit systemischer Berichte nicht einfach folgen können. Es führt kein direkter Weg von dieser Rückmeldung zu irgendeiner Vorstellung darüber, was ein sich anschließendes Handlungsprogramm sein könnte.

Herbert Altrichter: In meiner Beobachtung wurden in manchen europäischen Ländern Schulen – in einem nächsten Schritt der Entwicklung der Schulinspektion – differenzierter behandelt. Das heißt vor allem, die Qualität von Schulen vorab durch *Desk Research* zu eruieren und dann nur „kritische“ Schulen intensiveren Inspektionen zu unterziehen. Daten für solche Vorab-Analysen wären in Hamburg ja in reichem Maße vorhanden. Ist das ein System, das auch für Hamburg passt?

Norbert Maritzen: Ja, das könnte passen. Wir haben vor Beginn des zweiten Zyklus gesagt, wir sollten uns hierfür eine Veränderung der Strategie überlegen. Also nicht *Full Inspection* in gleicher Weise für alle Schulen, sondern nach unterschiedlichen Kriterien ausdifferenziert. Wir haben uns die nationalen und internationalen Modelle sehr genau daraufhin angeschaut und der Behördenleitung damals einen entsprechenden Vorschlag mit möglichen Optionen gemacht. Es ist dann politisch entschieden worden, dass eine zweite Runde Vollinspektion für alle durchgeführt wird. Wir haben die Vorschläge und Optionen, die wir vor drei Jahren schon einmal formuliert haben, jetzt auch noch mal analysiert und weiter modifiziert und erste Überlegungen für ein verändertes, stärker fall- und ausgangslagenbezogenes Vorgehen vorgelegt. Da gibt es die unterschiedlichsten Varianten, in denen man das machen kann. Allerdings verändern sich mit einer solchen Strategie auch die methodischen Voraussetzungen dafür, Aussagen über die Systementwicklung zu machen. Man wird dann nicht mehr dadurch, dass man aus der Grundgesamtheit der inspeziierten Schulen eine repräsentative Stichprobe für jedes Jahr zieht, über deren Ergebnisse systemübergreifende Aussagen machen können.

Herbert Altrichter: Was ist noch wichtig zum Verständnis der Entwicklung der Hamburger Schulinspektion, das wir bisher nicht angesprochen haben?

Norbert Maritzen: Was uns auch wichtig ist – das sage ich nicht nur dir als Schulentwicklungsforscher – ist, dass wir eigentlich immer einen engen Bezug zur Schulforschung, zur Schulinspektionsforschung, zur Forschung im Bereich des Monitorings, zu den Kolleginnen und Kollegen aus dem SteBis-Zusammenhang (Steuerung im Bildungssystem, Anm. der Hrsg.) in diesem Bereich hatten. Es ist uns immer sehr wichtig gewesen, von dort Impulse zu beziehen. Die Zügigkeit, mit der wir im Inspektionsteam Vorschläge für eine Weiterentwicklung, Ausdifferenzierung von Inspektionsverfahren haben entwickeln können, hängt auch ganz stark mit der Einbindung in Forschungsdiskurse zusammen. Das ist ein Spezifikum, das in Hamburg eine Tradition seit den 90er Jahren hat.

Herbert Altrichter: Das trifft sich mit meiner Beobachtung, dass man Mitglieder der Hamburger Schulinspektion auch bei wissenschaftlichen Tagungen sieht. Die wissenschaftlichen Mitarbeiterinnen und Mitarbeiter sind wahrscheinlich für die Inspektion eine Brücke zur Bildungsforschung.

Norbert Maritzen: Ja, und diese Diskurse sind innerhalb der Inspektion und des Instituts nicht einfach, sondern sehr kontrovers, weil die Perspektiven andere sind. Die Kollegen, die die wissenschaftliche Unterstützungsleistung entwickeln, die haben häufig eine andere Perspektive als der Inspektor, der unter dem täglichen Handlungsdruck zu entscheiden hat. Aber ein solches Spannungsmoment ist, glaube ich, unverzichtbar, damit es Entwicklung geben kann.

Herbert Altrichter: Herzlichen Dank für das Gespräch.

Aufbau und Rolle der Schulinspektion Hamburg

Martina Diedrich

Zusammenfassung

Der Beitrag beschreibt den Auftrag, die Struktur und die Funktionsweise der Hamburger Schulinspektion im ersten Zyklus. Grundlegende Funktionen von Schulinspektionen sind demnach 1. die Unterstützung von Schulentwicklung, 2. die Erkenntnisgenerierung, 3. die staatliche Gewährleistung und 4. die Normendurchsetzung. Schulinspektion wird als besondere Form der Evaluation aufgefasst, die unterschiedliche Paradigmen verbindet, womit unweigerlich Zielkonflikte einhergehen. Ausgehend vom Orientierungsrahmen Schulqualität werden die Bewertungsgrundlagen der Schulinspektion und davon abgeleitet die eingesetzten Instrumente und das Verfahren beschrieben. Mit Blick auf die Professionalität der Inspektorinnen und Inspektoren erläutert der Beitrag die erforderlichen Kompetenzen, das Ausbildungscurriculum und die Formate zur Weiterentwicklung der Schulinspektion. Die Hamburger Schulinspektion hat sich ein umfassendes System des internen Qualitätsmanagements gegeben, das interne und externe Formen der Evaluation miteinander verbindet. Abschließend wird die besondere Rolle der Schnittstellenpartner im System, insbesondere der Schulaufsicht und des Fortbildungs- und Unterstützungssystems, diskutiert.

Gegründet im Jahr 2006, blickt die Schulinspektion Hamburg mittlerweile auf eine achtjährige Geschichte zurück, in deren Verlauf sich trotz der Konstanz des grundsätzlichen Auftrags ihre normativen Grundlagen und ihr Vorgehen, aber auch ihr Selbstverständnis zum Teil deutlich gewandelt haben. Mit fortschreitender Erfahrung und sich wandelnden Rahmenbedingungen verändern sich die an die Inspektion gestellten Erwartungen, erfahren die zugrunde liegenden Gewissheiten leichte Verschiebungen oder auch Revisionen und entstehen neue Handlungsformen. Vor diesem Hintergrund ist die nachfolgende Betrachtung sozusagen historisch angelegt: Sie beschreibt, wie der behördliche Auftrag, alle Hamburger Schulen standardisiert extern zu evaluieren, im ersten Inspektionszyklus umgesetzt wurde. Wie die Schulinspektion aktuell im zweiten Inspektionszyklus verfährt, wird im Beitrag „Ausblick: Zweiter Zyklus der Schulinspektion Hamburg“ in diesem Band geschildert.

1. Auftrag und Ziele

Die Schulinspektion Hamburg wurde im Jahr 2006 gegründet, zunächst als eigenständige Einrichtung. Kurz darauf wurde sie als Abteilung in das ebenfalls neu gegründete Institut für Bildungsmonitoring integriert. Ihr Auftrag war und ist, alle staatlichen Hamburger Schulen mit einem standardisierten Verfahren innerhalb von vier Jahren extern zu evaluieren. Damit ist die Schulinspektion Teil einer Gesamtarchitektur schulischer Qualitätssteuerung, die in etwa zeitgleich in allen Bundesländern als Reaktion auf das schlechte Abschneiden deutscher Schülerinnen und Schüler bei PISA aufgesetzt wurde und insbesondere auf die Steuerung durch Ergebniserwartungen setzte (sog. „Output-Steuerung“, vgl. z.B. Quesel et al. 2011). Eng mit dieser Entwicklung verbunden ist der Paradigmenwechsel zur *Selbstverantworteten Schule*, der nicht nur in Hamburg der Gründung der Schulinspektion vorausgegangen war, insbesondere weil damit eine gestiegene Bedeutung von Verfahren der Rechenschaftslegung einherging (vgl. Maritzen 2009).

Die Aufgabe der Schulinspektion liegt in der Erhebung, Bewertung und Rückmeldung einzelschulischer Qualität. Damit wird unweigerlich auf die Notwendigkeit verwiesen, Qualität inhaltlich zu bestimmen (vgl. Heid 2000). In Hamburg erfolgt eine solche Festlegung entlang der Qualitätsdimensionen, -bereiche und -merkmale des sogenannten *Orientierungsrahmens Schulqualität*. Mit ihm wird eine normative und zugleich empirisch gestützte Grundlage hergestellt, damit sich die Zuschreibung schulischer Qualität auf transparente, nachvollziehbare und intersubjektiv geteilte Kriterien stützt. Daneben soll die Schulinspektion schulische Qualität auf der Ebene des gesamten Hamburger Schulsystems sichtbar machen, indem sie die einzelschulischen Daten im Rahmen ihres Jahresberichts aggregiert und übergreifende Tendenzen, Stärken und Problemlagen ausweist.

Drei Funktionen konturieren die Zweckbestimmung der Schulinspektion in ihrem ersten Zyklus:

1. *Katalyse*: Schulinspektion soll insbesondere die *Schulentwicklung* befördern. Dazu vermittelt sie den inspizierten Schulen ein datengestütztes Bild ihres gegenwärtigen Entwicklungsstands, das auf verschiedensten Datenquellen und den Perspektiven aller relevanten Akteursgruppen fußt (Schulleitung, Lehrkräfte, Schülerinnen und Schüler, Eltern, Technisches und Verwaltungspersonal sowie Betriebe bei den berufsbildenden Schulen). Indem sie ein differenziertes Profil der Stärken und Schwächen der Schule erstellt, will die Schulinspektion sie darin unterstützen, Handlungsfelder zu identifizieren, Prioritäten zu setzen und einen Ausgangspunkt für die weitere Messung der Qualitätsentwicklung zu bestimmen.
2. *Staatliche Gewährleistung*: Mit der Sichtbarmachung der schulischen Qualität gegenüber den beteiligten Akteuren, der Schulöffentlichkeit und den verantwortlich Steuernden leistet die Schulinspektion einen Beitrag dazu, dass

Schulen *Rechenschaft* ablegen und einen Nachweis über die bisher geleistete Qualitätsentwicklung erbringen. Damit trägt die Schulinspektion dazu bei, dass schulische Qualität im Angesicht schulischer Selbststeuerung nicht dem Belieben der Einzelschule verfällt, sondern ein zwischen allen Schulen vergleichbarer Standard gesichert wird.

3. *Erkenntnis*: Indem sie die empirischen Befunde über alle Einzelschulen hinweg aggregiert und zu Aussagen über die *Qualität des Gesamtsystems* verdichtet, zeigt die Schulinspektion bereits vorhandene Stärken auf und identifiziert zugleich übergreifende Handlungsfelder der schulischen Qualitätsentwicklung. Damit erhält insbesondere die Steuerungsebene eine Rückmeldung über das bisher Erreichte ebenso wie über notwendige Steuerungsimpulse bzw. -korrekturen.

Landwehr (2001) ergänzt diese Trias um die Funktion der *Normendurchsetzung*. Diese Funktion wird durch die wahrgenommene Praxis der Schulinspektion bestätigt: Ohne dass es hierfür gesicherte empirische Belege gäbe, entspricht es doch der häufig berichteten Erfahrung von Inspektoren, dass Schulen sich insbesondere mit Blick auf die anstehende Inspektion intensiv mit den im Orientierungsrahmen kodierten normativen Erwartungen an eine gute Schule auseinandersetzen. Gärtner & Wurster (2010) gelangen zu dem Schluss, dass Veränderungen im Sinne von Qualitätsentwicklung insbesondere im Zeitraum *vor* dem Inspektionsbesuch erfolgen, was ebenfalls die Bedeutung der Normendurchsetzung durch die Schulinspektion unterstreicht.

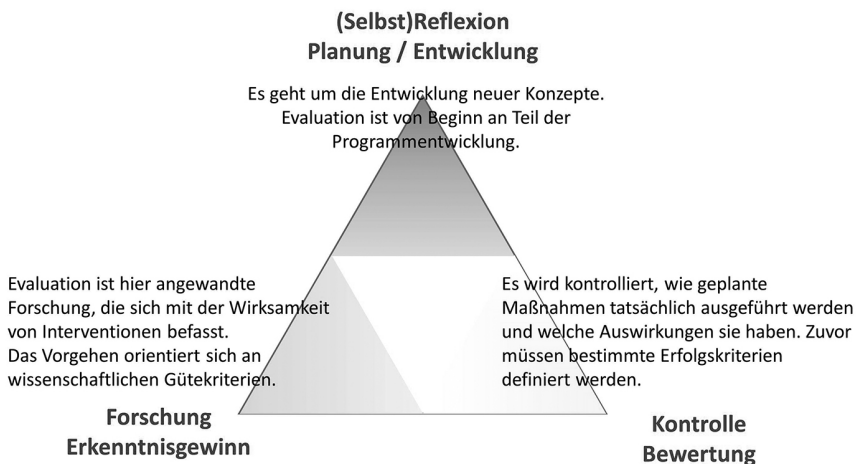
Kritisch ist festzuhalten, dass insbesondere die Frage, wie durch die Rückmeldung von Stärken und Schwächen sowohl auf der einzelschulischen Ebene als auch auf der Systemebene Veränderungen im Sinne der Qualitätsentwicklung und -verbesserung ausgelöst werden (Katalysefunktion), bislang nicht befriedigend geklärt ist. So ist es zwar unbenommen, dass die Schulinspektion mit ihrer umfangreichen Datengenerierung und -auswertung eine umfassende Wissensbasis schafft. Empirische Studien zeigen jedoch ein sehr uneinheitliches Bild mit Blick auf die faktische Nutzung solcher Ergebnisse durch die Schulen (vgl. zum Beispiel Pietsch et al. 2014; von Ackeren et al. 2013). Im Vorgriff auf die Weiterentwicklung der Inspektion kann deshalb festgestellt werden, dass das Verhältnis zwischen dem durch die Inspektion zur Verfügung gestellten Wissen einerseits und den Möglichkeiten der einzelschulischen wie auch der systemischen Schulentwicklung andererseits noch genauer auszutariieren ist.

1.1 Schulinspektion als Evaluation

Versucht man eine programmtheoretische Einordnung der Schulinspektion, so findet man am ehesten Anleihen in der Systematisierung verschiedener Evaluations-

ansätze. In der Literatur finden sich drei Typen bzw. Paradigmen der Evaluation (vgl. Chelimsky 1997; Abs & Klieme 2005; siehe Abbildung 1): 1. Im *Entwicklungsparadigma* sind Evaluationen neuer Konzepte und Programme verortet; es geht um den Gedanken, dass die Selbstreflexion und -vergewisserung von Beginn an Bestandteil der Qualitätsentwicklung ist. Evaluation ist hier eng mit der Entstehungsgeschichte der jeweiligen Intervention verwoben und soll Hinweise auf Möglichkeiten zur Verbesserung bzw. Optimierung geben. 2. Im *Kontrollparadigma* geht es primär um die Kontrolle des tatsächlich Erreichten im Verhältnis zu zuvor bestimmten Erfolgskriterien. Insoweit ist Evaluation hier Teil der Qualitätssicherung. 3. Im *Forschungsparadigma* dienen Evaluationen primär dem Gewinn wissenschaftlich abgesicherter Erkenntnisse über die Wirksamkeit und Wirkungsweise der untersuchten Intervention. Hier zielt Evaluation insbesondere auf die Identifizierung generalisierbarer Aussagen, die nach Möglichkeit Hinweise auf zugrunde liegende Wirkmechanismen oder Erfolgsbedingungen liefern.

Abbildung 1: Paradigmen der Evaluation nach Chelimsky (1997)



Schulinspektion trägt Züge aller drei unterschiedenen Paradigmen: Ihr primäres Interesse gilt dem Anstoß von Schulentwicklung und hebt damit auf die weitere Optimierung ab (Entwicklungsparadigma). Durch die Überprüfung der Schulen anhand zuvor definierter Qualitätskriterien ermöglicht sie nicht nur die Rechenschaftslegung der Schulen, sondern leistet auch einen Beitrag zur Kontrolle und Ausweisung, ob und inwieweit die Schulen ihrem Auftrag nachkommen (Kontrollparadigma). Und durch die Formulierung übergreifender Erkenntnisse auf Systemebene verweist sie auf über die Einzelschule hinausgehende Problemlagen, die

auf der Grundlage wissenschaftlicher Analysemethoden identifiziert und anhand wissenschaftlicher Gütekriterien empirisch abgesichert werden (Forschungsparadigma).

Mit dieser Multiperspektivität gehen verschiedene latente wie auch offene Konflikte einher, die bislang nicht befriedigend aufgelöst werden konnten. Oder, um es direkter zu sagen: Schulinspektion steht in der Gefahr, verschiedensten Seiten und sich zum Teil widersprechenden Bedürfnissen gerecht werden zu sollen, ohne dass das Verhältnis der unterschiedlichen Ansprüche mit Blick auf Vorrangigkeit, mögliche Hierarchien oder denkbare Gewichtungen geklärt wäre. Am ehesten lassen sich diese Widersprüche an dem Spannungsfeld „Entwicklung vs. Kontrolle“ festmachen. Der Auftrag und Anspruch, gleichzeitig Entwicklung anzustoßen wie auch Kontrolle auszuüben, erzeugt notwendigerweise Ambiguitäten: Ersteres erleben beide Seiten – sowohl die Schulinspektion als auch die Schulen – als den Schulen zuträglich und sie unterstützend, Letzteres nehmen Schulen, wenn nicht als bedrohlich, so zumindest als distanziert und abgewandt wahr. Für Inspektorinnen und Inspektoren geht damit ein latenter Nähe-Distanz-Konflikt einher, der erhebliche Anforderungen an ihre professionelle Selbstreflexion stellt (s. u.). Mit Blick auf die Schulen entsteht die Gefahr, dass der erlebte Widerspruch sie skeptisch bis misstrauisch macht und der Impuls zur Schulentwicklung im Sande verläuft. Da nicht davon auszugehen ist, dass Aspekte der Rechenschaftslegung oder aber Ansprüche der Qualitätsentwicklung künftig aufgegeben werden, wird dieser Zielkonflikt die Schulinspektion auch weiterhin prägen.

1.2 Selbstverständnis

Insbesondere da es sich um einen neuen Akteur in einer bereits existierenden Akteurskonstellation handelte, war es für die Schulinspektion wichtig, ihre eigene Rolle frühzeitig zu reflektieren und ihr Selbstverständnis zu definieren. Folgende Aspekte haben ihr Handeln im ersten Zyklus maßgeblich bestimmt:

- Vergleichbarkeit und Nachvollziehbarkeit von Inspektionsergebnissen sind ein wesentlicher Anspruch an die eigene Arbeit. Deshalb kommt dem Orientierungsrahmen Schulqualität eine zentrale Bedeutung zu (s. o.), da er die grundlegenden Bewertungsmaßstäbe definiert.
- Schulinspektion sollte als komplementärer Blick von außen auf ein System verstanden werden, das selbst kritisch innehält und eine regelhafte Evaluation wichtiger Entwicklungsschritte betreibt. Insbesondere angesichts der Stärkung der Selbstverantwortung bauen Instrumente der externen Evaluation darauf, dass die Schulen ihre Qualitätssicherung und -entwicklung eigenständig vorantreiben.

- Um sicherzustellen, dass die Werturteile der Schulinspektion gerechtfertigt sind, stützt sie ihre Bewertungen stets auf mehr als nur eine Datenquelle und berücksichtigt nach Möglichkeit die Perspektiven sämtlicher beteiligter Akteursgruppen. Damit folgt sie dem Grundsatz der Triangulation (vgl. Flick 2008).
- Ein zentrales Moment im Hinblick auf die Wirkungsweise der Hamburger Schulinspektion im ersten Zyklus ist die Rollenklarheit gegenüber anderen systemischen Akteuren, insbesondere im Verhältnis zur Schulaufsicht und zum Landesinstitut für Lehrerbildung und Schulentwicklung. Schulinspektion ist in diesem Sinne stets als Ergänzung der bereits bestehenden systemischen Funktionen (Kontrolle – Aufsicht – Beratung – Unterstützung) zu verstehen, sie folgt dem Prinzip der Arbeitsteilung.
- Diesem Gedanken folgend, baut Schulinspektion auf eine gemeinsame Ergebnisverantwortung der verschiedenen Akteure. Will Schulinspektion Wirkung entfalten, so darf es nicht egal sein, was mit den Ergebnissen passiert. Schulinspektion wird nur dort als sinnvoll erlebt, wo sie verbindliche Folgen hat.
- Aufgabe der Schulinspektion ist es, die Qualität der Einzelschule sichtbar zu machen. Damit nimmt sie sogleich eine systemische Perspektive ein in der Weise, dass nicht einzelne Personen Gegenstand der Evaluation sind, sondern die Schule als System. Dort, wo einzelne Personen in den Blick geraten – insbesondere die Schulleitung –, geschieht dies ausschließlich im Hinblick auf ihre Funktion, die sie für das Gelingen der gesamtschulischen Qualität wahrnehmen.

2. Das Inspektionsverfahren

Die Schulinspektion hat in sehr kurzer Zeit nach ihrer Gründung ein hoch standardisiertes, kriteriengeleitetes und auf wissenschaftlichen Methoden basierendes Verfahren erarbeitet, das eine gleichbleibend hohe Qualität der einzelschulischen Inspektionen gewährleisten soll. Wesentlich dazu beigetragen hat auch, dass neben den Inspektorinnen und Inspektoren von Beginn an Mitarbeiterinnen und Mitarbeiter der Verwaltung sowie ein Team aus Wissenschaftlerinnen und Wissenschaftlern beteiligt waren. Nachfolgend werden die Bewertungsgrundlagen und der Ablauf einer Inspektion beschrieben.

2.1 Bewertungsgrundlagen

Mit dem Orientierungsrahmen Schulqualität sind die normativen Erwartungen an eine gute Schule umfassend beschrieben. Um auf dieser Grundlage eine Inspektion durchzuführen, bedarf es allerdings der Operationalisierung, insbesondere mit dem Ziel einer standardisierten Erhebung und Berichtslegung. Dazu gehört auch, dass

unter Berücksichtigung der praktischen Umsetzbarkeit eine Auswahl zu bewertender Qualitätsmerkmale vorgenommen wird. In einem iterativen Prozess wurde deshalb im Verlauf des ersten Inspektionszyklus ein Bewertungsraster entwickelt, das analog zur Struktur des Orientierungsrahmens Schulqualität Qualitätsdimensionen, Qualitätsbereiche und Qualitätsmerkmale unterschied. Die Operationalisierung erfolgte dann durch die Ausweisung von Indikatoren. Sie sollte gewährleisten, dass alle Schulen auf der Grundlage gleicher Maßstäbe beurteilt werden. Jeder Indikator wurde durch das Inspektionsteam, sofern entsprechende Informationen verfügbar waren, auf einer vierstufigen Bewertungsskala (1 = deutlich mehr Schwächen als Stärken, 2 = eher mehr Schwächen als Stärken, 3 = eher mehr Stärken als Schwächen, 4 = deutlich mehr Stärken als Schwächen) bewertet, um durch Aggregation zu einer Bewertung der Qualitätsbereiche zu gelangen. Tabelle 1 weist aus, welche Qualitätsbereiche die Inspektion im ersten Zyklus bewertet hat. Es zeigt sich auf der Ebene der Qualitätsbereiche eine weitgehende Deckung mit dem Orientierungsrahmen Schulqualität mit Ausnahme des Qualitätsbereichs Bildungslaufbahnen und Kompetenzen, der mangels verfügbarer Daten nicht betrachtet wurde.

Tabelle 1: Betrachtete Qualitätsbereiche der Inspektion im ersten Zyklus

Führung und Management
1.1 Führung wahrnehmen
1.2 Personal entwickeln
1.3 Finanz- und Sachmittel gezielt einsetzen
1.4 Profil entwickeln und Rechenschaft ablegen
Bildung und Erziehung
2.1 Das schuleigene Curriculum entwickeln
2.2 Unterrichten, Lernen, Erziehen
2.3 Organisatorische Rahmenbedingungen sichern
2.4 Leistungen beurteilen
2.5 Prozesse und Ergebnisse evaluieren
2.6 Förderkonzepte entwickeln
2.7 Beratungsangebote gestalten
2.8 Die Schulgemeinschaft beteiligen
Wirkungen und Ergebnisse
3.1 Zufriedenes Personal, zufriedene Schülerschaft, Eltern und Betriebe

Neben dem Bewertungsraster wurden ebenfalls Vorkehrungen zur Standardisierung der Inspektionsberichte getroffen: So wurde ein Berichtsraster entwickelt, das für alle Stufen der Bewertungsskala Textbausteine vorgegeben hat. Die Bausteine waren entsprechend den Besonderheiten der jeweiligen Schule anzupassen. Auch

diese Maßnahme zielte vor allem darauf ab, Vergleichbarkeit zwischen den unterschiedlichen Inspektionen herzustellen und zugleich verbindliche Qualitätsstandards zu sichern.

2.2 Ablauf einer Inspektion

In ihrem Vorgehen hat sich die Schulinspektion Hamburg stark von der niederländischen Inspektion inspirieren lassen (vgl. <http://www.onderwijsinspectie.nl/english>). Grundlegende Prinzipien sind vor allem Multiperspektivität durch die Einbeziehung aller Akteure und die Daten- und Methodentriangulation (vgl. Flick 2008). Bei der Generierung und Analyse unterschiedlichster Daten werden qualitative und quantitative Verfahren genutzt. Abbildung 2 beschreibt einen typischen Inspektionsablauf. Demnach lässt sich die Inspektion grob in drei bzw. vier Phasen einteilen (wobei bei der vierten Phase die Inspektion selbst nicht mehr unmittelbar beteiligt ist).

Abbildung 2: Prototypischer Ablauf einer Inspektion



1. Vorbereitung und Vorerhebung

In den ersten sieben Wochen erfolgt die Kontaktaufnahme mit der Schule und die Vorbereitung der Inspektion. Dazu gehören neben dem Einholen von allgemeinen Informationen zur Schule und zu den Schulbeteiligten insbesondere die Sichtung und Analyse relevanter schulischer Dokumente:

- Schulprogramm/Leitbild,
- Ziel- und Leistungsvereinbarungen zwischen Schule und Schulaufsicht,
- Fortbildungsplanung,
- schulinterne Curricula.

Darüber hinaus werden alle schulischen Akteursgruppen schriftlich befragt:

- Schulleitung,
- Lehrkräfte,
- Schülerinnen und Schüler,
- Eltern,
- bei berufsbildenden Schulen die Betriebe.

Auf dieser Grundlage wird der Schulbesuch vorbereitet, wozu die Bildung erster Hypothesen auf der bestehenden Datengrundlage ebenso wie die Identifikation vertiefend zu beobachtender Qualitätsaspekte gehören.

2. Schulbesuch

Der zweitägige (an den berufsbildenden Schulen dreitägige) Schulbesuch wird in der achten Woche von einem in der Regel dreiköpfigen Team durchgeführt: zwei hauptamtliche Inspektorinnen und Inspektoren, von denen eine bzw. einer die Teamleitung innehat, sowie eine sogenannte Schulformexpertin bzw. ein sogenannter Schulformexperte; diese sind Leitungsmitglieder derselben Schulform wie die jeweils inspizierte Schule. In den berufsbildenden Schulen kommt darüber hinaus eine Wirtschaftsvertreterin bzw. ein Wirtschaftsvertreter hinzu, die bzw. der vor allem die Perspektive der Betriebe in den Blick nimmt. Während der Schulbesuchstage werden Interviews mit Vertreterinnen und Vertretern aller Schulbeteiligten durchgeführt: Schulleitung, Lehrkräfte, Schülerinnen und Schüler, Eltern, Technisches und Verwaltungspersonal, Betriebe (bei berufsbildenden Schulen). Außerdem erfolgen Unterrichtsbesuche: An den kleineren Grundschulen werden 40 zwanzigminütige Unterrichtssequenzen eingesehen, ansonsten 80, an den berufsbildenden Schulen 100. Dazu werden die Inspektionsteams durch in der Unterrichtsbeobachtung ausgebildete Fachleitungen anderer Schulen ergänzt, um die erforderliche Menge an Unterrichtssequenzen in der verfügbaren Zeit zu erreichen.

3. Auswertung und Ergebnisrückmeldung

In der neunten und zehnten Woche erfolgt die Auswertung sämtlicher erhobener Daten und auf dieser Grundlage die Berichtslegung. Ausgangspunkt ist das gemeinsame Auswertungsgespräch des Inspektionsteams am Tag nach dem Schulbesuch. Hier werden die Ergebnisse der Vorerhebungen ebenso wie die des Schulbesuchs mit Blick auf das standardisierte Bewertungsraster (s. o.) analysiert und

zu Bewertungen verdichtet. Ziel ist, im Inspektionsteam zu einer konsensualen Bewertung der 13 Qualitätsbereiche auf einer vierstufigen Skala (s. o.) zu gelangen. Die Letztentscheidung über die Bewertung ist der Teamleiterin bzw. dem Teamleiter vorbehalten.

Der etwa 35-seitige Bericht enthält zu allen Qualitätsbereichen detaillierte Beschreibungen des Beobachteten und Begründungen der Bewertungen, darüber hinaus einen Anhang mit weiterführenden Angaben zu den Unterrichtsbeobachtungen und den Ergebnissen der schriftlichen Befragung sowie zum Verfahren. Die Ergebnisse werden zunächst im Gespräch mit der Schulleitung zurückgemeldet; anschließend erfolgt in der Regel eine PowerPoint-gestützte Präsentation gegenüber der Schulöffentlichkeit. Letztere ist ein zusätzliches Angebot der Schulinspektion, das von nahezu allen Schulen in Anspruch genommen wurde, wobei der Kreis der Teilnehmenden durchaus variierte.

4. Auswertung und Nachbearbeitung der Ergebnisse

Aus Sicht der Schulinspektion ist die Inspektion mit der Ergebnismeldung nahezu abgeschlossen. Die Schule erhält die Möglichkeit, sachliche Fehler im Bericht zu korrigieren, sodass binnen zwei Wochen nach der Präsentation die Endfassung des Berichts erstellt werden kann. Darüber hinaus bittet das Inspektionsteam die Schule um eine schriftliche Evaluation des Inspektionsprozesses, um eigenes Optimierungspotenzial zu erkennen.

Danach geht der Prozess in die Verantwortung der Schule und der Schulaufsicht über. Zunächst setzt sich die Schulleitung, ggf. unter Einbeziehung der Schulgremien, mit den Ergebnissen auseinander und bereitet das sog. *Response-Gespräch* mit der Schulaufsicht vor. Dieses erfolgt ca. drei Monate nach der Ergebnismeldung; dabei werden möglicher Handlungsbedarf identifiziert und Maßnahmen der Schulentwicklung definiert. Im Idealfall münden die Vereinbarungen in eine Ziel- und Leistungsvereinbarung. Nach etwa einem Jahr wird im Bilanzgespräch zwischen Schule und Schulaufsicht ausgewertet, ob die Schule seit der Inspektion eine entsprechende Schulentwicklung vorangetrieben und etwaige Schwächen bearbeitet hat.

3. Professionalität der Inspektorinnen und Inspektoren

Schulinspektion stellt an die Handelnden vielfältige und komplexe Anforderungen. Inwieweit es sich bei den Inspektorinnen und Inspektoren bereits um eine eigenständige Profession handelt, kann derzeit noch nicht abschließend beurteilt werden. In einer zusammenfassenden Darstellung nennt Böttcher (2009) vier wesentliche Merkmale für die Profession:

- eine besonders anspruchsvolle Ausbildung mit entsprechend anspruchsvollen Qualifikationen,
- das Arbeiten an echten Problemen individueller Personen und das Lösen von Problemen in komplexen Interaktionskontexten,
- die Erfüllung zentraler gesellschaftlicher Funktionen, und damit verbunden
- die Sozialisation in ein spezifisches Wertesystem.

Diesen Ansprüchen folgend, lassen sich für die Berufsgruppe der Schulinspektorinnen und Schulinspektoren notwendige Kompetenzen in folgenden Felder identifizieren:

Fachkompetenz

- Schulische Feldkenntnis: Inspektorinnen und Inspektoren bedürfen des erfahrungsgesättigten Wissens über schulische Abläufe und die den Schulen inhärenten Funktionslogiken. Dieses kann in unterschiedlicher Weise erworben werden (beispielsweise nicht zwingend aus der Lehrerrolle heraus), ist aber eine wichtige Voraussetzung, um Beobachtetes angemessen einordnen und bewerten zu können.
- Kenntnis über Qualitätssicherung und -entwicklung: Schulinspektion folgt in ihrem Verfahren der allgemeinen Logik von Qualitätsmanagementsystemen, indem sie sich auf verbindliche Qualitätsmaßstäbe stützt, die als Kriterien expliziert und standardisiert sind, indem sie ein Bündel von Daten generiert und analysiert und indem sie die Ergebnisse in differenzierter Form an das System zurückmeldet. Wissen und Erfahrung im Bereich des Qualitätsmanagements sind deshalb zentrale Bestandteile des „Handwerkszeugs“ von Inspektorinnen und Inspektoren.

Methodenkompetenz

- Empirisches Grundwissen: Schulinspektion ist ein vor allem erfahrungsgeleitetes Verfahren, sie stützt sich auf die Anwendung wissenschaftlicher Methoden der empirischen Sozialwissenschaft. Ein Grundbestand an Wissen über Empirie oder zumindest die Bereitschaft, sich auf sie einzulassen und ihrer Logik zu folgen, sind für das Inspektionshandeln deshalb unerlässlich.
- Organisationsfähigkeit: Schulinspektion setzt ein hoch organisiertes, systematisches Handeln voraus, insbesondere deshalb, weil je Schule eine Vielzahl von Informationen zu verarbeiten und von Handlungsschritten zu unternehmen sind und zudem mehrere überlappende Inspektionen gleichzeitig koordiniert werden müssen. Dies gelingt nur bei einem in hohem Maße stringenten und wohl organisierten Vorgehen.

Selbst- und Sozialkompetenz

- **Beziehungs- und Kommunikationskompetenz:** Ob eine Inspektion gelingt oder misslingt, ist oft weniger eine Frage der rein fachlichen Expertise der jeweiligen Beteiligten, sondern hängt eher davon ab, ob es ihnen gelingt, einen sinnstiftenden, wertschätzenden Prozess zu initiieren und durchzuführen. Schulinspektion ist somit vor allem eine Aufgabe der Gestaltung sozialer Beziehungen (vgl. auch Brüsemeister in diesem Band). Inspektorinnen und Inspektoren bedürfen deshalb der Fähigkeit, sich auf unterschiedliche Personen einzulassen, ihre Besonderheiten zu erkennen und zu berücksichtigen und ihre Kommunikation spezifisch auf sie auszurichten.

Reflexionsfähigkeit: Da die persönliche Beziehungsgestaltung eine so wichtige Bedeutung hat, ist die Fähigkeit, das eigene Handeln selbstkritisch zu hinterfragen und zu reflektieren, eine, wenn nicht sogar die zentrale Kompetenz von Inspektorinnen und Inspektoren. Dies gilt umso mehr, da sowohl das System Schule als auch die Schulinspektion selbst einem kontinuierlichen Wandlungsprozess unterworfen sind. Nur wer bereit ist, das eigene Tun zur Disposition zu stellen und ggf. auf neue Anforderungen abzustimmen, dem gelingt auf Dauer der Umgang mit einer solchen Dynamik.

Da es sich hiermit um ein vielschichtiges Anforderungsprofil handelt, das nur wenige Personen aufgrund ihrer Berufsbiografie bereits mitbringen, wurde mit Gründung der Inspektion auch ein entsprechendes Ausbildungscurriculum entwickelt. Es sieht die etwa viermonatige Ausbildung von Inspektorinnen und Inspektoren vor, die theoretische wie praktische Elemente beinhaltet und mit einer eigenständig durchgeführten Prüfungsinspektion endet. Sie ist modulartig aufgebaut und wird in Abhängigkeit von den Vorkenntnissen der betreffenden Person angepasst. Wesentliches Element ist die Betreuung durch eine als Schulinspektorin erfahrene Mentorin bzw. einen als Schulinspektor erfahrenen Mentor. Konkret beinhaltet die Ausbildung folgende Bausteine:

1. Qualifizierung

- a. **Externe Qualifizierung:** Hierbei geht es um grundlegende Themen und Fertigkeiten, die eine besondere Kompetenz der Referentinnen und Referenten und ein spezielles Setting erfordern. Im Einzelnen werden folgende Themen – auch in Abhängigkeit von der Vorerfahrung der Person – behandelt: Rolle als Schulinspektorin bzw. Schulinspektor, Interviewführung, Berichte schreiben, Präsentationstechniken, Stilberatung.
- b. **Interne Qualifizierung:** Sie ist auf den Kern der Inspektion bezogen, es geht um die umfassende Vermittlung aller einschlägigen Aspekte der Inspektion: Quali-

- tätsmanagement, Instrumente, normative, theoretische und empirische Grundlagen, Abläufe. Diese werden vor allem durch die Mentorin bzw. den Mentor vermittelt, aber auch durch weitere Kolleginnen und Kollegen aus dem Institut.
2. *Shadowing*: Parallel zu den theoretischen Modulen lernen die neuen Inspektoreninnen und Inspektoren das Verfahren praktisch im Shadowing kennen. Ziel ist die Aneignung des Verfahrens, der Bewertungsstandards und die Übernahme der Inspektorenrolle. Dazu nehmen die Auszubildenden an allen Schritten einer Inspektion teil und übernehmen nach und nach einzelne Bestandteile, bis sie alle Elemente einmal selbstständig durchgeführt haben.
 3. *Mentoring*: Mentorin bzw. Mentor und Mentee treffen sich während der Ausbildung regelmäßig, in der Regel einmal wöchentlich. Gegenstand ist die gemeinsame Reflexion der Tätigkeit, der Erfahrungen und Eindrücke in der Inspektion und vor Ort in den Schulen. Alle Fragen der Ausbildung und zur Tätigkeit können hier angesprochen werden.
 4. *Ausbildungsgespräche*: Neben der Mentorin bzw. dem Mentor steht der neuen Inspektorin bzw. dem neuen Inspektor in größeren Abständen die Leitung der Schulinspektion für begleitende Gespräche zur Verfügung. Themen sind dabei unter anderem der Verlauf der Ausbildung und notwendige Unterstützungsschritte.
 5. *Selbststudium*: Die neue Inspektorin, der neue Inspektor arbeitet sich in viele Standards der Inspektion und in Hintergrundwissen selbstständig ein. Zur Vor- und Nachbereitung der internen Qualifizierungsmodule können Themen für das Selbststudium vereinbart werden.

Noch während der Ausbildung übernimmt die bzw. der Auszubildende die Rolle der Teamleitung. Die zweite Teamleitung gilt als Prüfungsinspektion und wird von der Leitung begleitet. Damit gilt die Ausbildung als abgeschlossen.

Über die Ausbildung hinaus bedarf es allerdings der kontinuierlichen Weiterentwicklung der Professionalität als Schulinspektorin bzw. als Schulinspektor. Neben der fachlichen Weiterentwicklung des Verfahrens geht es dabei insbesondere um die Vergegenwärtigung sich wandelnder Anforderungen, die sich aus der Veränderung des schulischen Kontexts ergeben. Damit einher geht aber auch die Notwendigkeit, das eigene Selbstverständnis, Qualitätsstandards und Rollenbilder immer wieder neu zu reflektieren, um weiterhin in angemessener Form agieren zu können. Die Schulinspektion bearbeitet diese Themen in unterschiedlichen Formaten:

- Einmal monatlich kommt sie in *Teamsitzungen* zusammen, in denen neben dem Informationsaustausch insbesondere die fachliche Weiterentwicklung des Verfahrens zum Gegenstand gemacht wird. Hier zeigt sich der Vorteil des Stadtstaats gegenüber einem Flächenland: Zusammenkünfte und damit auch die gemeinsame Arbeit des gesamten Teams sind weitaus häufiger möglich, was auch der Verständigung über gemeinsame Standards zuträglich ist.

- Im sogenannten *Wochenmarkt* werden einmal monatlich einzelne Verfahrenselemente thematisiert, unterschiedliche Praktiken sichtbar gemacht und Mindeststandards für das verbindliche Handeln festgelegt.
- *Teamtage* dienen der vertiefenden inhaltlichen Bearbeitung ausgewählter Themen, die in den bestehenden Formaten nicht befriedigend untergebracht werden können, weil sie zu zeitaufwendig sind.
- In den *Präsenztagen* (den letzten drei Tagen der Sommerferien) trifft sich die Schulinspektion an einem Ort außerhalb Hamburgs, um dort losgelöst vom Alltag das neue Schuljahr vorzubereiten und größere Themenblöcke zu bearbeiten.
- Um schwierige Situationen und mögliche Rollenkonflikte zu reflektieren, besteht für alle Inspektorinnen und Inspektoren die Möglichkeit, etwa einmal im Quartal an einer *Teamsupervision* teilzunehmen.
- Daneben begleitet die *Leitung* die Inspektorinnen und Inspektoren individuell durch Hospitationen, Rückmeldegespräche und im Rahmen von jährlichen Mitarbeiter-Vorgesetzten-Gesprächen.

4. Qualitätsmanagement

In dem Maße, wie die Überprüfung und Attestierung von Qualität den Kern der Schulinspektion bilden, liegt es auf der Hand, dass auch die Schulinspektion sich selbst Verfahren zur Überprüfung und Sicherung der Qualität ihrer eigenen Arbeit gibt. Im Sinne einer Lernenden Organisation sind die kontinuierliche Selbstreflexion und Selbstvergewisserung ebenso wie die Überprüfung eigener Qualitätsstandards selbstverständlich. Deshalb hat die Schulinspektion nach Abschluss ihrer Gründungsphase ein internes Qualitätsmanagement aufgebaut, das den Rahmen setzt, damit die Mitarbeiterinnen und Mitarbeiter der Schulinspektion qualitativ hochwertig und damit sinnstiftend arbeiten können (vgl. Bernt & Diedrich 2014). Sie orientiert sich dabei an dem Schweizer Q2E-Modell (Qualität durch Evaluation und Entwicklung, vgl. Landwehr & Steiner 2003). Dieses basiert wesentlich auf einem *Qualitätsleitbild* (vgl. den Beitrag von Ulrich in diesem Band), das typischerweise den ersten Schritt der systematischen Ausrichtung aller qualitätssichernden Verfahren darstellt. Ähnlich wie der Orientierungsrahmen Schulqualität definiert, was als gute Schule zu verstehen ist, beschreibt das Qualitätsleitbild der Schulinspektion, was sie als eine gute Inspektion versteht. Einer ähnlichen Unterscheidung in Qualitätsdimensionen, -bereiche und -merkmale folgend, werden unterschiedlichste Aspekte festgelegt, deren Erfüllung im Rahmen regelmäßiger Evaluationen geprüft wird. Bei erkannten Änderungsnotwendigkeiten, die intern verursacht sein können, aber auch durch Veränderungen des Kontexts (s. o.), werden entsprechende Änderungsschritte ergriffen, erprobt und nach einem zuvor festgelegten Erprobungszeitraum erneut evaluiert. Die Schulinspektion folgt damit idealiter einem

Entwicklungskreislauf des „*Plan, Do, Check, Act*“, der als allgemeines Kennzeichen gelingender Qualitätsentwicklung gesehen werden kann.

Neben der regelmäßigen Evaluation gehört ein fest installiertes *Feedbacksystem* ebenfalls zu den Grundpfeilern des Qualitätsmanagements der Schulinspektion. Dazu ist vorgesehen, dass jeweils die zuständige Teamleiterin bzw. der zuständige Teamleiter nach Abschluss einer Inspektion von der zweiten Inspektorin bzw. dem zweiten Inspektor ein umfassendes Feedback nach vorgegebenem Schema erhält. Dieses dient der persönlichen Weiterentwicklung jeder und jedes Einzelnen. Darüber hinaus sind die Inspektorinnen und Inspektoren aufgefordert, Kolleginnen und Kollegen oder auch die Leitung zu einer Inspektionshospitation einzuladen, sofern sie einen persönlichen Entwicklungsschwerpunkt vertieft in den Blick nehmen wollen.

Mit Blick auf die Qualität der Inspektionsberichte sind das *Bewertungsraster* und das *Berichtsraster* entscheidende Instrumente zur Gewährleistung verbindlicher Standards (s.o.). Im Konfliktfall – das heißt, wenn es zu nicht klärbaren Meinungsverschiedenheiten über Bewertungen in einem Inspektionsteam kommt oder wenn eine Schulleitung Bewertungen eines Berichts in erheblichem Maße moniert – nimmt die bzw. der Qualitätsbeauftragte eine Prüfung der Datenlage und der Berichtslegung vor und formuliert eine Stellungnahme an die Teamleitung, die diese in ihrer Letztentscheidung berücksichtigt. Im ersten Durchlauf der Inspektion aller Hamburger Schulen trat dieser Konfliktfall allerdings in deutlich weniger als einem Prozent der Inspektionsverfahren auf.

Das Q2E-Modell setzt auf eine Verbindung interner und externer Evaluationsverfahren. Deshalb hat sich die Schulinspektion bereits recht bald nach der Etablierung ihres internen Qualitätsmanagements dafür entschieden, ergänzend zu den Schritten der internen Evaluation ein *Peer Review* durchzuführen: „Kritische Freunde“ haben auf der Grundlage eigener Erkundungen Rückmeldungen zu zentralen Qualitätsaspekten gegeben. Nach einem umfassenden Prozess, in dem vier Kolleginnen und Kollegen aus anderen Bundesländern unterschiedlichste Daten zusammengetragen und zahlreiche relevante Personen einbezogen haben, ist ein Bericht entstanden, auf dessen Grundlage die Inspektion ihre weitere Qualitätsentwicklung strukturieren und priorisieren konnte.

Neben diesen originär auf ihr Qualitätsmanagement bezogenen Aktivitäten hat die Schulinspektion in Verantwortung ihres Wissenschaftsteams eine umfassende *Zufriedenheitsstudie* durchgeführt (vgl. Pietsch 2011). Dazu wurden im Frühjahr 2011 alle Schulleitungen an Hamburger Schulen, die in den Jahren 2007 bis 2011 durch die Schulinspektion Hamburg evaluiert wurden, schriftlich befragt. 166 Schulleitungen (59,5 % aller 279 Befragten) nahmen teil und ermöglichten mit ihren kritischen Angaben eine Einschätzung zur Nutzung und Nützlichkeit der Schulinspektion Hamburg. Es wurde deutlich, dass mit Blick auf die oben dargestellten Evaluationsparadigmen die Schulinspektion vor allem in den Ent-

wicklungskontext gestellt wird, weniger wird sie als Kontrollinstrument zur Rechenschaftslegung gegenüber Dritten gesehen. Generell attestieren die Befragten der Arbeit der Schulinspektion eine hohe Qualität und Glaubwürdigkeit, was andere Befunde bestätigt (vgl. z.B. Gärtner et al. 2009). Inwiefern die Inspektionsergebnisse tatsächlich für die Schulentwicklung genutzt werden, hängt aber offenbar vor allem davon ab, ob Schulleitungen ihnen eine Relevanz zuschreiben: Die Chance, dass konkrete Maßnahmen aus den Inspektionsbefunden abgeleitet werden, steigt in dem Maße, wie sie als relevant betrachtet werden. Allerdings hat die Studie auch deutlich gemacht, dass die Nutzung von Inspektionsergebnissen durch die Schulen nur recht bedingt durch inspektionsseitige Faktoren beeinflusst wird: Weitere Faktoren, wie das innerschulische Veränderungssetting und extern bereitgestellte Unterstützungsmaßnahmen, scheinen ebenfalls eine wichtige Rolle dabei zu spielen, ob es gelingt, Schulentwicklungsprozesse in Folge der Schulinspektion zu stimulieren.

5. Arbeitsteilung im System

Wie oben beschrieben, setzt die Schulinspektion auf Arbeitsteilung und eine gemeinsame Verantwortung unterschiedlicher Akteure bei der Schulentwicklung. Ihren Schwerpunkt sieht die Schulinspektion bei der datengestützten Bestandsaufnahme bzw. Diagnose und in der Rückmeldung ihrer Einschätzungen an die Schule. Beratung vorzunehmen, Empfehlungen zu geben oder den anschließenden Prozess der Ergebnisverarbeitung zu begleiten, fällt nicht in ihren Zuständigkeitsbereich. Um dennoch erfolgreiche Qualitätsentwicklung in den Schulen zu ermöglichen, kommt einer klar definierten Rollen- und Arbeitsteilung mit den systemischen Partnern der Schulinspektion eine zentrale Bedeutung zu. Dies sind vor allem die Schulaufsicht und das Unterstützungs- und Beratungssystem des Landesinstituts für Lehrerbildung und Schulentwicklung. Die Schulaufsicht ist die Instanz, an die die Schulinspektion den Prozess der Ergebnisverarbeitung unmittelbar nach der Präsentation der Ergebnisse übergibt. Sie begleitet die Auseinandersetzung der Schule mit den Ergebnissen und sichert deren Verbindlichkeit. Konkret sind die folgenden Schritte zwischen Schulinspektion und Schulaufsicht vereinbart:

- Die Schulaufsicht nimmt verbindlich an der öffentlichen Ergebnispräsentation der Schule teil und erfährt so die Ergebnisse aus „erster Hand“.
- Sie erhält als einzige Funktionsträgerin neben der Schulleitung den Inspektionsbericht.
- Sie führt spätestens zwölf Wochen nach der Ergebnissrückmeldung ein sogenanntes Response-Gespräch mit der Schulleitung. Hier tauscht sie sich mit der Schulleitung über ihre Einschätzungen und Bewertungen der Inspektionsergeb-

nisse aus und schließt verbindliche Vereinbarungen über die abzuleitenden Entwicklungsziele und Maßnahmen. Nach Möglichkeit sollten diese in eine neue Ziel- und Leistungsvereinbarung münden.

- Darüber hinaus verweist sie auf Unterstützungsangebote des Landesinstituts, insbesondere die Beratungsangebote der Agentur für Schulberatung und etwaige Fortbildungsmöglichkeiten.
- Ein Jahr nach Abschluss der Schulinspektion führt die Schulaufsicht mit der Schule ein sogenanntes Bilanzgespräch, in dem die wichtigsten Entwicklungen und weiterer Veränderungsbedarf evaluiert und reflektiert werden.

Um diese Form der Arbeitsteilung nachhaltig zu sichern und die Verbindlichkeit der einzelnen Schritte zu erhöhen, stehen sowohl die Leitungs- als auch die Arbeitsebene beider Institutionen in einem regelmäßigen Austausch. Ebenso besteht zwischen Schulinspektion und Landesinstitut ein regelmäßiger Kontakt, in dem unter anderem die Entwicklungsimpulse aus dem Jahresbericht der Schulinspektion aufgegriffen werden. Im Rahmen der oben zitierten Wirkungsstudie (Pietsch 2011) wurde allerdings sichtbar, dass eine strikte Rollentrennung nicht unbedingt im Sinne der Schulen ist: Die Mehrheit der befragten Schulleitungen würde es demnach begrüßen, wenn sie eine Beratung durch die Inspektionsteams während oder am Ende der Inspektion erhielten und somit die in Hamburg bisher praktizierte Trennung der institutionellen Akteure in die drei Bereiche Diagnose, Beratung und Begleitung sowie Unterstützung zukünftig reduziert würde.

Eine besondere Maßnahme, die aus der Kooperation aller drei Einrichtungen hervorgegangen ist, stellte die Unterstützung für Schulen mit besonderem Handlungsbedarf dar, die im Jahr 2010 initiiert wurde. Sie entstand aus der gemeinsamen Wahrnehmung, dass ein kleiner Teil der Schulen eine so schwerwiegende Problemkonstellation aufwies, dass sie aus eigener Kraft keine gelingende Schulentwicklung bewirken konnten, sondern massiver Unterstützung von außen bedurften. Für die Definition als „Schule mit besonderem Handlungsbedarf“ wurde ein spezifisches Bedingungsgefüge identifiziert, wobei der aktuell erreichte Stand der Schulentwicklung zugrunde gelegt wurde – d. h. die sozialen und bildungsbezogenen Voraussetzungen der Schülerschaft wurden explizit außen vor gelassen. Die Schulinspektion meldete eine Schule als eine Schule mit besonderem Handlungsbedarf an die Schulaufsicht, wenn schwache Urteile (Bewertungsstufe 2 oder schlechter) in folgenden Bereichen des Orientierungsrahmens Schulqualität festgestellt wurden: 1.1 Führung wahrnehmen und 2.2 Unterrichten, Lernen, Erziehen sowie in drei weiteren Bereichen aus 1.2 Personal entwickeln, 2.1 Das schuleigene Curriculum gestalten, 2.4 Leistungen beurteilen und 2.6 Förderkonzepte entwickeln. Im ersten Zyklus waren 24 Schulen (das entspricht ca. 6% aller Schulen) von einem solchen Bedingungsgefüge betroffen. Erfolgte eine entsprechende Meldung an die zuständige Schulaufsicht, holte diese zunächst weitere datengestützte Informationen wie

zum Beispiel Ergebnisse der zentralen Leistungsmessungen über die Schule ein. Mit dem Wissen der Schule, aber ohne ihre Beteiligung wurde dann seitens des Leiters des Amts für Bildung eine Fallkonferenz einberufen, an der die zuständige Schulaufsicht, die Aufsichtsleitung, die bzw. der für die Schule zuständige Teamleiterin bzw. Teamleiter der Inspektion, die Leitung der Schulinspektion und das Landesinstitut beteiligt waren. Nach einem Austausch über die verschiedenen Einschätzungen zur Schule wurde eine auf den Einzelfall bezogene, flexible Intervention festgelegt und unter Verantwortung des Landesinstituts durchgeführt. Nach etwa eineinhalb Jahren (in Abhängigkeit von der jeweiligen Maßnahme) erfolgte die Nachinspektion der betroffenen Schule. Dabei wurde ausschließlich auf die in der Erstinspektion monierten Qualitätsbereiche geschaut, Unterrichtsbesuche fanden nicht statt. Alle Schulen, die ein weiteres Mal inspiziert wurden, sind positiv bewertet und aus der Maßnahme entlassen worden.

6. Ausblick

Der erste Zyklus der Schulinspektion wurde im Frühjahr 2013 abgeschlossen. Die vorangegangenen Ausführungen sollen deutlich machen, dass es in sehr kurzer Zeit gelungen ist, ein hoch standardisiertes, normenbasiertes und kriteriengeleitetes Verfahren zu entwickeln, mithilfe dessen alle Hamburger Schulen in staatlicher Trägerschaft extern evaluiert werden konnten und eine Rückmeldung zu ihren Stärken und Schwächen erhalten haben. Trotz anfänglicher Skepsis seitens unterschiedlicher schulischer Akteure und einer gegenüber der Schulinspektion nicht ausnahmslos ausgeprägten Willkommenskultur in den Schulen ist es geglückt, die Schulinspektion als Verfahren in der Hamburger Schullandschaft zu etablieren. Auch wenn Kritiker nach wie vor Zweifel an der Sinnhaftigkeit von Schulinspektion äußern (vgl. z. B. Akbaş 2013), scheint die Akzeptanz der Institution sowie des Verfahrens mittlerweile recht hoch. Dennoch sollte dies nicht darüber hinwegtäuschen, dass einige der skizzierten Spannungsfelder nach wie vor nicht abschließend gelöst sind und die Schulinspektion auch weiterhin in gewisser Weise mit der Quadratur des Kreises konfrontieren. Dies betrifft insbesondere drei Entwicklungslinien:

1. *Die Gleichzeitigkeit von Entwicklungs- und Kontrollfunktion:* Dieses Spannungsfeld wird durch die bildungspolitische Entscheidung, die Inspektionsberichte im zweiten Zyklus im Internet zu veröffentlichen (s. Beitrag „Ausblick: Zweiter Zyklus der Schulinspektion Hamburg“ in diesem Band), eher weiter zugespitzt. Schulinspektion rückt dadurch in der Wahrnehmung der Schulen stärker in den Kontext von Rechenschaftslegung und öffentlicher Kontrolle, sodass es nicht einfacher werden dürfte, als wahrhaftig in dem Anliegen wahrgenommen zu werden, Impulse für die innere Schulentwicklung zu geben.

2. *Die Nutzung und Nutzbarmachung von Daten:* Die Hamburger Schulen erhalten fortlaufend datengestützte Rückmeldungen unterschiedlichster Herkunft, beispielsweise die Ergebnisse der landesweiten flächendeckenden Leistungsmessungen („KERMIT“), die Ergebnisse der zentralen Abschlussprüfungen, Erhebungen zum Sozialindex, regionalbezogene Aufbereitungen („Schule im Überblick“) zu Merkmalen ihrer Schülerschaft. Die Schulinspektion wird – angesichts der Vielzahl zu bewältigender Aufgaben, die zunächst wenig mit datengestützter Schulentwicklung zu tun zu haben scheinen – künftig verstärkt darauf hinwirken müssen, dass ihre Rückmeldungen an die schulische Arbeit noch anschlussfähiger als bislang sind.
3. *Die Gestaltung der systemischen Schnittstellen:* Auch wenn Verfahren der „Staffelübergabe“ zwischen Schulinspektion und insbesondere der Schulaufsicht definiert sind, ist angesichts des Fehlens eines gemeinsamen situativen Kontexts im Umgang mit Inspektionsergebnissen bislang nicht abschließend gewährleistet, dass eine konsistente Verwertung der Daten erfolgt: von der Generierung über die Rückmeldung an die Schule und die gemeinsame Deutung und Auswertung von Schule und Schulaufsicht bis hin zu etwaigen Maßnahmen am Landesinstitut. Insbesondere mit Blick auf den zuvor genannten Punkt 2 werden alle Akteure hier weiter an einem gemeinsamen Verständnis arbeiten müssen, damit Schulinspektion ihre qualitätsförderliche Wirkung vollständig entfalten kann.

Diese Entwicklungsaufgaben verstärkt aufzugreifen und anzugehen, prägt das Handeln der Schulinspektion im zweiten Zyklus, wie im Beitrag „Ausblick: Zweiter Zyklus der Schulinspektion Hamburg“ in diesem Band gezeigt wird.

Literatur

- Abs, H.J. & Klieme, E. (2005). Standards für schulbezogene Evaluation. In I. Gogolin, H.-H. Krüger, D. Lenzen & T. Rauschenbach (Hrsg.), *Standards und Standardisierungen in der Erziehungswissenschaft* (Zeitschrift für Erziehungswissenschaft: Beiheft 4, S. 45–62). Wiesbaden: VS Verlag für Sozialwissenschaften.
- van Ackeren, I., Binnewies, C., Clausen, M., Demski, D., Dormann, C., Koch, A. R., Laiser, B., Preisendörfer, P., Preuße, D., Rosenbusch, C., Schmidt, U., Stump, M. & Zlatkin-Troitschanskaia, O. (Autorengruppe EviS) (2013). Welche Wissensbestände nutzen Schulen im Kontext von Schulentwicklung? Theoretische Konzepte und erste deskriptive Befunde des EviS-Verbundprojektes im Überblick. In I. van Ackeren, M. Heinrich & F. Thiel (Hrsg.), *Evidenzbasierte Steuerung im Bildungssystem? Befunde aus dem BMBF-Förderschwerpunkt Steuerung im Bildungssystem (SteBis)* (Die Deutsche Schule: Beiheft 12, S. 51–73). Münster: Waxmann.

- Akbaş, M. (2013). *Warum fragt uns denn keiner? Was in der Schule falsch läuft*. München: C. Bertelsmann.
- Bernt, T. & Diedrich, M. (2014). Ist das QM oder kann das weg? – Qualitätssicherung in der Schulinspektion. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine daten- und theoriegestützte Schulentwicklung. Konzept und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)* (S. 51–65). Münster: Waxmann.
- Böttcher, W. (2009). Outputsteuerung durch Bildungsstandards. In H. Buchen & H.-G. Rolff (Hrsg.), *Professionswissen Schulleitung* (2. Aufl., S. 673–710). Weinheim: Beltz.
- Chelimsky, E. (1997). Thoughts for a new evaluation society. *Evaluation* 3 (1), 97–118.
- Flick, U. (2008). *Triangulation: Eine Einführung* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gärtner, H., Hüsemann, D. & Pant, H. A. (2009). Wirkungen von Schulinspektion aus Sicht betroffener Schulleitungen. Die Brandenburger Schulleiterbefragung. *Empirische Pädagogik*, 23 (1), 1–18.
- Gärtner, H. & Wurster, S. (2010). *Zweiter Ergebnisbericht: Längsschnittstudie zur Wirkung von Schulinspektion in Berlin*. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg e. V. (ISQ). Verfügbar unter: <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-347239> [Zugriff am 06.08.2014]
- Heid, H. (2000). Qualität. Überlegungen zur Begründung einer pädagogischen Beurteilungskategorie. In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich* (Zeitschrift für Pädagogik, 41. Beiheft, S. 41–51). Weinheim: Beltz.
- Landwehr, N. (2011). Thesen zur Wirkung und Wirksamkeit der externen Schulevaluation. In C. Quesel, V. Husfeldt, N. Landwehr & P. Steiner (Hrsg.), *Wirkungen und Wirksamkeit der externen Schulevaluation* (S. 35–70). Bern: hep-Verlag.
- Landwehr, N. & Steiner, P. (2003). *Q2E – Qualität durch Evaluation und Entwicklung. Konzepte, Verfahren und Instrumente zum Aufbau eines Qualitätsinstruments an Schulen. Heft 1: Das Q2E-Modell – Schritte zur Schulqualität. Aspekte eines ganzheitlichen Qualitätsmanagements an Schulen*. Bern: hep-Verlag.
- Maritzen, N. (2009). Schulinspektion und Schulaufsicht. In H. Buchen & H.-G. Rolff (Hrsg.), *Professionswissen Schulleitung* (2. Aufl., S. 1368–1392). Weinheim: Beltz.
- Pietsch, M. (2011). *Nutzung und Nützlichkeit der Schulinspektion Hamburg. Ergebnisse der Hamburger Schulleitungsbefragung*. Hamburg: Institut für Bildungsmonitoring. Verfügbar unter: <http://www.hamburg.de/contentblob/4022654/data/pdf-zufriedenheitsstudie-2011.pdf> [Zugriff am 12.08.2014]
- Pietsch, M., Jahnke, N., & Mohr, I. (2014). Führt Schulinspektion zu besseren Schülerleistungen? Difference-in-Differences-Studien zu Effekten der Schulinspektion Hamburg auf Lernzuwächse und Leistungstrends. *Zeitschrift für Pädagogik*, 60 (3), 446–470.
- Quesel, C., Husfeldt, V., Landwehr, N. & Steiner, P. (2011). *Wirkungen und Wirksamkeit der externen Schulevaluation*. Bern: hep-Verlag.

Erfahrungsbereiche der Hamburger Schulinspektion – Governance-analytische Programmtheorie

Thomas Brüsemeister

Zusammenfassung

In dem Artikel wird das Hamburger Institut für Bildungsmonitoring und Qualitätsentwicklung (IfBQ) mit Schwerpunkt auf die in ihm verankerte Schulinspektion untersucht. Orientiert an Analysen von Ehren et al. (2005) zur Schulinspektion in den Niederlanden wird die Programmtheorie der Hamburger Inspektion mithilfe der Governance-Perspektive analytisch erweitert. Dabei handelt es sich um vorläufige Befunde auf der Basis von acht offenen Interviews, die mit Expertinnen und Experten des IfBQ Anfang 2014 geführt und mit der Grounded Theory ausgewertet wurden, zunächst bezogen auf den Bereich der Bedingungen (Strategien und Konsequenzen werden ausgespart; zu den drei Bereichen vgl. Strauss & Corbin 1996, S. 132–139). Im Hauptteil des Artikels wird deutlich gemacht, dass das IfBQ mehrere vernetzende Aufgaben übernimmt: politische Verankerung, grenzüberschreitende Verständigung, Positionsklärung, Wissenser-schließung, Sichtbarmachung von Bildung. Diese Aufgaben werden mittels der Governance-Perspektive sowie nach Strauss (1978) als soziale Welten beschrieben und erklärt.

Einleitung

Die Programmtheorie der Hamburger Schulinspektion zu beschreiben, erscheint zunächst als gewinnbringendes Unterfangen, könnten doch, wie Ehren et al. (2005) für die Niederlande zeigten, Ungereimtheiten in einer derartigen Programmtheorie stecken. Einem derartigen Blick wird hier jedoch nicht gefolgt, denn wollte man wirklich eine Programmtheorie untersuchen, müsste dies, stützt man sich wie hier auf die Governance-Perspektive, noch viel weiter gefasst sein. Dies wird nachfolgend in einer Annäherung an eine später einmal auszuarbeitende Programmtheorie der Inspektion in Hamburg versucht, wobei mit Strauss (1978) mehrere konflikthafte Aushandlungsarenen sichtbar werden, in denen ein Institut wie das Institut für Bildungsmonitoring und Qualitätsentwicklung (IfBQ) tätig geworden ist. Nach dem Kodierparadigma von Strauss & Corbin (1996, S. 132–139), das Bedingungen, Strategien und Konsequenzen unterscheidet, begrenzt sich der Artikel – auf Basis einer kleinen Zahl von Interviews mit Akteuren des IfBQ – nur auf grundlegende

Bedingungen, beansprucht jedoch, darlegen zu können, dass eine Programmtheorie bereits in diesem Punkt mehrdimensional ist, da Erfahrungen in verschiedenen Bereichen gewonnen werden. Diese werden im Hauptteil kenntlich gemacht: als politische Verankerung des Monitorings, Verständigung und Zusammenarbeit, Positionsklärung, Wissenserschließung und Sichtbarmachung von Bildung. Insofern beanspruchen diese explorativen Befunde eine erste Annäherung an die Programmtheorie der Hamburger Schulinspektion.¹

1. Zur Programmtheorie

Ehren et al. untersuchten im Jahr 2005 die niederländische Inspektion dahin gehend, welche – oft impliziten – Unterstellungen die Gesetze und das Qualitätsmanagement des Inspektorats umfassten. Dies nannten sie „Programmtheorie“.² Für die Inspektion in den Niederlanden hielten sie z. B. als Annahmen fest (vgl. Ehren et al. 2005, S. 64): Würde das Inspektorat die Qualität einer Schule beurteilen, so würde die Schule mehr Variablen offerieren, die Aussagen über das Lernen von SchülerInnen erlauben. Weiter wird angenommen, dass sich dieses Offerieren intensiviert, je mehr eine Schule besucht wird. Das Gleiche würde passieren, würde das Inspektorat die Ergebnisse veröffentlichen. Ehren et al. stellen damit zahlreiche Annahmen über kausale Prozesse heraus, die die Akteure ursächlich mit der Schulinspektion verbinden.

Des Weiteren stellten sie im Programm der Inspektion Inkonsistenzen fest. So sei weder klar, wie Schulen die Feedbacks der Inspektion nutzen sollten noch welche Rolle die Inspektion in diesem Prozess hätte. Hier gäbe es den Widerspruch, dass auf der einen Seite das Programm die Schule in der Verantwortung für ihre eigene Entwicklung sehe. Auf der anderen Seite werde argumentiert, Schulen müssten durch die Inspektion angeleitet werden (ebd., S. 65). Gleichzeitig erwartet das Programm, dass Schulen Kapazitäten haben, eigene Fehler zu identifizieren und zu beheben (ebd., S. 66).

1 Ich bedanke mich bei Moritz Sowada für seine guten Hinweise zum Text sowie bei den InterviewpartnerInnen des IfBQ für die sehr freundliche Aufnahme und die konstruktiven Gespräche.

2 Eine Programmtheorie besteht nach Giel (2013, S. 118) aus verschiedenen Theorieversatzstücken, z.B. Alltagstheorien oder *Folk Theories*, Ad-hoc-Theorien, dem gesunden Menschenverstand und auch wissenschaftlichen Hypothesen. Theoretikerinnen und Theoretiker zur Programmtheorie sind sich uneins darüber, wie weit überhaupt *Ziele* als Bestandteile von Programmen angesehen werden dürfen; zumindest sieht Weiss (1998) Programme in derart schnellen Veränderungsprozessen, dass auch das Verfolgen einheitlicher Ziele kaum mehr unterstellt werden kann. Sie geht davon aus, „dass mehrere Theorien nebeneinander bestehen und dass diese durchaus falsch sein können“ (Giel 2013, S. 127) – eine Ansicht, der hier gefolgt wird.

Entsprechend untersucht eine Programmtheorie der Inspektion nach Ehren et al., wie konsistent, vollständig und wie realistisch die Annahmen sind, die der Inspektion unterlegt sind. Es sind Fehler in der Programmtheorie möglich, die gerade dann sichtbar werden, wenn die Inspektion nicht die erwarteten Effekte erzielt (ebd., S. 68–69). Damit verweisen sie auf mögliche Unklarheiten der niederländischen Inspektion, u. a. weil nicht genau geregelt sei, wie man sich autonome Schulen in Beziehung zu einer Deregulierung seitens der Regierung denke (ebd., S. 69). Ehren et al. sprechen hier von einer „*Ambiguity*“ (ebd.). Eine weitere Ambiguität wird im Referenzrahmen („*in the framework for inspection*“) gesehen, in dem Unvereinbares zusammengemischt sei. So müssten Schulen auf Gesetze reagieren, aber im Sinne eines Qualitätsmanagements sei ihnen die Art des Reagierens freigestellt (ebd.). Zusammengefasst fallen folgende Aspekte auf:

- Man kann sagen, dass die angenommenen Erwartungen für eine Programmtheorie zu speziell bzw. zu eng sind. Denn z. B. werden Erwartungen anderer Akteure, die an der Inspektion beteiligt sind oder die im weiteren Kontext mit ihr zu tun haben, überwiegend von der Forschung ausgeblendet (als Gegenbeispiel: Brüsemeister et al. 2014). Dies gilt insbesondere für die Akteure Politik und Schulaufsicht. Jedoch hat die Aufsicht mit dem Inspektionsbericht zu tun, insofern er die Grundlage von Zielvereinbarungen mit den Schulen sein soll.
- Auch lässt sich die Programmtheorie der Inspektion der Niederlande nur bedingt auf Deutschland übertragen, weil es weniger Elterneinflüsse gibt, dagegen in Deutschland der staatliche Kontext für die Inspektion eine viel größere Bedeutung hat.
- Für eine „echte“ Programmtheorie müsste nicht nur der „*Talk*“ (Brunsson 1989) der Programme, also das, was im „Programmheft“ steht, angeschaut werden, sondern auch die „*Action*“, also was in Interaktionen und Handlungen mit dem Programm unternommen wird³ – allein dies schließt aus, hier eine Programmtheorie der Hamburger Schulinspektion schreiben zu können, da dies eine eigene empirische Untersuchung erfordern würde. Jedoch warnen in diesem Zusammenhang Ehren et al. (2005, S. 72) davor, auf eine zu enge Programmtheorie zu setzen. Insbesondere würden mögliche Nebeneffekte ausgelassen, z. B. dass Schulen nicht immer ein Interesse daran haben, gerade während der Zeit der Visitation ihr Können zu ‚performen‘ bzw. zu zeigen (ebd., S. 71).

3 Auch Ehren et al. schreiben (ebd., S. 72): „Elte (1988, p. 2) also underlined the importance of looking at real (implemented) means and goals by saying that informal assumptions mostly influence the functioning and results of organizations instead of formal intentions.“

Es gibt also gute Gründe, eine Programmtheorie noch viel weiter zu fassen, als nur nach Interaktionseffekten des Programms mit den Schulen zu fragen, denn diese Interaktionseffekte könnten auch durch weiter gefasste Bedingungen beeinflusst sein, die in der bisherigen Analyse der Programmtheorie nach Ehren et al. zu kurz kamen.

Während die bisherige Programmtheorie nach Ehren et al. – in guten Absichten – eine Blickverengung vornimmt, indem sie Inkonsistenzen nur eines einzigen Programms untersuchen, kann mit der Governance-Perspektive genau umgekehrt eine Ausweitung des Blicks vorgenommen werden. So wird sichtbar werden, dass die Schulinspektion sogar noch viel voraussetzungsvoller ist, da die tragenden Institutionen – wie ein IfBQ – vernetzende Aufgaben übernehmen, d. h. ein Netz von verschiedenen Funktionen bedienen und bearbeiten.

Dies veranschaulicht folgende Tabelle, die als Grundlage von Kapitel 3 dort ausführlich expliziert wird, hier jedoch schon angeführt sei, allein um die Breite der Aufgaben zu verdeutlichen.

Tabelle 1: Erfahrungsbereiche der Schulinspektion in Hamburg

Erfahrungsbereich	Zentrale Elemente
Politische Verankerung des Monitorings	<ul style="list-style-type: none"> • Unterstützung durch Politik und Verwaltungsspitze • Bildung im Kontext der weiteren Ziele, Strategien und Leitbilder des Standorts • Priorität bildungspolitischer Ziele am Standort
Grenzüberschreitende Verständigung und Zusammenarbeit	<ul style="list-style-type: none"> • Kommunikation über Organisationsebenen und -grenzen hinweg • Zusammenarbeit unterschiedlicher Professionen • Vermittlung zwischen unterschiedlichen Selbstverständnissen und Traditionen
Positionsklärung	<ul style="list-style-type: none"> • Verständigung über Umfang und Möglichkeit der Beteiligung • Klärung von Rollen, Zuständigkeiten und Grenzen • Etablierung von Mechanismen für Konfliktbearbeitung und Interessenausgleich
Wissenserschließung	<ul style="list-style-type: none"> • Identifizierung und Erschließung von notwendigem Wissen • Verfügung von Wissen • Adressatengerechte Aufbereitung
Sichtbarmachung von Bildung	<ul style="list-style-type: none"> • Vermittlung von Zielen und Aktivitäten des Monitorings • Sichtbarkeit von Entwicklungen und Nutzen

Akteurs-, Daten-, Instrumenten-, Programmkonstellationen gemäß Governance-Perspektive

In der Governance-Perspektive dominiert die analytische Sicht auf *Akteurskonstellationen*, die dafür verantwortlich gemacht werden, wie sich Programme durchsetzen lassen, blockiert werden oder relativ wirkungslos sind, da Akteure nicht genügend involviert wurden (zu Akteurskonstellationen in der Schulinspektion vgl. Brüsemeister et al., in Vorbereitung; Brüsemeister et al. 2014).

Mit der Governance-Perspektive lässt sich weiter berücksichtigen, dass im Bildungssystem ja nicht nur ein, sondern mehrere „Datensätze“ gleichzeitig benutzt werden sollen. Neben Daten aus *Large Scale Assessments*, aus dem Monitoring und der Überprüfung durch Bildungsstandards erfolgt die Sicherung von Qualitätsstandards durch zentrale Prüfungen, Vergleichsarbeiten und Lernstandserhebungen (in Hamburg KERMIT) (vgl. Heinrich et al. 2014). Es herrscht also mit der Akteurskonstellation auch eine Daten- und eine Instrumentenkonstellation vor. Anders als bei dem vor vielen Jahren noch unterstellten Technologiedefizit im Bildungsbereich haben wir es nach Nittel (2011, S. 54) mittlerweile sogar mit einem „Technologieüberhang“ zu tun, erst recht für ein Institut wie das IfBQ, das verschiedene Dateninstrumente handhaben und aufeinander abstimmen muss.⁴

„Programmschachtelungen“

Gleichzeitig scheint das Nebeneinander vieler Einzelmaßnahmen Politik und Administration kaum bewusst, denn es werden immer wieder einzelne Programme dazu genommen und andere Programme werden schon nach einiger Zeit nicht mehr stringent weiterverfolgt, nachdem sie erst mit großer Aufmerksamkeit eingeführt wurden. Kurz: Es gibt Karrieren von Programmen verschiedenster Art, von denen die meisten bislang nur unzureichend orchestriert sind.

Nach der Governance-Perspektive lässt sich ein einzelnes Programm nicht isoliert betrachten, sondern ein Programm steht – aus Sicht der Akteure – in Beziehung, mitunter auch in Konkurrenz zu anderen Programmen – es gibt also auch eine Programmkonstellation – und in diesen anderen Programmen gibt es wiederum spezifische Akteurs-, Instrumenten- und Datenkonstellationen. Akteure verhalten sich hierbei nie nur zu einem einzigen Programm, sondern haben Erfahrungen

4 Der Terminus des Technologiedefizits – herzlichen Dank an Moritz Sowada für diesen Hinweis – bezieht sich eigentlich auf die pädagogische Praxis, an die insbesondere aggregierte Daten nicht heranreicht, jedoch umstellen diese Daten gleichsam massiv zumindest das Schülerverhalten und beeinflussen damit vermutlich auch einen Teil der pädagogischen Praxis.

mit vielen Programmen, die ihre Sichtweisen und Strategien zu einem speziellen Programm rahmen und beeinflussen.

Das Programm der Schulinspektion wird zum Beispiel beeinflusst durch das frühere Programm einer Schulaufsicht, die ohne Inspektion auskam, oder durch ein Monitoringprogramm oder noch durch ein allgemeineres Programm der Wissenschaftsorientierung oder noch allgemeiner durch das Programmzeitalter einer „*Audit Society*“ (Power 2000), die *Audits* in jedem gesellschaftlichen Bereich für sinnvoll hält. Bereits hier sind kulturelle Überzeugungen am Werk, die eine einzelne Programmtheorie beeinflussen. Zum Beispiel halten wir Akteure es in der *Audit Society* generell kulturell für besser, wenn eine Evaluation, ein *Audit* durchgeführt wird, als wenn dies nicht geschehen würde.

Die Akteure „schachteln“ ihre Erfahrungen aus vielen solcher Programme. Dies ermöglicht ihnen auch weiterzuarbeiten, sollte ein einzelnes Programm beendet werden. Und die Erfahrungen motivieren, längerfristige Entwicklungen im Blick zu haben und ein Programm als Baustein auf dem Weg dorthin zu instrumentalisieren. Zum Beispiel scheint Mecklenburg-Vorpommern derzeit die externe Evaluation zurückzufahren. Jedoch ist dies vielleicht nur vorübergehend, und die übergreifenden Überzeugungen, dass Evaluationen richtig sind, sind vermutlich nicht verschwunden. Andere Akteure haben Jahre gebraucht, um die Politik davon zu überzeugen, dass es mittel- und langfristig keine Alternativen zu einer datenbasierten Steuerung des Bildungswesens gibt. Einzelne kleine Programme, deren Funktionieren gezeigt wurde, waren der Einstieg in diese größere Sicht der Dinge, die national und international hoch relevant ist.

Wenn man die Art, dass und wie Akteure Programme schachteln, eine Kontextbetrachtung nennt, dann wird der empirischen Bildungsforschung ein anderer Blick eröffnet: weg von der Evaluation und Begleitung einzelner Programme hin zu Schachtelungen von Programmen und zu den diesbezüglichen Erfahrungen von Akteuren, die sich auch auf teilweise erfolgte Prozesse der Institutionalisierung beziehen.

Man könnte derart zum Hamburger IfBQ sagen: Die größeren, übergreifenden Programme, die vom IfBQ verfolgt werden, lauten „Wissenschafts-“ und „Daten- und Informationsorientierung“. Man benutzt diese übergreifenden Programme, klinkt sich in sie ein, und gleichzeitig nährt das IfBQ diese Programme, indem es die Schulinspektion systematisch aufbaut und verankert hat.

Hierbei besteht in Hamburg die Besonderheit, dass die Schulinspektion, lokale Bildungsberichte, KERMIT⁵, der Sozialindex, schulstatistische Analysen, der regionale Bildungsatlas u. a. m. in einem einzigen Institut gebündelt wurden. Mit diesem Bündel von Instrumenten werden dann verschiedene Teilfunktionen wahr-

5 KERMIT steht für „Kompetenzen ermitteln“. Das Programm ermittelt und unterstützt die Qualitätsentwicklung der Schulen; vgl. Lücken et al. 2014, S. 127 ff.

genommen, z. B. Dienstleistungen für die Politik⁶ und die beginnende Professionalisierung der Öffentlichkeitsarbeit (siehe dazu unten Kap. 3).

Dienstleistungsorientierung, Wissenschaftsorientierung, Daten- und Informationsorientierung etc. sind allgemeine Trends in verschiedenen gesellschaftlichen Teilsystemen, die sich Akteure zunutze machen, die sie aktivieren, wenn sie ein umgrenzteres, kleineres Programm wie die Schulinspektion initiieren. Man könnte hier von größeren Kulturprogrammen sprechen; mehr Dienstleistung, mehr Wissenschaft, mehr Öffentlichkeitsarbeit und vor allem: Mehr Daten sind allgemeine Trends, denen niemand widerspricht und die für die Legitimierung und Ausstaffierung einzelner Programme mitlaufen.

Akteure, die sich solcher allgemeineren Trends bewusst sind, verlieren ihr Herz nicht komplett an ein kleineres Programm wie die Schulinspektion, sondern sind in der Lage, dieses kleinere Programm auch aufzugeben, wenn der Geist und die Funktionen mittels anderer Programme erfüllt werden können. Ebenfalls sind sie in der Lage, ein Set von Instrumenten zu bündeln, wie es in Hamburg mit den verschiedenen „daten- und wissenschaftsorientierten“ Instrumenten geschieht.⁷

Es gilt also im Folgenden, sich die in Hamburg gestalteten „Programmschachtelungen“ anzuschauen, dies als Erweiterung des Blicks und anstelle einer nicht geschriebenen Programmtheorie der Inspektion. Dennoch bleibt das Interesse des vorliegenden Artikels darauf gerichtet, wie maßgebliche Akteure insbesondere die Schulinspektion innerhalb einer solchen „Programmschachtelung“ wahrnehmen.

2. Methode: Soziale Welten als „Erfahrungsbereiche“

In der Erforschung von Organisationen mittels qualitativer Methoden der Sozialforschung wurde bereits seit den 1960er Jahren mit dem Ansatz der „sozialen Welten“ nach Strauss gearbeitet (vgl. Glaser & Strauss 1968; Strauss 1984; Strauss & Maines 1991), der hier nur in Grundzügen wiedergegeben sei. Dieser Ansatz wird für geeignet gehalten, eine komplexe Organisation wie das IfBQ, in der das Programm der Schulinspektion eine wichtige Rolle spielt, in Ansätzen zu erfassen.

- Das Konzept der „sozialen Welten“ beschreibt zunächst, ähnlich wie in einer systemtheoretischen Kennzeichnung von Organisationen, eine Selektion, da

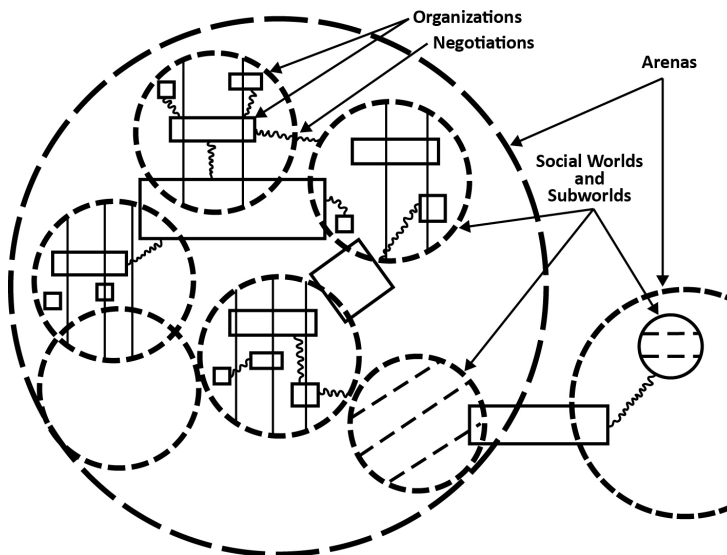
6 Auch Detlef Fickermann (2014, S. 296) spricht von wissenschaftlichen Dienstleistungen, die für die Politik erbracht werden, und beschreibt vermittelnde Funktionen, da es einerseits um die Verankerung von Monitoringabteilungen in den Wissenschaftsbetrieb geht, andererseits um Stellenstrukturen in der Verwaltung, sprich „personal- und haushaltsrechtliche Rahmenbedingungen“ (ebd., S. 297).

7 Eine solche Reflexionsfähigkeit ähnelt biografischen Handlungsmustern (nach Schütze) in der biografischen Forschung (vgl. Brüsemeister 2008, S. 118).

jede soziale Welt nach innen hin gemeinschaftsbildende Funktionen hat, während sie nach außen hin konflikthafte Arenen stiftet. „Soziale Welten stehen für den Zusammenhalt des Gleichgerichteten, Arenen für den Austausch des Divergenten.“ (Strübing 2007, S. 97) Jede soziale Welt differenziert eine Vielzahl solcher Konfliktarenen aus. Sie sind Bühnen für Aushandlungen mit anderen Akteuren (vgl. ebd., S. 59 f.; Niedlich & Brüsemeister 2011a), die wiederum dem Glauben ihrer eigenen Gemeinschaft folgen.

- Soziale Welten werden als Organisationen übergreifend verstanden. Wie man dem Modell von Clarke (1991, S. 123) entnehmen kann (vgl. Abbildung 1), bilden sich soziale Welten innerhalb und außerhalb von Organisationen, sie überlappen sich und sie bilden Subwelten aus:

Abbildung 1: Soziale Welten und Arenen



Quelle: Clarke 1991, S. 123

Entsprechend nützt es wenig, Funktionsweisen einer Organisation aus formalen Aufgaben und Rollen heraus verstehen zu wollen, vielmehr sind auf empirischem Wege die impliziten „Soziale-Welt“-Vorstellungen zu entschlüsseln. Die sozialen Welten lassen sich auch „Erfahrungsbereiche“ der Akteure nennen. Dieser Terminus ist modernisierungskritisch gemeint (vgl. Brüsemeister 2015, im Erscheinen), da die Akteure selbst gegen eine Verregelung durch hierarchisch strukturierte Verwaltungen und Professionen vorgehen und daraus teilweise durch Verwendung von Alltagssprache ausbrechen. Da qualitative Forschung

gegenstandsangemessen den Interpretationen von Akteuren nachgeht – nicht: nacherzählt –, sind ein Teil der Überlegungen der Akteure des Feldes auch in konzeptionelle wissenschaftliche Bezeichnungen eingegangen.

- In anderen Untersuchungen wurden für die neue Steuerung von Bildungssystemen ebenfalls gezeigt, dass multidimensionale Aufgaben bewältigt werden müssen. So verwies Clark (1997) für die Hochschulforschung darauf, dass das *New Public Management* gleichzeitig mehrere Stellschrauben verändern wollte, was dann in einem sogenannten Governance-Equalizer abgebildet wurde (vgl. de Boer et al. 2007; Schimank 2007). Wie bei einem Musikmischpult weist ein Equalizer mehrere Regler auf, die einzeln herauf- oder heruntergefahren werden können. Jeder einzelne Regler geht dabei mit Annahmen einher, ähnlich wie sie auch für eine Programmtheorie analysiert werden.⁸ Da die Kritik an den Reglern zudem war, dass sie zu speziell auf den Hochschulbereich zugeschnitten waren, wurden die Regler bewusst allgemeiner gehalten; eben dafür passt auch der umgangssprachliche Begriff von „Erfahrungsbereichen“. Der Clou der Regler wird sein, in späteren Projekten verschiedene Bildungsbereiche vergleichen zu können, z. B. mit folgenden Fragen: Wie intensiv sind die Erfahrungsbereiche jeweils in einem Bildungsbereich, um ein Programmset zum Laufen zu bringen? Wie stark sind Aufgaben und Einflüsse ausdifferenziert, wenn man sich einen einzigen Regler für verschiedene Bildungsbereiche anschaut?
- Die Erfahrungsbereiche beim aufzubauenden regionalen Bildungsmanagement (vgl. Lindner et al. 2015) wie auch die hier (in Ausschnitten) erfasste (und erweiterte) Programmtheorie zur Schulinspektion in Hamburg lassen sich mit Strauss (1984) als soziale Welten verstehen, die *Erfahrungen* von Akteuren in verschiedenen Arenen wiedergegeben, in denen um die Durchsetzung

8 Es würde hier zu weit führen, die einzelnen Regler für die Analyse von Hochschulen detailliert anzusprechen, da mittlerweile etwas veränderte Reglerdimensionen diskutiert werden, die sich auch auf andere Bildungsbereiche anwenden lassen (vgl. Niedlich & Brüsemeister 2011b). (Für Hochschulen soll immerhin so viel gesagt werden: Für sie lauteten die impliziten Wirkungsvorstellungen nach Clark vereinfacht: Eine Führung durch herausgehobene Einzelne ist möglich und bedeutet bessere Steuerung; der Rückzug des Staates auf Strategisches ist möglich und besser; Organisationen lassen sich durch externe Zielvorgaben bewegen; eine professionelle Selbstverwaltung bringt wenig und kann zurückgefahren werden usw.) Ebenfalls kann hier nicht darauf eingegangen werden, dass für die Erfahrungsbereiche (s. u.) eines kommunalen Bildungsmanagements und einen entsprechend analytisch breiteren Equalizer auch ein breiter gesellschaftlicher Modernisierungsdruck verantwortlich gemacht wird. Der bisherige analytische Regler aus dem Hochschulbereich folgte dagegen einem recht schmalen Band einer Modernisierung mittels Managementmethoden (New Public Management). Für die Analyse eines solchen Bildungsmanagements wurden die dezisionistisch von Clark gesetzten Regler durch empirische Regler ersetzt, die induktiv aus Interviewdaten und teilnehmenden Beobachtungen gewonnen wurden (vgl. Niedlich & Brüsemeister 2011b; Lindner et al. 2015).

eines Programms oder Programmsets gerungen wurde – in Auseinandersetzung mit Akteuren, die nicht derartige Überzeugungen hatten. Die maßgeblichen Akteure entwickeln eine entsprechende Soziale-Welt-Sicht, wenn man so will eine „*Theory of Change*“ (Weiss 1998), die Annahmen und Erfahrungen über Veränderungen aus den zurückliegenden sozialen Kämpfen enthält sowie zukünftige Erwartungen wiedergibt.

- Methodisch kann man soziale Welten mit der von Strauss entwickelten Grounded Theory herausfinden, in dem verschiedene Kategorien identifiziert werden (vgl. Strauss & Corbin 1996). Diese halten zentrale Strategien der Akteure fest sowie vorauslaufende Bedingungen für diese Strategien (außerdem werden Konsequenzen der Strategien festgehalten, z.B. für Nachhaltigkeit). Wichtig erscheint zunächst, die Bedingungen zu verdeutlichen, unter denen die Akteure arbeiten, unter denen sie soziale Welten mit den dazugehörigen konflikthaften Arenen aufbauen. Da dies ein längerfristiges Forschungsprogramm ist, werde ich nachfolgend einige zentrale Bedingungen vorstellen, die aber keinen Anspruch auf Vollständigkeit haben. Diese Bedingungen mögen etwas blass erscheinen, weil die eigentlichen Strategien (und die Konsequenzen) der Akteure hier nicht untersucht wurden; es wird also ungefähr ein Drittel einer möglichen (erweiterten) Programmtheorie behandelt.

3. Erfahrungsbereiche

Nun zu den die *Erfahrungen* von Akteuren in verschiedenen Arenen. Bezugspunkt meiner Annäherung an eine Programmtheorie sind damit soziale Welten, die ich mittels Governance-Perspektive als konflikttheoretische Auseinandersetzungen und Aushandlungen zwischen Akteuren verstehe. Um dies ansatzweise abbilden zu können, wird auf die Wahrnehmung zentraler Träger abgestellt – hier Akteure, die mit dem IfBQ zu tun haben und das Institut profilieren. Zugrunde liegen acht offene Interviews, die am 18. und 19.3.2014 erhoben und mit der Grounded Theory (vgl. Strauss & Corbin 1996) ausgewertet wurden. Kursiv gesetzt sind teils explizit geäußerte, teils implizite Theorien von ExpertInnen des IfBQ (und solchen, die mit ihm zusammenarbeiten), die Annahmen über Wirkungsweisen im Zusammenhang mit der Schulinspektion ausdrücken. Die Auswertung ist vorläufig und muss durch zusätzliche Kodierung weiter validiert werden.

3.1 Politische Verankerung

Die Politik davon zu überzeugen, dass eine einheitliche Adresse, *ein* Institut benötigt wird, das verschiedene Daten zu einem für die Schulen und die Politik sinnvol-

lem Set zusammenfügt: Dies ist sicher keine Aufgabe von einer Dekade, sondern sogar noch langfristiger. Da Wissen Differenzen erzeugt (und nicht abbaut), ist grundsätzlich eine gemeinschaftsbildende Sicht auf Datensysteme politisch nicht leicht herzustellen. Da sich Datenmöglichkeiten wie auch das soziale Verständnis von Daten beständig verändern, ist die Mobilisierung von politischer Unterstützung – um zu einer politischen Verankerung eines Monitorings zu gelangen – im Grunde eine permanente Aufgabe. So wie sich auch Präsidenten von Universitäten permanent profilieren müssen (vgl. Braun 2001), ist auch für ein Monitoringinstitut die Herausbildung von politischem Kapital, von politischer Sensibilität erforderlich. Dabei muss zugleich parteiübergreifend und unabhängig von Wahlen gearbeitet werden, da sich Datensysteme nur langfristig aufbauen lassen und vor allem der Umgang mit Daten und eine damit verbundene Bewältigung von Konflikten von allen Beteiligten erst erlernt werden muss.

Des Weiteren entsteht eine strukturell konflikthafte Arena dadurch, dass es nur eine „Chefin der Daten“ geben kann, was u. U. bestehende Ansprüche an politische Vorherrschaften infrage stellt. Jedoch lassen Daten in ihrer je eigenen Illusion, sie seien „neutral“, derartige Konflikte teilweise gar nicht aufkommen. Sprich: Die Organisation von Daten kann eine Zeit lang als neuer eigener Bereich wahrgenommen werden, als *Add-on*, der mit keinen bestehenden Verwaltungsabteilungen in Konflikt steht. Dass entsprechende Konflikte nicht geführt werden, könnte auch damit zu tun haben, dass die Politik sowohl ein neues Dateninstitut, als auch die bestehenden Einrichtungen politisch befürwortet.

Unterstützung der Politik, politische Planung im sozialen Wandel

Die Inspektion erscheint als eine Unterstützung für die Aufgabe der Politik, über das Schulsystem Rechenschaft abzulegen. Mit dem IfBQ ist der Hamburger Schulpolitik politische Planung möglich, wie sie datenbasiert mit den Schulen umgehen möchte.

Diese Planung hat sich in den letzten Jahren geändert, zuletzt 2013. Bedingt durch ein bürgerschaftliches Ersuchen werden die Inspektionsberichte nun veröffentlicht. Dafür wird ein Teil der Berichte einfacher gehalten und abgespalten. Soziologisch interessant erscheint an diesem Punkt, dass das Programm der Inspektion offen für sozialen Wandel ist. Das IfBQ reagiert auf veränderte Ansprüche der Politik.

3.2 Grenzüberschreitende Verständigung und Zusammenarbeit

Eine interviewte Person aus dem Institut zeichnet in ihrem Schaubild die eigene Einrichtung gleichsam als ‚Vernetzerin‘ zwischen verschiedenen anderen Instan-

zen ein. Es herrscht die Auffassung vor, das Institut sei ein nützlicher Agent für vernetzte Informationen.

Für eine derartige Zusammenarbeit mit anderen benötigt das Programm eine stabile, sich im selben Geist erneuernde Mannschaft mitsamt institutioneller Verankerung. In diesem Sinne ist die Etablierung des IfBQ zu sehen. Die Idee eines datengestützten Monitorings ist eine Leitidee nach Lepsius (2009), die sich mit institutionell verankerten Interessen verbinden muss, oder anders gesagt: Gibt es eine institutionelle Verankerung, ist auch das Interesse längerfristig verankert. Die Politik drückt aus, in dem sie sich ein IfBQ ‚leistet‘, dass es an dem Thema länger am Ball bleiben will. Eine herausragende Leistung des Programms, wenn man bedenkt, dass z. B. Mecklenburg-Vorpommern die Inspektion wieder zurückfährt.

Ein Mitglied des IfBQ legt dar, man hätte nur ein Institut wie das IfBQ für alle Bundesländer gebraucht. Das ist nicht im Sinne von Überheblichkeit gemeint, sondern drückt den enormen Aufwand bei der Datenbeschaffung und -anwendung aus, den man keinem zweiten Bundesland wünscht. Dass es eine Arena des Austausches und der Verständigung gibt, zeigt dabei, dass der Umgang mit Daten nach weiterer Vernetzung der Expertinnen und Experten ruft, der sogar über die Bundesländer hinausgeht. Die Institute sorgen selber für einen Austausch, vernetzen sich aus eigenem Antrieb über die Bundesländer hinweg, stehen offensichtlich kurz davor, einen eigenen Professions-Verband zu bilden. Es gibt einen bundesweiten, ja internationalen Diskurs zum Monitoring, der die Bemühungen wie in Hamburg unterstützt. Die Politik kann sehen: Was das IfBQ unternimmt, folgt einem internationalen Trend. *Das Vorhandensein eines übergreifenden Diskurses* ist ein wichtiger Faktor einer Institutionalisierung (vgl. Uske 2014), hier von Monitoringeinrichtungen. Umgekehrt nährt das IfBQ auch diesen Diskurs, um gegenüber der Politik auch die eigene Bedeutung zu unterstreichen, da gezeigt werden kann, dass auch andere Bundesländer einen ähnlichen Weg verfolgen.

Dass das Programm eine institutionelle Verankerung, sprich Stellen benötigt, ist eine weitere implizite Wirkungsannahme. In der Politik versteht jeder, dass ein ernsthaft verfolgtes Programm auch Stellen benötigt. Andersherum würden sich Programme unglaublich machen, würden sie keine Stellen erfordern.

3.3 Positionsklärung

Die Verankerung einer datenbasierten Steuerung in der Politik benötigt eine eindeutige Positionierung, was dem IfBQ gelingt, indem es seine Konzepte fachlich in der Wissenschaft verankert und sich dortigen Standards, u. a. Peer Reviews, unterzieht. Die Wissenschaft ist dabei Ausgangspunkt, gleichsam Heimat für die eigentlichen Strategien der Politikberatung und für Serviceleistungen der Schulen

(die Strategien dazu sind im vorliegenden Artikel weitgehend ausgespart). Diese Strategien sind selbst keine Wissenschaft, gehen aber von ihr aus.

In der übergreifenden Aufgabe, Schulen zu helfen, wird die Politik mit den Augen der Wissenschaft beraten und Wissenschaft mit den Augen der Politik gedeutet. Wie Kussau & Brüsemeister (2007, S. 76; Herv. i. O.) intermediäre Einrichtung beschreiben, lässt sich sagen, „dass die Kommunikationsarten von Politik und Wissenschaft aufeinander bezogen werden müssen. Ein Monitoring beinhaltet insofern, *Wissenschaft mit den Augen der Politik zu deuten* und *Politik mit den Augen der Wissenschaft zu beraten*. Es liefert der Politik wissenschaftsgestützte Instrumente. Ein Ministerium benötigt einen internen Stab von Fachleuten, der wissenschaftliche Argumente auf ihre politische Machbarkeit hin überprüft. Dies bedeutet, dass die Akteure im Rahmen von Monitoring gleichsam wie in einer ‚Reagenzglassituation‘ verschiedene ‚Stoffe‘ zu einer ‚Mixtur‘ zusammenbringen: auf der einen Seite wissenschaftliche Argumente, auf der anderen Seite eine Sichtung und Gewichtung dieser Argumente mit Blick auf Möglichkeiten der Politik.“

Aufbauarbeit

Einer der Befragten erläutert die Genese und die Aufbauarbeit des IfBQ. Es mussten bereits Jahre vor Arbeitsbeginn Stellen, die in der Schulaufsicht vakant wurden, für das neue Institut reserviert werden, um die notwendige Manpower zu haben. In dieser Zeit durften die abgebenden Instanzen nicht verprellt werden. Politische Unterstützung und Feingefühl bei gleichzeitig stringent verfolgten Absichten waren über Jahre hinweg notwendig. Im Hintergrund läuft ein entscheidender Gedanke der Nachhaltigkeit mit: Denn nur mit genügend Ressourcen kann man überhaupt in den Bereich von Wirkmächtigkeit gelangen.

Die Ausweitung des IfBQ ist geschichtlich bedingt, da Hamburg in den 1990er Jahren, vorangetrieben durch den damaligen Staatsrat Lange, mittels LAU (Studie „Aspekte der Lernausgangslage und der Lernentwicklung“), danach mit KESS („Kompetenzen und Einstellungen von Schülerinnen und Schülern“) und der Einführung eines Sozialindex einen wichtigen Beitrag zur „empirischen Wende“ in der Bildungsforschung leistete.

Meilenstein Institutionalisierung

Die ersten Jahre waren vom Aufbau geprägt und vom Geist beseelt, etwas völlig Neues zu schaffen. Aus Angst, dies alles könnte gefährdet sein – schließlich war und ist die Existenz von Monitoringabteilungen andernorts durchaus wacklig –, folgte aus dem Institut heraus nach einigen Jahren die Strategie der Absicherung

und Institutionalisierung. Und zwar indem die Schulinspektion als wichtiger Leistungsbaustein einem Qualitätsmanagement und Peer Review unterzogen wurde (wissenschaftliche Publikationen, Besuche von *Critical Friends*). Der neue Player will sich prospektiv absichern, er schätzt seine Position realistisch ein, er macht sich nichts vor. Aktuell verlangt die Politik dem Institut ‚Beiträge für die Zukunft‘ ab, auch im IfBQ müssen Stellen gekürzt werden (aktuell gibt es acht Schulinspektoren). Das interne Qualitätsmanagement ist angesichts dessen auch eine Absicherungsstrategie, die jedoch bereits früher eingeführt wurde.

Diese Absicherungsstrategie wird als ein erster wichtiger Meilenstein hervorgehoben. Ein zweiter Meilenstein war die Schulreform 2010, die das IfBQ zum Anlass nahm, bereits über den zweiten Zyklus der Inspektion nachzudenken. Auf Basis der gewonnenen Erfahrungen ließen sich notwendige Folgeaufgaben antizipieren, die das Institut bewerkstelligen konnte. Der dritte Meilenstein war, dass man diese Vorarbeiten nutzen konnte, als dann tatsächlich der zweite Zyklus kam.

Positionsklärung ohne Positionsklärung

So gut auch das IfBQ seine Position für sich gefunden hat und artikulieren kann und auch von anderen anerkannt wird, so dauerhaft brüchig bleibt die Beziehung zur Schulaufsicht. Der Grund dafür ist, dass sich die Schulaufsicht – in der Sichtweise des IfBQ – beständig in Rollenklärungsprozessen befindet. Überspitzt gesagt fällt damit dieser Akteur für eine dauerhaft verlässliche Kooperation aus. Das IfBQ hat dagegen eine relative Klarheit über sich und darüber, was es macht. Die zentrale Schnittstelle Inspektion/Schulaufsicht kann jedoch in der Akteurkonstellation nur bedingt bearbeitet werden.

So bleibt es auch in der Luft hängen, wer wann welche Konsequenzen aus den Daten zieht. Das IfBQ kann sie mangels Ressourcen und mangels Zuständigkeit nicht ziehen und die Schulaufsicht ist dafür nicht gebaut bzw. bislang dafür nicht umgebaut worden.

Als Ersatz versteckt man sich hinter Daten und äußert gegenüber den Schulen sinngemäß: ‚Die Konsequenzen aus den Daten müssen Sie selber ziehen.‘ Dem Institut gefällt diese eigene Position nicht, noch weniger, wenn sie auch beim zweiten Zyklus noch aufrechterhalten werden müsste. Die Alternative wäre, alle Mitspieler des Systems darauf zu verpflichten, den Schulen stärker dabei zu helfen, Konsequenzen aus den Daten zu ziehen. Dies steht unausgesprochen am Horizont.

Soziale Aushandlungen von Daten

Wie betont, vereint das IfBQ mehrere Aufgaben und Instrumente unter einem Dach: Systemmonitoring, lokale Bildungsberichte, den Sozialindex, den regionalen Bildungsatlas und KERMIT u. a. m. Damit können sich Schulen vielfach informieren. Zuerkannt wird, dass sie dies erst noch lernen müssen und dass man ihnen dabei helfen muss, dieses Lernen zu lernen. In einer Perspektivenverschränkung wird in Richtung der Schulen argumentiert: *Der Status der Urteile über Daten bewegt sich nicht auf der Ebene von wahr oder unwahr, richtig oder falsch, sondern die Schule muss selber prüfen, ob sie die hinter den Befunden stehenden normativen Urteile und die operativen Grundlagen übernimmt, sie muss einen sozialen Aushandlungsprozess betreiben.*

Schulen werden also als befähigt zu sozialen Aushandlungsprozessen gesehen. Ihnen wird damit etwas zugetraut, was ihrem Autonomiestatus entspricht. Zwar wird dieser bereits seit Jahren den Schulen zugeschrieben. Für das IfBQ ist nun jedoch die Frage konkreter, nämlich was es zu tun gedenkt, nach dem die Inspektion die Schulen besucht hat. Wie wird nun mit dem Inspektionsbericht umgegangen? Mit dem Verweis auf Aushandlungsprozesse wird dabei den Schulen der Hinweis auf eine soziale, nicht auf eine technische Seite einer internen Verarbeitung gegeben. Das Institut beweist sich darin als Grenzgängerin. Da sich die Schulen inspizieren lassen müssen, bestehen Interdependenzen zwischen dem IfBQ und den Schulen. Jedoch wird – im Zuge von sozialen Aushandlungen – das eher technische Instrument bzw. Verfahren umgedeutet in ein soziales Verfahren.

Verbesserung der Inspektion

Zugleich wird das Inspektionsverfahren teilweise als verbesserungsfähig gesehen – was die Reflexionsfähigkeit über die eigene Arbeit zeigt. Zum Beispiel werde dem Inspektionsteam nicht genügend Zeit gegeben, als Team der Schule ein paar Stunden lang am Ende der Inspektion Eindrücke zurückzumelden und ausführlich zu erklären, wie es zu seinen Urteilen gekommen ist. Dies ist eine Inkonsistenz der Programmtheorie. Zwar treten die Schulen bereits lange vor dem Inspektionsbesuch mit den Inspektoren in Kontakt, nach der Rückmeldesituation jedoch bricht die Kommunikation ab. Klar ist, dass für die Schulen keine Beratung gemacht werden kann, auch kein Coaching. Den Schulen würde es jedoch helfen, würden sie erfahren, *wie* genau das Inspektionsteam zu seinen Bewertungen gekommen ist. Dies könnte man Schulen summativ zurückmelden.

In einem Qualitätsmanagement würde man dies machen: Man würde direkt am Ort des Produktionsgeschehens über Fehler und Gelingen diskutieren. Wird jedoch ein Qualitätsmanagement innerhalb einer Hierarchie gemacht, besteht immer die

Möglichkeit, auf andere Akteure und Ebenen zu verweisen: Für die Entwicklung ist jetzt die Aufsicht zuständig, für Coaching die Schulberater, für weiteres Wissen die Fortbildner. Solche Verweise sind geradezu Aufgabe der Hierarchie, auch um weitere Akteure mit Arbeit zu versorgen. Nur die Schule soll sich außerhalb der Hierarchie denken und all die Dinge aufnehmen, die sie betreffen.

Das IfBQ ist in diesem Punkt erneut eine Grenzüberschreiterin: Umgeben von Hierarchien, soll es Schulen nichthierarchisches Denken beibringen. Offensichtlich wird die Funktion des Dritten, die die Inspektoren haben, nicht durchgehalten; die Inspektion hat diese Funktion des Dritten, da sie neben die Schule und die Aufsicht tritt. Der Staat lässt dies zu, aber dann nach der Rückmeldesituation doch wieder nicht genug. Die kommunikative Kompetenz des Teams als Dritten kommt nicht voll zum Tragen.

3.4 Wissenserschließung

Eine Person aus dem Institut erläutert die Wissenserschließung als Entwicklung weg von quantitativen Daten hin zu qualitativen. Dies stellt eine massive Veränderung dar.

Es geht dem IfBQ um eine Verbreiterung und Vernetzung des Wissens, zusätzlich zu den immer aktualisierten Leistungsdaten zum Unterricht. Hierbei soll das Wissen auch andere relevante Abteilungen erreichen. Das IfBQ vernetzt diese Akteure mit seinem Wissen, ist eine Art Dienstleisterin, ohne sich subordinieren zu wollen. Diese Vernetzung ist ein hohes Gut; es gilt, alle anderen Institutionen mitzunehmen.

Das ist eine ganz andere Vorstellung von Steuerung: Sie ist nicht hierarchisch, sondern wandert gleichsam kreuz und quer zu verschiedenen Akteuren eines Mehrebenensystems und der (politischen) Öffentlichkeit, dies in Rückkopplungsschleifen. Eine Vernetzung kann auf tagesaktuelle Politiken ebenso reagieren, wie eigene Ziele längerfristig im Auge behalten, und insbesondere auch die Wissenschaft als die zentrale Heimat.

Die Aufgabe der innerbehördlichen Verbreitung des Monitoringwissens ist groß, gilt es doch, andere Instanzen auf dem Laufenden zu halten und erneuerte Programnteile mit den anderen Akteuren zusammen einzubauen, d. h., den Wandel zu managen und gleichzeitig für die eigenen Perspektiven und Deutungsangebote zu kämpfen.

Das Wissen für alle beteiligten Behörden und für die Politik zu verbreitern, gelingt über die – stets auszubauende – Professionalisierung des IfBQ. Diese Professionalisierung aber darf nicht zu spezialisiert, nicht von den verschiedenen Behördenpraktiken und Lebenswelten der Öffentlichkeit abgehoben sein, muss von außen verstehbar bleiben, muss transferierbar bleiben und von anderen ge-

nutzt werden können, ohne wiederum abgeschliffen zu werden auf ein Jedermann-Wissen.

Wir-Gefühl und Commitment

Im Institut mussten verschiedene Generationen „persönliche Bewährungsproben oder auch Krisen in der Geschichte der Einrichtung bewältigen“ (Nittel 2011, S. 55). Sie identifizieren sich stark „mit dem ‚Geist‘ der Einrichtung“ (ebd.). Im IfBQ hat der intergenerationelle Austausch funktioniert, da Leitungspositionen auch mit Jüngeren besetzt wurden.

Die soziale Welt des IfBQ wird wesentlich von einer Schicksalsgemeinschaft bestimmt, weil man als eine der ersten Einrichtungen konsequent auf Datennutzung abgestellt hat, die ohne Alternative ist. Dies konnte verschiedenen Abnehmern deutlich gemacht werden. Das IfBQ dient damit auch der politischen Profilierung.

Der Preis ist, dass in den verschiedenen sozialen Welten dauernd die „Verzahnung von Begründungs-, Entscheidungs-, Realisierungs- und Evaluationsaktivitäten“ (ebd., S. 54) geleistet werden muss. Hierbei sind die Logiken teilweise schwer vereinbar, was aber nicht an den mangelnden Leistungen des Instituts liegt, sondern an teilweise schwer vereinbaren Anforderungen in der Umwelt.

Ein Beispiel hierzu ist der erneuerte Inspektionsbericht. Durch parteipolitische Bemühungen kam es 2013 zu einer Änderung des Schulgesetzes, wodurch der Inspektionsbericht in einen Berichtsteil für die Öffentlichkeit und einen für die Schulentwicklung differenziert wurde. Für den öffentlichen Teil gibt es Vorgaben, den Bericht knapp zu halten.

Da sich auch der Qualitätsrahmen verändert hat, haben Inspektoren Angst, dass ein Teil der Berichtspunkte nicht mehr auftaucht, an dem sich jedoch die Schulen im ersten Zyklus orientierten. Dies wird als Argument gegen die Änderung des Qualitätsrahmens angeführt. Obwohl die Schulen nach den gleichen Kriterien genau angeguckt werden, ist der öffentliche Berichtsteil schwammiger. Die Schulen wissen weniger, was sie damit anfangen sollen. Dies bemängeln die Inspektoren teilweise, was für ihre Professionalität spricht.

Es wird als schwierig empfunden, zwei Arten von Rückmeldungen zu geben: eine schwammige für die Öffentlichkeit und eine für die Schulen. Mit der ersten Art wird im Prinzip der Boulevard bedient; Schulen können an den Pranger gestellt werden, auch wenn dies faktisch noch nicht passiert ist. Entwicklungsarbeit mit den Schulen lässt sich jedoch nicht auf einem öffentlichen Marktplatz machen.⁹

9 Die Schulleitung kann einzig mit jeder einzelnen Lehrkraft Gespräche führen. Dafür muss der Staat einen Vertrauensschutz herstellen. Derartige Gedanken kamen vor einigen Jahren auch in Schweizer Kantonen auf, als die Macht deutlich wurde, die Monitoringdaten haben

3.5 Sichtbarmachung von Bildung

Nach der Auffassung von Nittel weist „das Verhältnis zwischen pädagogischem Sorgediskurs und den von der Gesellschaft erbrachten Anerkennungsleistungen [...] seit Jahrzehnten eine notorische Schieflage auf“ (Nittel 2011, S. 57). Anders gesagt besteht eine Lücke zwischen dem, was Bildungseinrichtungen für Leistungen erbringen, und dem, was die Öffentlichkeit davon wahrnimmt. Darauf reagieren neue Einrichtungen wie das IfBQ, in dem sie die öffentliche Informationslage pädagogisch Tätiger erheblich verbessern, und zwar auf einer wissenschaftlichen Basis. Damit wird auch dem medialen Auf und Ab eine langfristige, fundierte und abgesicherte Berichterstattung entgegengestellt.

Die Verbindungen zum Pressesprecher der Politik sind gut, was stark an Personen hängt und an deren Interessen, Bildungsfragen publik zu machen. Der letzte Jahresbericht wurde gut mit dem Pressesprecher zusammen in mehreren Sitzungen vorbereitet; ein Wissen, was offensichtlich *on the job* erlernt wird. Hier lässt sich ein Nachholbedarf erkennen, insofern Monitoringinstitute vielleicht eine eigene Pressestelle haben müssten.

Die Notwendigkeit dazu ergibt sich aus der im IfBQ geäußerten Ansicht, dass die *medialen Reaktionen auf die Produkte, die das IfBQ liefert, relativ unsteuerbar seien*. Während die Jahresberichte niemanden interessieren, gab es vor einiger Zeit eine Empörung in der Zeitung, als die gleichen Daten im Bildungsbericht auftauchen, sodass sich die Politik zum Einschreiten genötigt sah.

Hier mögen graduelle Differenzen zwischen Bildungspolitik und IfBQ zum Ausdruck gekommen sein, die mit den unterschiedlichen Funktionen dieser beiden Akteure zu tun haben könnten. Überblickt man jedoch insgesamt die Leistungen, die ein Institut wie das IfBQ hinsichtlich der Sichtbarmachung von Bildung unternimmt, wird die besondere Stellung solcher Institute umso deutlicher. Die erste Leistung besteht darin, dass Bildung nicht mehr nur als Selbstwert für sich steht, sondern dass ihr Wert über verschiedene Leistungsdaten – Schulinspektion, lokale Bildungsberichte, KERMIT, Sozialindex, schulstatistische Analysen, regionaler Bildungsatlas – öffentlich konkretisiert wird.

Damit wird zweitens auch die Legitimation der politisch Verantwortlichen indirekt unterstützt, denn die Bildungspolitik steht als Letztverantwortlicher in der Pflicht und ein Institut wie das IfBQ arbeitet der Politik in diesem Punkt zu. Durch die ‚Sichtbarmachung von Bildung‘ eröffnet das Institut der Politik ein zusätzliches Feld der Legitimation, das zu der bestehenden Legitimation der Politik

können. Dies wurde dann unter dem Begriff der *Ownership* gelöst (vgl. Bähr 2003). Dem Staat gehören aggregierte Daten auf Schulebene, um die Leistung des Gesamtsystems der Öffentlichkeit belegen zu können. Die Entwicklungsdaten (Klassendaten) gehören jedoch nur der Schule bzw. der Lehrkraft.

hinzukommt. Gleichzeitig trägt das Institut durch die Sichtbarmachung auch zur eigenen Legitimation bei.

4. Fazit und Ausblick

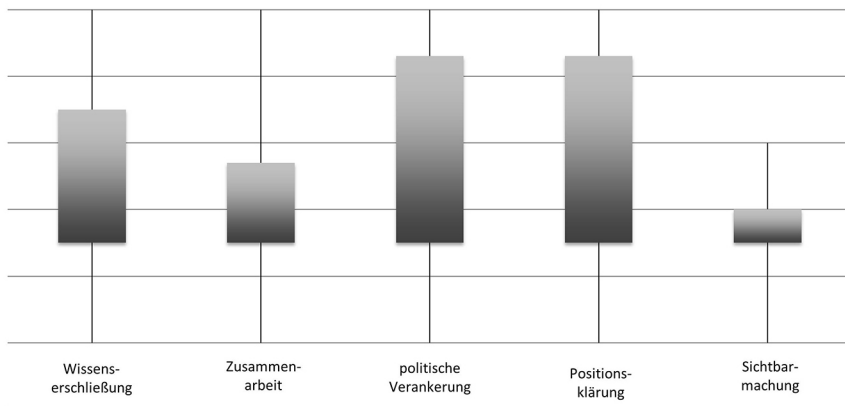
(1) Das IfBQ hat Erfahrungen in verschiedenen Bereichen gewonnen: politische Verankerung des Monitorings, Verständigung und Zusammenarbeit, Positionsklärung, Wissenserschließung, Sichtbarmachung von Bildung.

Die Erfahrungsbereiche weisen auf komplexe Handlungsfelder des Monitorings hin. Diese kann die Politik fördern, wodurch die Politik ihre Legitimität erhöht. Die Politik hat das IfBQ akzeptiert und nutzt insbesondere Inspektionsberichte, KERMIT und die Zielvereinbarungen. Probleme scheint eher die Schulaufsicht zu haben, die sich permanent in Rollenfindungen verhangen sieht, die zum Teil durch das klare Aufgabenbild des IfBQ mit ausgelöst scheinen.

Die Wissenschaftsorientierung des IfBQ ist ihre Heimat und verschafft der Institution ein eindeutiges Profil. Das IfBQ unterzieht die Schulinspektion einem Qualitätsmanagement – und könnte damit Vorreiterin für andere Behördenteile sein. Dienstleistungen gegenüber Schulen sind vielfältig und regen, nach dem Selbstverständnis des Instituts, Schulen im Umgang mit Daten zu sozialen Aus handlungsprozessen an.

(2) Bildet man die Erfahrungsbereiche als Regler ab (vgl. Abb. 2), sieht man, dass die Regler unterschiedlich weit hochgefahren sind. Daraus lassen sich Hypothesen ableiten, die weiter validiert werden müssen:

Abbildung 2: Erfahrungsbereiche in verschiedenen Handlungsbereichen



- Am umfangreichsten waren in der obigen Darstellung die Ausführungen zum Regler bzw. Erfahrungsbereich „Positionsklärung“. Auch die „politische Verankerung“ spielt eine große Rolle. Allein aus dem Umfang von Ausführungen lässt sich zwar nicht auf Wichtigkeit und Bedeutung schließen, da auch ein kleiner Aspekt eine enorme Bedeutung haben kann. Dennoch lässt sich die – weiter zu prüfende – These aufstellen, dass diese beiden Aspekte in Hamburg die größte Rolle spielen, erst dann folgen die „eigentliche Arbeit“ der Wissenserschließung, die Zusammenarbeit und, auf dem letzten Platz, die Sichtbarmachung von Bildung. Bei der „Positionsklärung“ und der „politischen Verankerung“ ist das IfBQ auch mit starken Interdependenzen anderer Akteure konfrontiert, wobei diese Interdependenzen kommunikativ bearbeitet werden müssen – während die Autonomie in der „Wissenserschließung“ deutlich höher ist und dem Institut dort niemand etwas vormacht.
- Dass Positionsklärungen den breitesten Raum einzunehmen scheinen, überrascht nicht, da Monitoring und Inspektion in der Bildungslandschaft neue Phänomene sind und ein dazugehöriges tragendes Institut auch.

Damit gäbe es erste Hinweise, dass die Programmtheorie allein schon bezüglich dieser fünf grundlegenden Handlungsbereiche bereits deutliche Profile aufweist, von denen geklärt werden müsste, ob sie alle so gewollt waren. Die Richtung der Antwort wurde bereits einleitend paradigmatisch gegeben, nämlich dass sich dies nicht rein auf angezielte Strategien zurückführen lässt, da ein Dickicht von Akteurs-, Daten-, Instrumenten- und Programmkonstellationen besteht, auf das zu unterschiedlichen Zeiten unterschiedlich reagiert werden muss.

Zu untersuchen wäre weiter, ob „Positionsklärungen“ und Aufgaben der „politischen Verankerung“ a) in späteren Jahren zurückgefahren werden können, da die Anfangsinvestitionen dann geleistet wurden und sich Erwartungen in der Akteurkonstellation eingespielt haben. Untersucht werden könnte b) weiter, ob (und wenn ja in welchem Ausmaß) „Positionsklärungen“ Programme, Personen und Ressourcen binden, die dann für die ‚eigentlichen Aufgaben‘ fehlen. In Hamburg jedenfalls scheint dies kaum der Fall gewesen zu sein, da die Positionsklärung erfolgreich war (und eher andere Akteure Probleme mit der Stringenz des neuen Akteurs andeuten). Analytischen Gewinn könnte auch die Analyse von These c) haben: Denn es könnte bezweifelt werden, dass „Positionsklärungen“ bloße Durchgangsstadien sind, sondern sie könnten umgekehrt neue Sinnhorizonte erschaffen, da den Akteuren durch solche Klärungen wechselseitig neue Dinge einfallen; zumindest ließen sich oben im Abschnitt zu den Positionsklärungen Hinweise dazu finden.

(3) Zum Schluss lässt sich fragen: Was nutzt es der Politik, wenn sie sich die vorhandenen Erfahrungsbereiche des IfBQ anschaut (vgl. Tab. 2)?

Tabelle 2: Erfahrungsbereiche des IfBQ und Nutzen für die Politik

Erfahrungsbereich	Nutzen für die Politik
Politische Verankerung des Monitorings	<ul style="list-style-type: none"> • Es lassen sich langfristige bildungspolitische Ziele verfolgen • Mehr Entscheidungen auf der Grundlage von Daten • Versachlichung • Weniger „Bauchentscheidungen“ • Stabilität der Bildungspolitik durch Datenunterstützung
Grenzüberschreitende Verständigung und Zusammenarbeit	<ul style="list-style-type: none"> • Vernetzende Kommunikation durch das IfBQ zu anderen Behördenteilen und der Öffentlichkeit via Daten • Kommunikation über Organisationsebenen und -grenzen hinweg • Zusammenarbeit unterschiedlicher Professionen • Vermittlung zwischen unterschiedlichen Selbstverständnissen und Traditionen
Positionsklärung	<ul style="list-style-type: none"> • Verständigung über Umfang und Möglichkeit einer Datennutzung • Klärung von Rollen, Zuständigkeiten und Grenzen • Künftig: Etablierung von Mechanismen für Konfliktbearbeitung und Interessenausgleich
Wissenserschließung	<ul style="list-style-type: none"> • Identifizierung und Erschließung von notwendigem Wissen • Verfügung von Wissen • Adressatengerechte Aufbereitung
Sichtbarmachung von Bildung	<ul style="list-style-type: none"> • Vermittlung von Zielen und Aktivitäten des Bildungsmonitorings und der Schulinspektion • Sichtbarkeit von Entwicklungen und Nutzen • Gewinnung von Reputation, Ansehen

Hervorzuheben ist, dass durch stabile Datengrundlagen eine Versachlichung des politischen Blicks auf Schulen möglich ist. Die Programmtheorie ist dabei für tagesaktuelle Ereignisse responsibel, wiewohl sie längerfristig Daten- und Instrumenten-Sets im Auge behält, diese erneuert und so insgesamt auf sozialen Wandel reagiert.

Es findet eine vernetzende Arbeit zu anderen Abteilungen und auch Bundesländern statt, die mit zu einer Stabilität eines übergreifenden Diskurses zur Notwendigkeit und Sinnhaftigkeit von Monitoringwissen beiträgt. Dies wiederum kann

sich positiv auf eine politische Verankerung auswirken, da die Politik sehen kann, dass nicht ein Bundesland allein an den entsprechenden Problemen und Problemlösungen einer datenbasierten Steuerung arbeitet.

Die Einführung der Inspektion und des Instituts als einer intermediären Einrichtung nötigt im Grunde allen Akteuren eine eigene Positionsklärung ab, ohne dass es dabei zu bemerkbaren Dysfunktionalitäten kommt. Im Gegenteil werden am Horizont mögliche Optionen für den weiteren Umbau des Schulsystems sichtbar.

Literatur

- Bähr, K. (2003). Die Rolle von Schulleistungstests für das Qualitätsmanagement im Bildungswesen, in Schulen und Klassenzimmern. In T. Brüsemeister & K.-D. Eubel (Hrsg.), *Zur Modernisierung der Schule. Leitideen – Konzepte – Akteure* (S. 217–224). Bielefeld: transcript.
- de Boer, H., Enders, J. & Schimank, U. (2007). On the Way Towards New Public Management? The Governance of University Systems in England, the Netherlands, Austria, and Germany. In D. Jansen (Hrsg.), *New Forms of Governance in Research Organizations. Disciplinary Approaches, Interfaces and Integration* (S. 137–152). Dordrecht: Springer.
- Braun, D. (2001). Regulierungsmodelle und Machtstrukturen an Universitäten. In E. Stöling & U. Schimank (Hrsg.), *Die Krise der Universitäten* (Leviathan Sonderheft 20, S. 243–262). Wiesbaden: Westdeutscher Verlag.
- Brunsson, N. (1989). *The Organization of Hypocrisy: Talk, Decisions and Actions in Organizations*. Chichester: Wiley.
- Brüsemeister, T. (2008). *Qualitative Forschung. Ein Überblick*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Brüsemeister, T. (2015, im Erscheinen). Verschiebungen von Machtbalancen im Bildungswesen. Kontrastiver Vergleich zwischen „Schulinspektion“ und „Lernen vor Ort“ mittels Governance-Reglern. In U. Steffens (Hrsg.), *Schulqualitätsdiskussion in Deutschland. Eine Zwischenbilanz anlässlich der 25. Fachtagung des Arbeitskreises Qualität von Schule*. Wiesbaden: Springer VS.
- Brüsemeister, T., Preuß, B. & Wissinger, J. (2014). Schulentwicklung als Governance. Herausforderungen datenbasierter Schulentwicklung. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine daten- und theoriegestützte Schulentwicklung. Konzeption und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)* (S. 215–234). Münster, New York: Waxmann.
- Brüsemeister, T., Gromala, L., Preuß, B. & Wissinger, J. (in Vorbereitung). Schulinspektion im regionalen und institutionellen Kontext. Qualitative Befunde zu schulinspektionsbezogenen Akteurkonstellationen. In Arbeitsgruppe Schulinspektionsforschung (Hrsg.), *Schulinspektion als Steuerungsimpuls?* Wiesbaden: Springer VS.

- Clark, B.R. (1997). *The Entrepreneurial University: Demand and Response*. Keynote. Speech presented at the 19th Annual EAIR Forum, University of Warwick (mimeo).
- Clarke, A.F. (1991). Social Worlds/Arenas Theory as Organizational Theory. In A. Strauss & D.R. Maines (Hrsg.), *Social organization and social process. Essays in honor of Anselm Strauss* (S. 119–158). New York: de Gruyter.
- Ehren, M.C.M., Leeuw, F.L. & Scheerens, J. (2005). On the Impact of the Dutch Educational Supervision Act. Analyzing Assumptions Concerning the Inspection of Primary Education. *American Journal of Evaluation*, 26 (1), 60–76.
- Fickermann, D. (2014). Einrichtungen zur Qualitätssicherung und -entwicklung als „nachgeordnete Dienststellen besonderer Art“. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine daten- und theoriegestützte Schulentwicklung. Konzept und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung* (S. 291–298). Münster, New York: Waxmann.
- Giel, S. (2013). *Theoriebasierte Evaluation. Konzepte und methodische Umsetzungen*. Münster, New York: Waxmann.
- Glaser, B.G. & Strauss, A.L. (1968). *Time for dying*. Chicago, IL: Aldine Publishing.
- Heinrich, M., Lambrecht, M., Böhm-Kasper, O., Brüsemeister, T. & Wissinger, J. (2014). Funktionen von Schulinspektion? Zum Governance-Programm der Vergewisserung und Weiterentwicklung der Qualität schulischer Arbeit. In C. Fischer (Hrsg.), *Damit Unterricht gelingt. Von der Qualitätsanalyse zur Qualitätsentwicklung* (S. 19–51). Münster, New York: Waxmann.
- Kussau, J. & Brüsemeister, T. (2007). *Governance, Schule und Politik. Zwischen Antagonismus und Kooperation*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lepsius, R.M. (2009). *Interessen, Ideen und Institutionen*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lindner, M., Niedlich, S., Klausning, J. & Brüsemeister, T. (2015). Regelungsbereiche des kommunalen Bildungsmanagements im Programm „Lernen vor Ort“ aus Sicht der Governance-Forschung. In T. Coelen, A.J. Heinrich & A. Uttke (Hrsg.), *Stadtbaustein Bildung*. Wiesbaden: Springer VS.
- Lücken, M., Thonke, F., Pohlmann, B., Hofmann, H., Golecki, R., Rosendahl, J., Benzing, M. & Poerschke, J. (2014). KERMIT – Kompetenzen ermitteln. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine daten- und theoriegestützte Schulentwicklung. Konzept und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung* (S. 127–154). Münster, New York: Waxmann.
- Niedlich, S. & Brüsemeister, T. (2011a). Bildungsmonitoring zwischen Berichterstattung und Steuerungsanspruch – Entwicklungslinien und akteurtheoretische Implikationen. In A. Wacker, U. Maier & J. Wissinger (Hrsg.), *Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung – Empirische Befunde und forschungsmethodische Implikationen* (S. 131–153). Wiesbaden: Springer VS.
- Niedlich, S. & Brüsemeister, T. (2011b). Modelle regionalen Bildungsmanagements – Ansätze zur Behebung sozialer und bildungsbezogener Bildungsungleichheiten. In N.

- Thieme, F. Dietrich & M. Heinrich (Hrsg.), *Neue Steuerung – alte Ungleichheiten?* (S. 201–218). Münster, New York: Waxmann.
- Nittel, D. (2011). Von der Profession zur sozialen Welt pädagogisch Tätiger? Vorarbeiten zu einer komparativ angelegten Empirie pädagogischer Arbeit. In W. Helsper & R. Tippelt (Hrsg.), *Pädagogische Professionalität* (Zeitschrift für Pädagogik, 57. Beiheft, S. 40–59). Weinheim, Basel: Beltz.
- Power, M. (2000). *The Audit Society. Rituals of Verification*. Oxford: University Press.
- Schimank, U. (2007): Die Governance-Perspektive: Analytisches Potenzial und anstehende konzeptionelle Fragen. In H. Altrichter, T. Brüsemeister & J. Wissinger (Hrsg.), *Educational Governance – Handlungskoordination und Steuerung im Bildungssystem* (S. 231–260). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Strauss, A. (1978). *Negotiations. Varieties, Contexts, Process, and Social Order*. San Francisco: Jossey-Bass.
- Strauss, A. (1984). Social Worlds and Their Segmentation Processes. In N.K. Denzin (Hrsg.), *Studies in symbolic interaction. A research annual* (S. 123–139). Greenwich, Conn: JAI Press.
- Strauss, A. & Corbin, J. (1996). *Grounded Theory. Grundlagen Qualitativer Sozialforschung*. Weinheim: Beltz.
- Strauss, A. & Maines, D.R. (Hrsg.) (1991). *Social organization and social process. Essays in honor of Anselm Strauss*. New York: de Gruyter.
- Strübing, J. (2007). *Anselm Strauss*. Konstanz: UVK.
- Uske, H. (2014). *Zwischenergebnisse des Transferprozesses im Rahmen des Projekts „Lernen vor Ort“ im Kreis Recklinghausen 2009–2014*. Universität Duisburg-Essen: unveröffentlichtes Manuskript.
- Weiss, C. (1998). *Evaluation*. Upper Saddle River: Prentice-Hall (zuerst 1972).

Das Qualitätsleitbild der Schulinspektion Hamburg

Susanne Ulrich

Zusammenfassung

Die Schulinspektion Hamburg hat im Jahr 2010 ein Qualitätsleitbild verabschiedet, das sie in einem partizipativen Prozess erarbeitet hat, der alle Mitarbeiterinnen und Mitarbeiter, die an der Schulinspektion beteiligt sind, einbezogen hat. Das Qualitätsleitbild ist integraler Bestandteil des Qualitätskonzepts der Schulinspektion Hamburg, das die bestehende Vielzahl von qualitätssichernden Aktivitäten konzentriert abbildet und systematisiert. Das Qualitätsleitbild formuliert den Anspruch der Inspektion an die Qualität der eigenen Arbeit und des eigenen Handelns und bildet damit die Grundlage für die Selbst- und Fremdevaluation. Das Qualitätsleitbild dient zum einen der Identitätsdefinition der noch recht jungen Organisation „Schulinspektion“, zum anderen erwächst aus dem Evaluationscharakter, der mit dem Leitbild verbunden ist, die Entwicklungsdynamik für das Inspektionsverfahren und die einzelnen an der Inspektion beteiligten Mitarbeiterinnen und Mitarbeiter.

Im Oktober 2006 gegründet, hatte die Schulinspektion die klare Zielsetzung, im Februar 2007 mit den ersten Schulbesuchen zu beginnen. In den dazwischen liegenden Monaten war die Inspektion mit der Entwicklung des hamburgspezifischen Inspektionsformats und der Qualifizierung ihrer Mitarbeiterinnen und Mitarbeiter für diese neue Aufgabe beschäftigt. Bereits in dieser frühen Phase wurde deutlich, dass parallel zur inhaltlichen Arbeit auch das eigene Rollenverständnis entwickelt werden musste. Als neue Institution hatte die Schulinspektion ihren Platz neben den etablierten Akteuren wie der Schulaufsicht und dem Landesinstitut zu finden. Auch die Notwendigkeit, ein gemeinsames Verständnis für das eigene Agieren zu entwickeln und damit einen Standard zu setzen, stand im Raum. Die Frage nach der eigenen Identität mündete schließlich in die Fragen: „Was ist eine gute Inspektorin bzw. ein guter Inspektor?“ und „Was ist eine gute Inspektion?“. In der Auseinandersetzung mit dem Thema kristallisierte sich heraus, dass es keine raschen Antworten geben konnte. Ziel war nicht, ein vorzeigbares Papier zu erstellen, sondern in einem partizipativen Verständigungs-, Reflexions- und Selbsteinschätzungsprozess, der alle Mitarbeiterinnen und Mitarbeiter der Schulinspektion einbezog, zu einer Identitätsdefinition zu gelangen, die dann in einem Qualitätsleitbild festgehalten werden sollte.

Qualitätskonzept und Qualitätsleitbild

Die Entstehung des Qualitätsleitbildes ist in engem Zusammenhang mit der des Qualitätskonzepts zu sehen. 2009 entwickelte die Inspektionsleitung gemeinsam mit dem Qualitätsbeauftragten ein Qualitätskonzept zur Etablierung eines umfassenden Qualitätsmanagements. Maßgeblich für diese Vereinbarung war der Gedanke, dass das Konzept alle qualitätssichernden Tätigkeiten systematisch erfassen soll, die die Qualität definieren und Nachhaltigkeit sichern. Als neu gegründete Institution hatte die Schulinspektion seit ihrer Gründung zahlreiche Qualitätssicherungsschritte wie den Workflow, das Drehbuch, Checklisten, Protokolle, Beschlusslisten usw. entwickelt. Hinzu kamen viele Einzelabsprachen zu qualitätssicherndem Arbeiten. Das Qualitätskonzept sollte nun die Vielzahl der qualitätssichernden Aktivitäten verorten, zueinander in Beziehung setzen und konzentriert abbilden. Mit dieser Systematisierung legt die Schulinspektion den Standard für ihr qualitäts- und entwicklungssicherndes Vorgehen fest, macht es nachvollziehbar und überprüfbar und schafft den Rahmen für ihre interne Reflexion. Über die inspektionsinterne Funktion hinaus dient das Qualitätskonzept der Rechenschaft nach außen und verfolgt damit auch eine externe Zielrichtung. Als Organisation, die die Qualität anderer Organisationen – Schulen – beurteilt und in deren Beurteilungstableau das Qualitätsmanagement enthalten ist, muss sich die Schulinspektion selbst diesen Ansprüchen stellen. In dem Qualitätskonzept manifestiert die Schulinspektion wie und nach welchen Prämissen sie ihre Qualität sichern und voranbringen will.

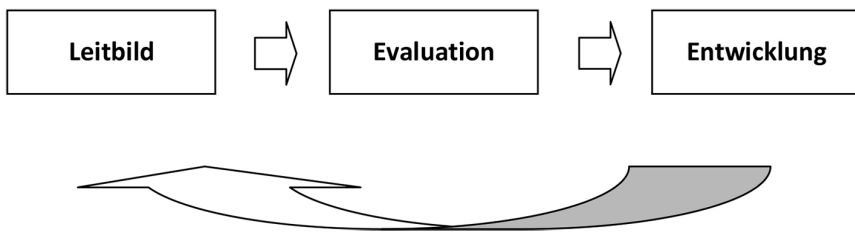
Das Qualitätskonzept ist in Anlehnung an das Q2E-Modell (Qualität durch Evaluation und Entwicklung; Landwehr & Steiner 2003) entwickelt worden, weil es einer entwicklungsorientierten Qualitätssteuerung, wie sie von der Schulinspektion intendiert ist, entgegenkommt. Das Q2E-Modell bietet der Schulinspektion eine Grundlage, an der sie ihre Qualitätssicherung und -entwicklung sowohl institutionell als auch individuell orientieren kann. Ausgangspunkt für diese Prozesse ist ein Qualitätsleitbild, das integraler Bestandteil des Qualitätskonzepts ist. Die im Leitbild niedergelegten Ansprüche werden durch Selbst- und Fremdevaluation überprüft und weiterentwickelt. Ziel ist dabei, den im Leitbild postulierten Ansprüchen durch kontinuierliche Überprüfung und Verbesserung in der Praxis schrittweise näher zu kommen. Im Qualitätsleitbild manifestiert die Schulinspektion den Qualitätsanspruch, dem sie gerecht werden will und an dem sie sich messen lassen muss. Das Leitbild wird damit zum Motor für die Qualitätsentwicklung des Inspektionsverfahrens wie auch für die Kompetenzentwicklung der einzelnen Beschäftigten, die in der Schulinspektion tätig sind.

Definition des Qualitätsleitbilds

Dem Gedanken folgend, dass die Qualitätsverbesserung eines Gegenstandes durch Evaluation initiiert wird, wurde für das Qualitätsleitbild der Schulinspektion festgelegt, dass es

- ein dynamisch wachsendes Dokument ist,
- aus evaluierbaren Aussagen besteht,
- der Identifizierung von Entwicklungsbedarfen dient,
- der (Selbst-)Verständigung über die schulinspektionsspezifischen Standards und ihrer Entwicklung dient und
- ein Instrument der systematischen Qualitätsentwicklung der Schulinspektion ist.

Abbildung 1: Entwicklung durch Evaluation



Quelle: Landwehr & Steiner 2003

Entstehung des Qualitätsleitbildes

Im Gegensatz zum Qualitätskonzept, das von der Inspektionsleitung und dem Qualitätsbeauftragten entworfen worden ist, war für die Entwicklung des Qualitätsleitbildes von Anfang an klar, dass die Erarbeitung im Gesamtteam stattfinden soll. Leitung und Qualitätsbeauftragtem kam in diesem Prozess nur steuernde oder zusammenfassende Funktion zu. Es ging darum, ein Selbstverständnis zu entwickeln, das konsensual von allen Mitarbeiterinnen und Mitarbeitern der Inspektion getragen wird, d. h. von dem Inspektorenteam, dem Wissenschaftsteam und dem Verwaltungsteam. Hintergrund war der Gedanke, dass die Arbeit der unterschiedlichen Berufsgruppen im Inspektionsverfahren eng miteinander verwoben ist und in diesem Rahmen das Wissen aller Spezialistinnen und Spezialisten zusammenkommen muss und nicht separiert werden darf. Dieser Anspruch legte ein mehrschrittiges Vorgehen nah, in dessen Zentrum gemeinsame Sitzungen standen, die

der Verständigung und der konkreten Aushandlung dienen. Den Start für die inspektionsinterne Diskussion des Leitbilds bildete im Sommer 2009 die Vorlage des Qualitätskonzepts. Richtungsweisend für die Auseinandersetzung waren die beiden Fragen „Was ist eine gute Inspektorin bzw. ein guter Inspektor?“ und „Was ist eine gute Inspektion?“. Im Klärungsversuch entfaltete sich schnell das immense Interpretationsspektrum, das die Notwendigkeit und Schwierigkeit einer Verständigung über einen gemeinsamen Qualitätsbegriff umso deutlicher machte. Im Vorfeld der entwicklungsorientierten Sitzungen wurden Zielaussagen, Aussagen zum Selbst- und Rollenverständnis, informell bestehende Qualitätsansprüche usw. gesammelt, um sie dort in konzentrierter Form zu diskutieren, zu präzisieren, zu verwerfen und weiterzuentwickeln. Durch die strukturierte Diskussion konnten die Ergebnisse konsensual gesichert und in überprüfbare Kriterien überführt werden. Der Prozess zielte auf die Erarbeitung eines gemeinsamen Qualitätsbegriffs, der in konkreten, überprüfbaren Soll-Aussagen zur inspektionsbezogenen Arbeit mündet. Entstanden ist ein Produkt, das Ansprüche formuliert, die mit der Praxis abgeglichen werden können und so den Anstoß zur weiteren Entwicklung geben können. Damit gibt sich die Schulinspektion den Orientierungsrahmen, an dem sie ihr eigenes Handeln und ihre Weiterentwicklung ausrichten kann.

Die Struktur des Qualitätsleitbildes

Die Struktur des Qualitätsleitbilds orientiert sich an drei Qualitätsdimensionen – dem Input, dem eigentlichen Prozess und dem Output. Die Inputqualität beschreibt vorrangig die Voraussetzungen, unter denen die Schulinspektion arbeitet. Konkret geht es um den Auftrag, den die Inspektion 2006 erhalten hat, die ein- oder mehrjährig abgeschlossenen Ziel- und Leistungsvereinbarungen mit der Behörde, die vorhandenen personellen und sachbezogenen Ressourcen, die Ausbildung der Mitarbeiterinnen und Mitarbeiter usw. Die Prozessqualität wird unterteilt in einen Bereich, der übergeordnete Prozessqualitäten beschreibt – wie die der Leitung, der Verwaltung oder des Qualitätsmanagements –, und einen Bereich, der Aussagen zum eigenen Kerngeschäft, zum Qualitätsanspruch des Inspektionsprozesses macht. Die Outputqualität meint bezogen auf die Schulinspektion vor allem die Wirksamkeit der Schulinspektionsarbeit für die Schulentwicklung. Im Entstehungsjahr des Leitbilds war die Wirksamkeit noch nicht überprüfbar. Insofern formuliert das Leitbild an dieser Stelle erste Soll-Vorstellungen. Die im Jahr 2015 anstehende Weiterentwicklung des Leitbilds (s. Ausblick: Die weitere Entwicklung) wird eine Präzisierung möglich machen.

Die drei Qualitätsdimensionen Input – Prozesse – Output bilden die oberste Gliederungsebene des Qualitätsleitbilds. Unterhalb dieser Ebene gibt es die Ebene der Qualitätsbereiche, sie differenzieren die Inhalte weiter aus. Die Soll-Aussagen,

die unter den Bereichen subsumiert sind, sind 2009 im diskursiven Austausch aller Mitarbeiterinnen und Mitarbeiter entstanden. Die Sammlung der Qualitätsbereiche und ihrer Qualitätskriterien spiegelt den Entwicklungsstand des Jahres 2009 wieder, anders als die Qualitätsdimensionen sind sie flexibel angelegt, sodass sie bei Überarbeitungen die aktuellen Qualitätsansprüche aufnehmen können.

Ausblick: Die weitere Entwicklung

Das Leitbild gibt es in seiner jetzigen Fassung seit dem Jahr 2010. Zu glauben, dass es in seiner Gesamtheit fortwährender Gesprächsanlass im Inspektionsalltag wäre, ist unrealistisch. Aber die einmal postulierten Ansprüche sorgen bei allen Inspektionsmitarbeiterinnen und -mitarbeitern für eine BewusstseinsEbene, an der sie sich orientieren. Nach fünf Jahren wird der Veränderungsbedarf aus unterschiedlichen Gründen deutlich. Zum einen hat sich die Inspektionsarbeit verändert, insbesondere die Entwicklung des zweiten Inspektionszyklus hat zahlreiche Veränderungen mit sich gebracht, die sich in der Neufassung des Orientierungsrahmens Schulqualität, der Bewertungsgrundlage, der Instrumente, des Berichtsformats und der Tatsache, dass die Zusammenfassung der Ergebnisse im zweiten Zyklus im Internet veröffentlicht wird, abbilden (vgl. Diedrich in diesem Band). Am deutlichsten und eventuell auch am auswirkungsreichsten ist jedoch die Weiterentwicklung der Rollenwahrnehmung der Inspektorinnen und Inspektoren. Der Inspektion ist bereits im ersten Inspektionszyklus bewusst gewesen, dass sie nicht die Verkünderin einer unabänderlichen Wahrheit ist. Im zweiten Inspektionszyklus hat die Inspektion dieses Bewusstsein mit dem Begriff des Deutungsangebots in Worte gefasst. Sie hebt damit hervor, dass sie den Schulen mit ihren Ergebnissen das Angebot unterbreitet, die Inspektionsergebnisse mit den eigenen Wahrnehmungen und Erkenntnissen zu vergleichen. Durch den Vergleich können die Schulbeteiligten herausfiltern, inwieweit sie der Darstellung des Inspektionsberichts folgen oder aber ihre eigene Wahrnehmung dagegen setzen. Dieser Auseinandersetzungsprozess setzt die Offenheit aller Beteiligten voraus und ist der erste Schritt, um die nächsten Schritte der Schulentwicklung zu identifizieren. Die Haltung der Schulinspektion, ein Deutungsangebot zu unterbreiten und keine unabänderliche Wahrheit zu formulieren, beruht auf dem Wissen, dass die Bewertungen auf der Grundlage zuvor definierter Konstrukte und Kriterien zustande kommen, die im Inspektionsprozess spezifische Wahrnehmungen und Interpretationen durchlaufen (vgl. Diedrich in diesem Band). Aus Inspektionssicht bietet die Rollenentwicklung das wichtige Potenzial, mit den Schulbeteiligten intensiv ins Gespräch zu kommen, und damit die Chance, die Wirksamkeit der Inspektionsergebnisse zu erhöhen. Die Wirksamkeitsfrage hatte die Inspektion im Jahr 2010 zwar bereits beschäftigt, im zweiten Zyklus ist sie aber richtungsweisend geworden und muss damit im Qualitätsleitbild einen hö-

heren Stellenwert einnehmen. Letztlich muss das neue Qualitätsleitbild auch der Organisationsentwicklung des Instituts für Bildungsmonitoring und Qualitätsentwicklung Rechnung tragen, in dessen Organisationsgefüge die Schulinspektion eine Unterabteilung bildet. Angesichts der umfassenden Veränderungserfordernisse ist noch nicht zu ahnen, wie das neue Qualitätsleitbild inhaltlich aussehen wird. Fest steht aber, dass die Entwicklung wieder in einem partizipativen Prozess unter Beteiligung aller Mitarbeiterinnen und Mitarbeiter der Schulinspektion stattfinden wird: Wenn es in einer initiiierenden Sitzung in der zweiten Hälfte des Schuljahres 2014/15 erneut heißt „Was ist eine gute Inspektorin?“, „Was ist ein guter Inspektor?“ und „Was ist eine gute Inspektion?“.

Das Qualitätsleitbild der Schulinspektion Hamburg

Vorweg

Mit folgendem Qualitätsleitbild legen wir als Schulinspektion die Qualitäten (Werte, Normen, Standards) fest, nach denen wir uns beurteilen lassen bzw. deren Erreichung wir künftig selber evaluieren werden. Damit steht das Qualitätsleitbild im Zentrum der Qualitätsentwicklung, indem es darlegt, welche „Soll-Vorgaben“ sich die Schulinspektion selbst gibt, nach denen Qualitätsentwicklung und -evaluation künftig erfolgen sollen. Die unten stehenden Qualitätskriterien beschreiben zum einen unsere Augenblicksdefinition von professionellem Handeln und zum anderen Ziele, die wir aus eigener Kraft heraus absehbar erreichen können und deren Erreichungsgrad wir im kommenden Jahr evaluieren wollen.

Der Begriff „Qualitätsleitbild“ steht bewusst in Abgrenzung zum herkömmlichen Begriff des Leitbildes – wie wir es aus der Schulleitbilddiskussion kennen. Das Schulleitbild dient dazu, das wünschenswerte, teilweise visionäre Profil einer Schule für die Darstellung nach außen (im Sinne der Corporate Identity) möglichst eingängig zu kommunizieren. Das Qualitätsleitbild hingegen wirkt nach innen wie beschrieben in einem direkten und unmittelbaren Bezug zum Qualitätsmanagement. Zur lebendigen Arbeit am und mit dem Leitbild gehört es, dieses regelmäßig zu evaluieren, aus den Ergebnissen der Evaluation konkrete Entwicklungsschritte für unsere Arbeit zu formulieren und Formate bereitzustellen, die einen weiteren Ausbau des Qualitätsleitbilds mit konsensualen Qualitätskriterien ermöglichen.

Zum Selbstverständnis unseres Auftrages und Handelns

Die Verbesserung der Lernchancen von Schülerinnen und Schülern in Hamburg ist das Ziel schulischer Qualitätsentwicklung und -sicherung, als deren Teil sich die

Schulinspektion versteht. Ihr sind wir in erster Linie verpflichtet. Unser Referenzrahmen für unseren Begriff von „guter Schule“ ist dabei der Orientierungsrahmen Schulqualität.

Externe Evaluation ist entwicklungsorientiert. Wir betrachten Schulen als sich verändernde und entwickelnde Systeme, die aktiv lernende Individuen voraussetzen. Diese Grundhaltung spiegelt sich auch in unserer eigenen professionellen Entwicklung.

Qualitätsdimension 1: Input-Qualität

Qualitätsbereich 1: Rahmenvorgaben/Strategische Vereinbarungen

Die Schulinspektion ist eine Abteilung des Instituts für Bildungsmonitoring (IfBM), welches seinerseits eine Dienststelle der Behörde für Schule und Berufsbildung (BSB) ist. Sie erhält von dieser ihren staatlichen Auftrag und ist in sie personal- und dienstrechtlich eingebunden. Die Schulinspektion ist in der Bewertung der Qualität einzelner Schulen an Weisungen nicht gebunden. Als Teil des Bildungsmonitorings fertigt sie systemübergreifende Berichte zum Hamburger Schulsystem.

Die Schulinspektion definiert ihre Aufgabe der externen Evaluation Hamburger Schulen im System der Entwicklung und Gewährleistung schulischer Qualität. Die Zuständigkeiten und die Form des Kontakts und der Zusammenarbeit im Systemdreieck Schulaufsicht – Landesinstitut – Schulinspektion sind klar beschrieben. Es gibt im Jahresturnus erneuerte Absprachen zur gegenseitigen Konsultation mit weiteren behördlichen und außerbehördlichen Stellen und Gremien im Bildungsbereich.

Die Schulinspektion inspiziert entsprechend der zur Verfügung gestellten Mittel die zugesagte Zahl staatlicher Hamburger Schulen.

Qualitätsbereich 2: Personelle und strukturelle Voraussetzungen

Die Stelle der Schulinspektorin/des Schulinspektors ist in Hamburg definiert. Der Definition liegen formulierte Standards zugrunde, die die Professionalität einer Schulinspektorin/eines Schulinspektors beschreiben.

Innerhalb der Schulinspektion gibt es unterschiedliche berufliche Funktionen: Abteilungsleitung und stellvertretende Abteilungsleitung, Schulinspektorinnen und -inspektoren, wissenschaftliche Mitarbeiterinnen und Mitarbeiter, Verwaltungskräfte und studentische Hilfskräfte. Alle tragen durch ihre Arbeit zum Gelingen der Schulinspektion bei.

Es gibt Stellenbeschreibungen für alle Funktionen der Schulinspektion.

Schulinspektorinnen und Schulinspektoren haben eine professionelle Ausbildung durchlaufen. Sie sind vom Ausbildungsträger zertifiziert und übernehmen Verantwortung für die permanente eigene Fortbildung. Regelmäßige und prozess-

begleitende Weiterbildungen und Rückmeldungen seitens des Teams Wissenschaft gewährleisten die Professionalität der Inspektorinnen und Inspektoren in besonderer Weise. In der Entwicklung unserer Professionalität arbeiten wir zudem länderübergreifend und im internationalen Austausch.

Die Leitung der Schulinspektion gewährleistet für alle Berufsgruppen Fortbildungen entsprechend den Anforderungen.

Die Mitarbeiterinnen und Mitarbeiter der Schulinspektion arbeiten bereichsübergreifend im steten Prozess an der Gewährleistung eines internen Common Sense zu zentralen Fragen des professionellen Verständnisses und der Praxis der externen Evaluation.

Externe Teammitglieder (Schulformexpertinnen/-experten, Wirtschaftsvertreterinnen und -vertreter, beteiligte Eltern) durchlaufen verbindliche Trainingsmodule, bevor sie eingesetzt werden. Die Schulinspektion greift dabei stark auf die eigenen Kompetenzen der Inspektorinnen und Inspektoren sowie des Teams Wissenschaft zurück.

Es gibt ein Rahmenmodell zur Arbeitszeit, in dem die Aufgaben und ihr zur Verfügung stehender zeitlicher Umfang beschrieben sind.

Qualitätsbereich 3: Materielle und finanzielle Ressourcen

Es ist für die Mitarbeiterinnen und Mitarbeiter durch Berichte und Veröffentlichungen der Leitung nachvollziehbar, welche finanziellen und personellen Ressourcen welchem Aufgabenpensum zugeordnet werden.

Es wird in einem angemessenen Umfang bezüglich Verpflegung und Verköstigung dafür gesorgt, dass sich Schulformexpertinnen und -experten, Gäste, Kolleginnen und Kollegen anderer Behörden u. a. bei Terminen, Arbeitssitzungen o. ä. willkommen geheißen fühlen.

Qualitätsdimension 2: Prozessqualität

Schulinspektion

Qualitätsbereich 1: Leitung

Die Leitung definiert die grundlegenden Entwicklungsziele und Anforderungen der Schulinspektion und formuliert klare Erwartungen an die Mitarbeiterinnen und Mitarbeiter. Sie bindet diese in die Gestaltung von Prozessen und in Entscheidungsfindungen mit ein.

Die Leitung gibt ihren Mitarbeiterinnen und Mitarbeitern in festgelegtem Turnus regelhaft Feedback. Sie ermöglicht dem Einzelnen unter anderem über die Mitarbeiter-Vorgesetzten-Gespräche individuelle Entwicklungsmöglichkeiten.

Die Leitung verfügt über ein Konfliktmanagement, das es ermöglicht, Konflikte zwischen Mitarbeiterinnen und Mitarbeitern und Leitung sowie Konflikte im Team zeitnah und lösungsorientiert zu bearbeiten.

Die Leitung stellt Transparenz über erteilte Arbeitsaufträge intern sowie über Ziele und Anforderungen her, die vom Auftraggeber oder vom Institut formuliert werden (Ziel- und Leistungsvereinbarungen etc.)

Die Leitung verfügt über eine längerfristige Strategie zur Positionierung der Schulinspektion in Hamburg und sorgt dafür, dass die Arbeit und die Linie der Schulinspektion angemessen nach außen kommuniziert werden.

Qualitätsbereich 2: Organisation/Verwaltung

Organisatorische Zuständigkeiten und Abläufe zwischen Inspektionsteams und Verwaltung sowie zwischen Inspektionsteams und dem Team Wissenschaft, innerhalb der Inspektionsteams wie auch innerhalb der Leitung sowie zwischen Leitung und Mitarbeiterinnen und Mitarbeitern sind geregelt und transparent.

Die Verwaltung unterstützt die Teamleitungen im Rahmen der fixierten Abläufe und Zuständigkeiten bei der Durchführung des Inspektionsverfahrens und arbeitet dabei selbstständig und eigenverantwortlich.

Qualitätsbereich 3: Kollegiale Zusammenarbeit

Die Zusammenarbeit aller Personen, die in der Schulinspektion arbeiten, ist von Zuverlässigkeit, Wertschätzung und Verbindlichkeit geprägt.

Die Zusammenarbeit aller in der Schulinspektion Beschäftigten basiert auf der gegenseitigen Anerkennung der Kompetenzen und des Engagements jeder/s Einzelnen. Meinungen und Ideen anderer werden respektiert und als Möglichkeit verstanden, gemeinsame Lösungen zu entwickeln.

Regelmäßiges und regelhaftes Feedback ist eine Konstante kollegialer Zusammenarbeit der Schulinspektion. Zur guten kollegialen Zusammenarbeit gehört, dass Meinungsunterschiede deutlich benannt werden, Kritik auch formuliert wird und die persönliche Integrität des Gegenübers gewahrt bleibt. Konflikte und offene Fragen werden zeitnah geklärt.

Fall- und Teamsupervision finden als regelhafte Struktur von Austausch und Klärung im regelmäßigen Turnus statt. Die Inspektorinnen und Inspektoren verstehen die Supervision als wichtiges und gewinnbringendes Instrument der kollegialen Zusammenarbeit und der professionellen Entwicklung.

Qualitätsbereich 4: Interne Innovationsqualität

Für die Mitarbeiterinnen und Mitarbeiter der Schulinspektion wie ihre Leitung stellt die Bereitschaft zur Innovation eine notwendige Qualität innerhalb eines sich entwickelnden Systems der externen Evaluation dar. Ressourcen für die inhaltliche Fortentwicklung der Schulinspektion werden von der Leitung bereitgestellt. Regel-

mäßige Fortbildungen werden angeboten und wahrgenommen; Zeiten dafür sind im Rahmenmodell zur Arbeitszeit verortet.

Die Heterogenität von beruflichen Hintergründen, Erfahrungen und Kompetenzen wird von den verschiedenen Akteuren als Möglichkeit eingesetzt und genutzt, um Qualität zu entwickeln und zu sichern. Zweifeln, Nachfragen, Nachdenken haben ihren Platz und werden als Chance gesehen.

Innovation bedingt die systematische Reflexion der standardisierten Abläufe: Regelmäßig kommen dabei kollegiales Feedback, Formen der Qualitätskontrolle innerhalb der Teams (Gegenlesen der Berichte und Redaktionssitzungen) und Supervisionen im Gesamtteam zum reflektierenden Austausch über Arbeits- und Teamprozesse zum Einsatz.

Das Team Wissenschaft verantwortet die Weiterentwicklung der Instrumente.

Die Schulinspektion beteiligt die Mitarbeiterinnen und Mitarbeiter an der Weiterentwicklung von Verfahren und Instrumenten und nutzt so ihre Professionalität und Erfahrung.

Innovationsqualität stellt die Schulinspektion her durch die aktive bundesweite Vernetzung und Evaluation der Länderinspektionen. Das Team Wissenschaft sichert durch seine Arbeit und durch Kooperationen mit anderen Instituten einen hohen wissenschaftlichen Standard.

Auf veränderte Rahmenbedingungen der Schulen reagiert die Schulinspektion in angemessener Weise.

Qualitätsbereich 5: Qualitätsmanagement

Das Qualitätsmanagement wird von der Leitung gesteuert. Im Qualitätskonzept der Schulinspektion beschreibt sie die Ziele und Maßnahmen zur Sicherung und Entwicklung der Qualität der Schulinspektion als Akteur im Konzert der Hamburger Schulentwicklung und intern in der Entwicklung der Professionalität seiner Mitarbeiterinnen und Mitarbeiter und der Qualität seiner Instrumente und Verfahren.

Das QM-System wird von einer Stabstelle aufgesetzt. Die Mitarbeiterinnen und Mitarbeiter verstehen Qualitätsentwicklung und Qualitätssicherung als wichtigen Teil des Arbeitsprozesses und sind an der Umsetzung und der Evaluation des Qualitätsmanagementsystems aktiv beteiligt.

Das Team Wissenschaft verantwortet die Qualitätsprüfung der Instrumente.

Die Schulinspektion nutzt die Ergebnisse aus der Evaluation ihrer Arbeit durch Externe (Landesinstitut, Schulaufsicht, Peer-Besuche anderer Länderinspektionen, Schulen) zur eigenen Qualitätsverbesserung.

Die Schulinspektion hat ein Ideen- und Beschwerdemanagement mit notwendigen Strukturen und Verantwortlichkeiten für die im Inspektionsprozess befindlichen Schulen – vertreten durch die Schulleitungen – und andere beteiligte Funktionsträgerinnen und -träger und Institutionen (wie Schulformexpertinnen und -experten, Wirtschaftsvertreterinnen und -vertreter, die Schulaufsicht etc.) eingerichtet. Es ist

Bestandteil der externen Qualitätssicherung und hat zum Ziel, die Prozessqualität lösungsorientiert zu sichern und kontinuierlich zu verbessern. Die Ziele und die Funktion des Ideen- und Beschwerdemanagements sind den Prozessbeteiligten bekannt und sollen als Angebot zur aktiven Teilnahme und zur Kooperation verstanden werden. Die Leitung und der/die Qualitätsmanagementbeauftragte (QMB) werten die Gesamtheit der eingegangenen Anliegen systematisch aus und sorgen für einen regelhaften Diskussionsprozess im Team, um Verbesserungen und Weiterentwicklungen abzuleiten. In geeigneter Weise werden die Anliegen der Schulen, der Umgang damit und mögliche Veränderungen bekannt gemacht.

Die Schulinspektion erachtet es für wichtig, dass eine externe Evaluation unter Einbindung von Inspektorinnen und Inspektoren kooperierender Schulinspektionen und weiterer Expertinnen und Experten aus dem akademischen Umfeld erfolgt. Die Schulinspektion strebt an, nach fünfjähriger operativer Tätigkeit ein Peer-Review durchzuführen. Der primäre Untersuchungsgegenstand des Peer-Reviews ist die Sicherung und Verstetigung der Prozessqualität. Der Untersuchungsrahmen wird durch die Leitung und den QMB festgelegt. In geeigneter Weise und im Einvernehmen zwischen Institutsleitung, Leitung sowie Mitarbeiterinnen und Mitarbeitern sollen die Ergebnisse erläutert, diskutiert und einem vorher festzulegenden Adressatenkreis bekannt gemacht werden. Das Peer-Review, die Ergebnisse und mögliche Handlungsbedarfe dienen der Vorbereitung auf eine spätere Zertifizierung.

Inspektionsprozess

Qualitätsbereich 6: Externe Evaluation

Die Schulinspektion stellt ihre Ergebnisse in einer Weise zur Verfügung, die Schulen die Möglichkeit gibt, sich zu entwickeln.

Die Schulinspektion berücksichtigt Daten aus dem regionalen Schülerdatensatz und dem Bildungsmonitoring in ihren Berichten.

Die Inspektorinnen und Inspektoren treten der Schule gegenüber wertschätzend auf und halten dabei professionelle Distanz. Sie sind klar in ihrer Rolle als externe Evaluatoren gegenüber allen Mitgliedern der Schulgemeinschaft.

Die Inspektorinnen und Inspektoren arbeiten sorgsam, verantwortlich und die Würde Dritter achtend mit ihnen anvertrauten Informationen über Personen, Prozesse und Ergebnisse.

Die Instrumente der Schulinspektion (Fragebögen, Unterrichtsbeobachtungsbögen, Interviews, Auswertungsraster) entsprechen Standards und werden professionell eingesetzt.

Die Inspektorinnen und Inspektoren bewerten nach einheitlichen Bewertungskriterien (Qualität der Urteilsbildung). Sie beziehen sich auf das verbindliche Berichtsraster, das eine Anpassung der Kriterien des Orientierungsrahmens Schul-

qualität an die Bedarfe des Inspektionsprozesses darstellt. Über die Handhabe der Kriterien findet eine regelmäßige Verständigung im Gesamtteam statt. Die besondere Situation der einzelnen Schule findet dabei Berücksichtigung.

Die Bewertungen vermitteln sich in der Präsentation prägnant und gut verständlich.

Die Berichte sind gut lesbar, prägnant und funktional. Sie sind in Form, inhaltlicher Gestaltung, den Bewertungskriterien und den getroffenen Bewertungen miteinander vergleichbar.

Qualitätsbereich 7: Gewährleistungsfunktion

Der Bericht der Schulinspektion stellt für die Schule und die Schulaufsicht ein sinnvolles Instrument dar, um Schulqualität zu evaluieren, Stärken zu sichern und Entwicklungsbedarfe zu identifizieren.

Es gibt ein zwischen Schulinspektion, Schulaufsicht und Landesinstitut abgestimmtes Verfahren zur Identifikation von „Schulen mit besonderem Handlungsbedarf“ und dem folgenden Umgang mit diesen.

Qualitätsbereich 8: Bereitstellen von Systemwissen/Bildungsmonitoring

Die Schulinspektion stellt ihr Systemwissen in adäquater und vielfältiger Weise der Hamburger Bildungslandschaft zur Verfügung (themenbezogene Veröffentlichungen, Veranstaltungen und Vorträge).

Die Schulinspektion stellt allen relevanten Akteuren innerhalb der Behörde unter anderem mit dem Jahresbericht Wissen zur Verfügung, das Steuerung ermöglichen soll.

Qualitätsbereich 9: Dienstleistungsansatz/Kundenorientierung

Die Schulinspektion erfüllt ihren Auftrag dienstleistungsorientiert: Das heißt, ihre Rückmeldungen sind klar und präzise, dabei respektvoll und die Arbeit der Schulen würdigend. Die Präsentation des Qualitätsprofils der Schule im Bericht wie ggf. vor der Schulöffentlichkeit dient dem Ziel, Schulentwicklungsarbeit anzuregen.

Qualitätsdimension 3: Outputqualität

Qualitätsbereich 1: Kundenorientierung: Zufriedenheit mit dem Inspektionsprozess

Die Schulleitungen und – soweit es eine öffentliche Präsentation des Inspektionsberichts gibt – die Gruppen der Schulgemeinschaft erleben den Inspektionsprozess als im Ergebnis den notwendigen Aufwand rechtfertigend.

Qualitätsbereich 2: Anwendungsorientierung: Nutzen von Bericht und Daten

Schulleitungen/Schulen und Schulaufsichten verstehen den Inspektionsbericht als differenzierte und nachvollziehbare Darstellung von Stärken und Schwächen der Schule.

Die beigefügten Daten können von den Schulen gelesen, verstanden und weiter ausgewertet werden.

Qualitätsbereich 3: Ergebnisorientierung: Inspektion fördert Schulentwicklung

Die Schulinspektion stellt den Schulleitungen Inspektionsberichte zur Verfügung, aus denen diese wichtige Anregungen zur eigenen Schulentwicklung ziehen können. Durch das Angebot, den Inspektionsbericht öffentlich zu präsentieren, fördert die Schulinspektion den Einbezug der Gruppen der Schulgemeinschaft in die Schulentwicklung.

Die Schulinspektion stellt den Schulaufsichten Inspektionsberichte zur Verfügung, aus denen diese Erkenntnisgewinn zur Steuerung der Schulen ziehen können.

Qualitätsbereich 4: Systemorientierung: Netzwerk Inspektion – Schulaufsichtsbehörde – Unterstützungssysteme

Schulaufsichten und Unterstützungssystemen wird es durch die Netzwerkarbeit mit der Schulinspektion ermöglicht, Ergebnisse der Schulinspektion für ihre Arbeit zu nutzen.

Die Schulinspektion beteiligt sich aktiv an der stetigen Ausgestaltung eines Netzwerks, das die Arbeit der Akteure im Schnittbereich von Qualitätsüberprüfung und Qualitätsentwicklung an Hamburger Schulen verzahnt, aufeinander bezieht und institutionell klar definiert.

Qualitätsbereich 5: Zufriedenheit der Mitarbeiterinnen und Mitarbeiter der Schulinspektion

Die in der Schulinspektion Beschäftigten fühlen sich wohl an ihrem Arbeitsplatz. Sie erleben das Verhältnis von Erfüllung und Beanspruchung als angemessen.

Literatur

- Landwehr, N. & Steiner, P. (2003). *Q2E – Qualität durch Evaluation und Entwicklung. Konzepte, Verfahren und Instrumente zum Aufbau eines Qualitätsmanagements an Schulen*. Bern: hep-Verlag.

Teil 2

Die Arbeit der Schulinspektion

Wirkungen von Schulinspektion: Ein Rahmen zur theoriegeleiteten Analyse von Schulinspektionseffekten

Marcus Pietsch, Ann-Katrin van den Ham & Olaf Köller

Zusammenfassung

Schulinspektionen sollen die Qualität von Schule und Unterricht verbessern und hierdurch zur Steigerung von Schülerleistungen beitragen. Um die derzeit gängigen Schulinspektionsverfahren wissensbasiert weiterzuentwickeln, intendierte Wirkungen mit höherer Wahrscheinlichkeit als bislang zu erzielen und somit die Wirksamkeit der Verfahren zu erhöhen, ist es notwendig, empirische Evidenz zu wirksamkeitsförderlichen Bedingungen und Vorgehensweisen zu sammeln. Eine gute Möglichkeit, dies zu erforschen, bieten theoriebasierte Evaluationen. Da die bisherige Forschung zu Wirkungen von Schulinspektionen eher cursorisch erfolgt, wird im Rahmen des vorliegenden Beitrages ein dreischrittiges Verfahren zur systematischen Evaluation von Schulinspektionseffekten vorgeschlagen. Hierbei wird insbesondere dafür plädiert, dass im Rahmen theoriegeleiteter Evaluation von Schulinspektionseffekten die Komplexität und die Kompliziertheit der Intervention stärker berücksichtigt werden als bisher und die im Rahmen programmtheoretisch fundierter Analysen generierten Befunde mittels des Ansatzes des Interpretation/Use Argument nach Kane validiert werden, um so die Belastbarkeit der Studien und der daraus abzuleitenden Konsequenzen für die Weiterentwicklung von Inspektionsverfahren zu erhöhen.

1. Einleitung

Schulinspektionen sollen dazu beitragen, Schülerleistungen zu verbessern. Dies, so die Erwartung, soll gelingen, indem sie Schulen durch die Rückmeldung beobachteter Stärken und Schwächen Impulse für die Weiterentwicklung von Schule und Unterricht liefern. Hierfür wird nach normativen Vorgaben, die in der Regel in landesspezifischen Qualitätsrahmen oder Qualitätstableaus formuliert wurden, durch Schulinspektoren extern an Schulen evaluiert. Es wird dann davon ausgegangen, dass Schulverantwortliche die zurückgemeldeten Informationen für eine wissensbasierte Schul- und Unterrichtsentwicklung nutzen, die sich wiederum in verbesserten Schülerleistungen niederschlägt. Wie Pietsch, Schulze, Schnack und Krause (2011) herausstellen, knüpfen die diesbezüglichen Wirksamkeitserwartungen an Schulinspektionen vor allem an die Forschung zum zielorientierten

Feedback (vgl. Kluger & DeNisi 1996; Ramaprasad 1983; Visscher & Coe 2003) an. Entsprechend wird erwartet, dass das empirisch begründete Aufzeigen von Differenzen zwischen normativ vorgegebenen Soll- und empirisch beobachteten Ist-Zuständen dazu führt, dass in extern evaluierten Schulen infolge der Rückmeldung eine Handlungsoptimierung geplant wird, die es ermöglicht, anzustrebende Ziele in Zukunft besser zu erreichen.

Eine Vielzahl von Studien zur Wirksamkeit von Schulinspektionen zeigt jedoch, dass dieses Ziel häufig nicht erreicht wird oder sogar nicht-intendierte Nebeneffekte beobachtbar sind. Die Befundlage ist hierbei sehr heterogen. So zeigen einzelne Studien, dass sich Effekte auf die innerschulische Qualitätsentwicklung finden lassen, und andere Studien negieren genau diesen Befund (vgl. zur Übersicht z. B. de Wolf & Janssens 2007; Dederling 2012; Husfeldt 2011). Auch mit Blick auf die Verbesserung von Schülerleistungen weisen bislang nur zwei Studien aus den Niederlanden und Deutschland positive Effekte nach (vgl. Pietsch et al. 2014; Luginbuhl et al. 2009). Diejenigen Studien hingegen, die sich auf Effekte der englischen Schulinspektion OFSTED (*The Office for Standards in Education*) beziehen, zeigen, dass bezogen auf das Gesamtsystem keine oder sogar leicht negative Wirkungen auf Schülerleistungen als Folge von Schulinspektionen beobachtet werden können (vgl. de Wolf & Janssens 2007). Darüber hinaus wird in einzelnen Studien von nicht-intendierten Nebenwirkungen bei Schulinspektionen berichtet. Demnach steigt den Befunden zufolge das Stresserleben bei Lehrkräften während einer Inspektion (vgl. z. B. Brunson et al. 2006) oder es findet eine überzogene positive Selbstdarstellung der Schule (*Window Dressing*) statt (vgl. de Wolf & Janssens 2007).

Insgesamt findet sich somit auch zehn Jahre nach Einführung der ersten Schulinspektorate in Deutschland eine uneindeutige Befundlage, die keine kohärente Einschätzung darüber ermöglicht, ob Schulinspektionen ein Instrument sind, von dem erwartet werden kann, dass es zu einer Verbesserung von Schul- und Unterrichtsqualität und infolge dessen zu einer Steigerung von Schülerleistungen führt. Pietsch et al. (2014) haben vorgeschlagen, diesem Problem mithilfe einer Kombination kausalanalytischer Blackbox-Verfahren und theoriegeleiteter Evaluationen zu begegnen. Während die erste Herangehensweise Aufschluss darüber geben kann, ob Inspektion grundsätzlich die intendierten Effekte erzielt oder nicht, ist letztere insbesondere dann sinnvoll, wenn es um Entscheidungen zur weiteren Gestaltung, zu Optimierung, Veränderung oder Verbesserung von Inspektionsverfahren geht. Einen Punkt, den bereits Cronbach et al. (1980, S. 251) als besonders relevantes Kernelement von Evaluationen herausstellen:

„Knowing this week’s score does not tell the coach how to prepare next week’s game. The information that an intervention had satisfactory or unsatisfactory outcomes is of little use by itself; users [...] need to know what led to success or failure.“

Während die Evaluation mithilfe von Blackbox-Verfahren zwar methodisch anspruchsvoll, ansonsten jedoch geradlinig und überschaubar ist, erfordert eine theoriegeleitete Evaluation von Inspektionswirkungen eine komplexere Herangehensweise, um zu belastbaren Resultaten zu gelangen (vgl. Astbury & Leeuw 2010). Im Folgenden wird daher ein Rahmen für eine theoriegeleitete Analyse von Schulinspektionseffekten vorgestellt, die es ermöglichen soll, Befunde zur Wirkung von Schulinspektionen systematisch zu generieren. Hierzu werden zuerst Annahmen zur Wirksamkeit und den Funktionen von Schulinspektionen beschrieben. Anschließend wird dargestellt, welche Gründe es aus programmtheoretischer Perspektive dafür geben kann, dass Schulinspektionen ggf. nicht wirksam werden, um anschließend auf ein strukturiertes Vorgehen zur Evaluation von Schulinspektionswirkungen einzugehen.

2. Annahmen zur Wirksamkeit und zu Funktionen bei Schulinspektionen

2.1 Funktionen von Schulinspektionen

Schulinspektion hat die Aufgabe, Informationen aus dem Bereich Bildung und Erziehung systematisch zu beschaffen und aufzubereiten, um auf diesem Wege Akkreditierungen, Rechenschaftslegungen und Diagnosen für systemisches Lernen im Bildungssystem zu ermöglichen. Schulinspektionen sollen entsprechend dazu beitragen, elementare Standards von Bildungsqualität zu gewährleisten, einen verbesserten Service für das einzelschulische Qualitätsmanagement zu bieten und unterschiedlichen Akteuren im Bildungssystem verwertbare Informationen zu schulischen Prozessqualitäten bereitzustellen (vgl. Döbert et al. 2008; Maritzen in diesem Band; Pietsch et al. 2009, 2013). Dabei liegt der Schwerpunkt der Inspektionsarbeit auf der Evaluation einzelschulischer Prozesse (vgl. Böttcher & Kotthoff 2010; Döbert et al. 2008; Maritzen in diesem Band; van Ackern & Klemm 2009), wobei die Inspektionen in den Ländern trotz augenscheinlicher Unterschiede in Ausstattung, Ausgestaltung und ministerieller Anbindung „auf Grundlage von Verfahren mit wiederkehrenden Standardelementen“ erfolgen (Maritzen in diesem Band, S. 16).

Grundlage für die externe Einzelschulevaluation bilden dabei in allen Ländern die landesspezifischen Qualitätsrahmen oder Qualitätstableaus, die im Sinne eines Input-Prozess-Output-Modells (teilweise unter Berücksichtigung von Kontextfaktoren) Anforderungen an Schulqualität definieren (vgl. Ehren & Scheerens in diesem Band; Müller 2010). In allen Inspektionen wurden, basierend auf diesen Qualitätskatalogen, Leistungsindikatoren definiert, die es ermöglichen sollen, die Qualität einzelner Prozesse möglichst differenziert zu erfassen (vgl. Böttcher

& Kotthoff 2007). Geschulte Experten – meist Lehrkräfte, Schulleitungen oder Schulaufsichten – evaluieren anhand dieser Indikatoren die Schulen in Teams von zwei bis vier Personen (vgl. Döbert et al. 2008; Müller 2010). Die Datenerhebung selbst erfolgt über ein umfangreiches Repertoire an Methoden. Fragebögen und Interviews, gerichtet an verschiedene Schulbeteiligte, sowie Schulbegehungen und die Begutachtung schulinterner Dokumente und Statistiken gehören derzeit ebenso wie die Beobachtung von Unterrichtseinheiten zum länderübergreifenden Methodenstandard (vgl. Döbert et al. 2008). Die derart ermittelten Befunde werden in einem Bericht zusammengefasst und den Schulen sowie ggf. weiteren Beteiligten, wie z. B. Schulaufsichten, übergeben. In regelmäßigem Turnus werden darüber hinaus Berichte verfasst, die Befunde der Schulinspektionen auf Systemebene präsentieren.

Durch diese Rückmeldungen aus Schulinspektionen erhalten die Schulbeteiligten, ebenso wie Bildungsadministration und Bildungspolitik, Informationen zur extern wahrgenommenen Qualität innerschulischer Prozesse auf verschiedenen Ebenen. Obwohl dies häufig nicht explizit benannt wird, steht hinter dem Vorgehen, aufbereitete Informationen an verschiedene Gruppen auf Schul- und Schulsystemebene zurückzumelden, die Idee, dass die Rückmeldungen von Evaluationsergebnissen als Grundlage für Entwicklungen an den evaluierten Schulen genutzt werden, die in ihrer Konsequenz in verbesserte fachliche Schülerleistungen münden (vgl. de Wolf & Janssens 2007; Ehren & Visscher 2006; Ehren et al. 2005). Insbesondere in den deutschen Ländern zeichnet sich diesbezüglich aktuell ein Trend ab (vgl. Böttcher & Kotthoff 2010, S. 307), „demzufolge Schulinspektionen [...] als ein Unterstützungs- und Dienstleistungsangebot für die Schule verstanden werden, deren Erkenntnisse wirksam für die interne Schulentwicklung genutzt werden können bzw. sollen“. Folgerichtig betont Köller (2009), dass der Wert von Schulinspektionen insbesondere darin läge, dass sie Hinweise auf die Prozesse vor Ort gäben, mithilfe derer Hypothesen darüber generiert werden könnten, welche Veränderungen von Schule und Unterricht zu höheren Lernerträgen der Schülerinnen und Schüler führen können.

2.2 Annahmen zur Nutzung von Inspektion als Grundlage für Schulentwicklung

Entsprechend wird davon ausgegangen, dass Schulverantwortliche die Inspektion und ihre Berichte nutzen, um hieraus konkrete Maßnahmen für die Schul- und Unterrichtsentwicklung abzuleiten (vgl. Ehren & Visscher 2006, 2008). Im Zentrum eines solch entwicklungsfokussierten, anwenderorientierten Verständnisses von Evaluation, wie sie den Schulinspektionsverfahren in den deutschen Bundesländern zugrunde liegt, steht daher zwangsläufig der Evaluationsnutzen, der sich aus

der Untersuchung der Qualität von Schule und Unterricht für die Schulbeteiligten ergibt. Dabei ist es von grundlegender Bedeutung zu wissen, in welcher Art und Weise Informationen aus Schulinspektionsverfahren innerhalb von Schulen genutzt werden können, wenn das Ziel Qualitäts- und Schulentwicklung heißt, da „das Verwendungsinteresse der [...] Anwendersysteme [...] unterschiedlich ist, je nachdem ob die Evaluation eher zur Überprüfung/Kontrolle, zur Legitimation, zur Innovation/Entwicklung oder lediglich zur Bestätigung durchgeführt wird“ (vgl. Stamm 2003, S. 195 f.). Diesbezüglich lassen sich mit Blick auf die Nutzung von Evaluationsergebnissen grundsätzlich drei Nutzungsformen voneinander unterscheiden, die je einen spezifischen Zweck verfolgen (vgl. King & Pechman 1984; Weiss 1998):

- 1) Instrumentelle Nutzung (*Instrumental Use*): Die instrumentelle Nutzung von Informationen aus Evaluationen bezieht sich darauf, konkrete Probleme zu lösen bzw. konkrete Entscheidung zu treffen, die den Evaluationsgegenstand betreffen.
- 2) Konzeptionelle Nutzung (*Conceptual Use*): Die konzeptionelle Nutzung von Informationen aus Evaluationen findet statt, wenn diese indirekt genutzt werden, um das Wissen bzgl. des Evaluationsgegenstandes zu erweitern.
- 3) Symbolische Nutzung (*Symbolic Use*): Die symbolische Nutzung von Informationen aus Evaluationen findet statt, wenn Befunde zum Evaluationsgegenstand eingesetzt werden, um bereits getroffene Entscheidungen gegenüber Dritten zu legitimieren.

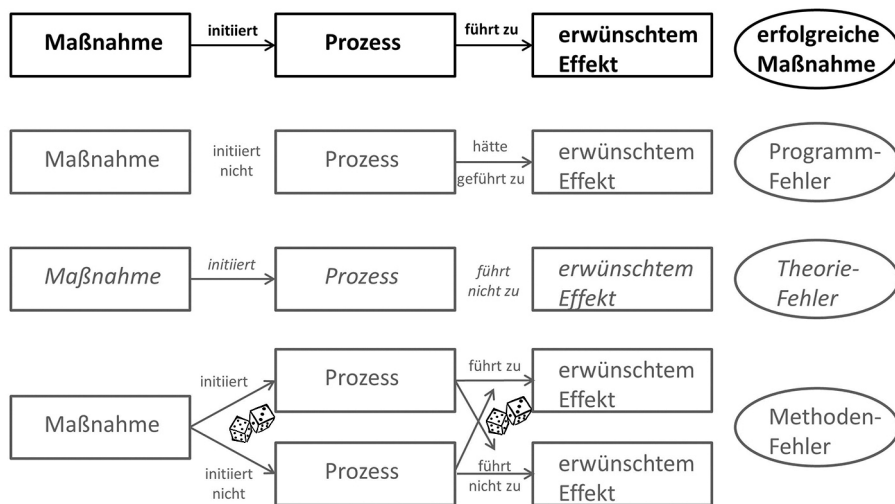
Dabei dient die instrumentelle Nutzung primär dem Zweck der konkreten und zielgerichteten Weiterentwicklung, die konzeptionelle Nutzung wiederum hat die indirekte Veränderung von Einstellungen, Meinungen und Kognitionen der an der Evaluation Beteiligten zum Ziel und die symbolische Nutzung dient der Bekräftigung von Positionen in Diskursen und argumentativen Auseinandersetzungen (vgl. Weiss 1998). Für die Nutzer der Evaluationsbefunde auf Ebene der Einzelschule, müssen die evaluativ gewonnen Informationen daher, wenn es um die konkrete Weiterentwicklung von Schule und Unterricht geht, in erster Linie einen instrumentellen, einen direkten praxisbezogenen Nutzen zur Entscheidungsfindung erbringen (vgl. Scheerens 2007). Einen Nutzen also, der es ermöglicht, inkrementelle Entscheidungen zu treffen, um so z.B. ein bestimmtes Programm zu beenden, es zu modifizieren, Aktivitäten zu verändern oder Weiterbildungsmaßnahmen zu gestalten (vgl. Weiss 1998). Eine solche instrumentelle Nutzung von Evaluationsbefunden kann jedoch nur unter einer der vier folgenden Voraussetzungen stattfinden (vgl. Weiss 1980): 1) wenn die Implikationen, die sich aus den Befunden ergeben, nicht kontrovers sind; 2) wenn die geplanten Veränderungen sich innerhalb des bisher geplanten Programms bewegen und nicht allzu umfassend sind; 3) wenn das

Umfeld des Programms stabil ist, es keine größeren Veränderungen im Bereich von Leitung, Finanzen und Teilnehmern gibt oder wenn 4) ein Programm in der Krise ist und niemand weiß, was getan werden muss. Die direkte instrumentelle Nutzung, die in Veränderungen und Entwicklung mündet, findet somit vor allem dort statt, wo der bestehende Status quo nicht gestört wird (vgl. Nutley et al. 2003).

3. Problemstellung

Wirksamkeits- und Wirkungsstudien untersuchen nun, ob Schulinspektionen durch Schulverantwortliche genutzt werden und welche Wirkungen erreicht werden. Intendiert ist eine instrumentelle Nutzung, die zu einer Verbesserung von Schülerleistungen und diesbezüglich vermittelnden Faktoren der Schul- und Unterrichtsqualität führt. Lassen sich diese Zusammenhänge nicht nachweisen, so kann dies theoretisch auf drei Gründe zurückgeführt werden (vgl. Pietsch et al. 2013; Stame 2010): Erstens kann ein Programm- resp. Implementations-, zweitens ein Theorie- und/oder drittens ein Methodenfehler vorliegen (vgl. Abb. 1).

Abbildung 1: Wirkungskette und mögliche Ursachen für Unwirksamkeit aus programm-theoretischer Perspektive



Während ein Programmfehler vorhanden ist, wenn es mittels einer Intervention nicht gelingt, eine intendierte Wirkung nachweisbar zu erzeugen, liegt ein Theoriefehler vor, sofern ein theoretisch postulierter Wirkungszusammenhang nicht ausreichend valide begründet wurde. Darüber hinaus kann eine nicht nachweisbare

Wirkung letztlich auch darauf zurückzuführen sein, dass Methodenfehler vorliegen. Die Nicht-Wirksamkeit einer Intervention rührt in diesem Falle daher, dass es im Rahmen der empirischen Kausalstudie aufgrund methodologischer Unzulänglichkeiten nicht gelingt, den angenommenen Wirkungszusammenhang empirisch verlässlich zu überprüfen.

Empirische Untersuchungen lassen sich zu allen drei Bereichen finden, wobei derzeit ein grundsätzliches Problem darin besteht, dass Programmtheorien, die eine theoriebasierte Evaluation von Inspektionswirksamkeit ermöglichen sollen, zu unterkomplex sind und die Kompliziertheit der Intervention Schulinspektion nicht berücksichtigen (vgl. de Wolf & Janssens 2007; Pietsch et al. 2013). Dies wiederum hat zur Folge, dass Aussagen zu Programmfehlern – also dazu, ob Schulinspektionen ggf. nicht zufriedenstellend implementiert sind und/oder ihrer Aufgabe nur unzureichend nachkommen – nur schwer möglich sind und die bislang vorliegenden Befunde daher nur wenig verlässliche Aussagen zur Wirksamkeit von Schulinspektionen auf die Entwicklung von Schule und Unterricht sowie auf Schülerleistungen ermöglichen. Entscheidungen zum weiteren Umgang mit Inspektionsverfahren sind daher aktuell nur schwer möglich.

Um die Wirkungen von Inspektionen systematisch zu erforschen, schlagen wir daher ein dreischrittiges Verfahren vor, das der Komplexität des Untersuchungsgegenstandes gerecht zu werden versucht. Hierbei soll in einem ersten Schritt gewährleistet werden, dass das Zielkriterium, der Effekt, um den es geht, sowie der logisch angenommene Wirkungszusammenhang vor Beginn einer Untersuchung klar und eindeutig definiert wird. Im zweiten Schritt geht es darum, die Wirkmechanismen im Rahmen eines logischen Modells zu beschreiben und zu untersuchen, das der Komplexität und der Kompliziertheit des Untersuchungsgegenstandes gerecht wird. In einem dritten Schritt sollen die Ergebnisse dieser Untersuchungen zusammengeführt werden, um hieraus belastbare Konsequenzen auf Systemebene, aber auch auf Ebene der einzelnen Inspektion ziehen zu können.

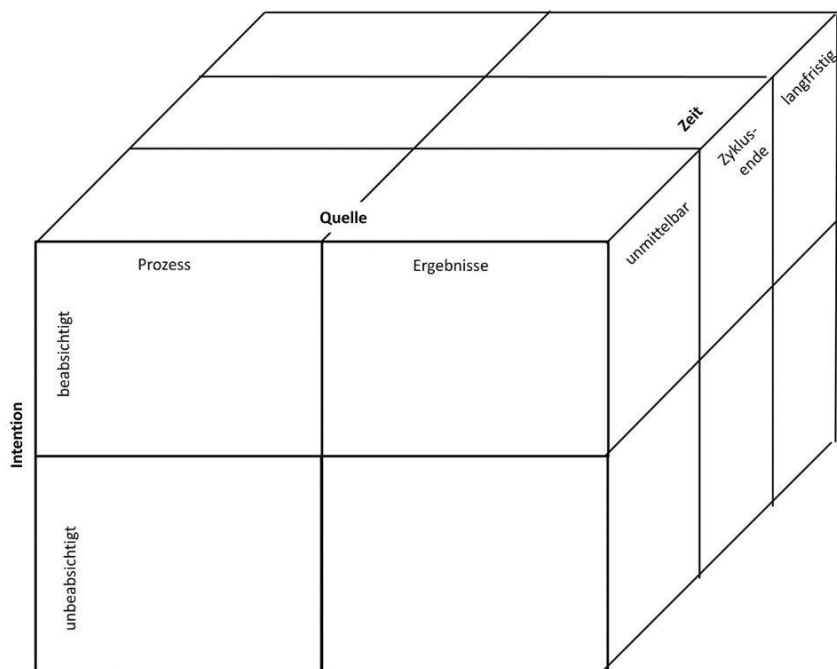
4. Ein konzeptioneller Analyserahmen zur Evaluation von Schulinspektionswirkung

4.1 Grundlage: Definition von Inspektionswirkungen mithilfe eines mehrdimensionalen Modells

Schulinspektionen können, wie beschrieben, sowohl intendierte als auch nicht-intendierte Wirkungen nach sich ziehen, aber auch vollkommen wirkungslos bleiben. Dabei geht es um den Einfluss (im fachwissenschaftlichen Diskurs wird diesbezüglich von ‚*Influence*‘ gesprochen, vgl. z. B. Herbert 2014), den Inspektio-

nen auf Schule, Unterricht und Schülerleistungen haben. Daher ist es im Rahmen einer theoriegeleiteten Evaluation von Schulinspektionswirkungen in einem ersten Schritt notwendig, zu klären, welche Wirkung in welchem Zeitraum generiert und durch welchen Prozess (Durchführung einer Inspektion oder Rückmeldung von Ergebnissen) erwartet wird. Möchte man also den Einfluss von Schulinspektionen auf Schülerleistungen sowie vermittelnde Faktoren wie Schul- und Unterrichtsentwicklung analysieren, benötigt man entsprechend elaborierte Modelle. Eben solche haben Kirkhart (2000) sowie Alkin und Taut (2003) vorgeschlagen. Mithilfe dieser dreidimensionalen Modelle lässt sich die Wirkung von Evaluationen – und entsprechend auch von Schulinspektionen – fassen (vgl. Abb. 2).

Abbildung 2: Modell zur Definition von Inspektionswirkungen (nach Kirkhart 2000)



Der Vorteil eines solchen Schemas liegt vor allem darin, dass er über den Begriff der Nutzung hinaus geht und es vor allem auch erlaubt, nicht-intendierte Wirkungen in die Analysen mit einzubeziehen und darüber hinaus Effekte resp. Veränderungen – und somit Wirkungen, nicht die Nutzung von Evaluationen – in den Mittelpunkt stellt. Weiterhin ermöglicht dieses Schema, zu berücksichtigen, dass Schulinspektionen, anders als z. B. Vergleichsarbeiten und Lernstandserhebungen, auch einen Einfluss durch das Wirken von Inspektorinnen und Inspektoren vor Ort

haben und nicht ausschließlich Effekte durch die Rückmeldung von Ergebnissen erzielen. Die Stärken eines solchen Modells stellt Herbert (2014, S. 394) entsprechend heraus:

- „influence provides a definition and a framework that reflects the full impact of evaluation and a cohesive way to organize theoretical and empirical knowledge of the effect evaluation can have on programs;
- by adopting this more comprehensive view, influence allows for the study of implicit mechanisms that affect change, including processes at the individual, interpersonal, and collective levels;
- influence frameworks are oriented around linkages to more developed constructs in other fields of literature such as attitude change, priming, skill acquisition, and persuasion;
- shifting to an influence framework allows for the study of pathways of influence and the study of situations where evaluation failed to affect change; and influence is built around social betterment as the ultimate goal of evaluation, rather than use.“

Untersucht werden kann in einem solchen Design somit: a) ob intendierte Effekte durch die Evaluation erzielt wurden oder nicht (*Intention*), b) ob diese durch die Ergebnissrückmeldung oder den Evaluationsprozess zustande gekommen sind (*Quelle*) und c) ob die Effekte unmittelbar, zum Ende eines Zyklus oder aber erst langfristig nachweisbar sind (*Zeit*).

4.2 Analyse: Theoriebasierte Evaluation mithilfe komplexer logischer Modelle

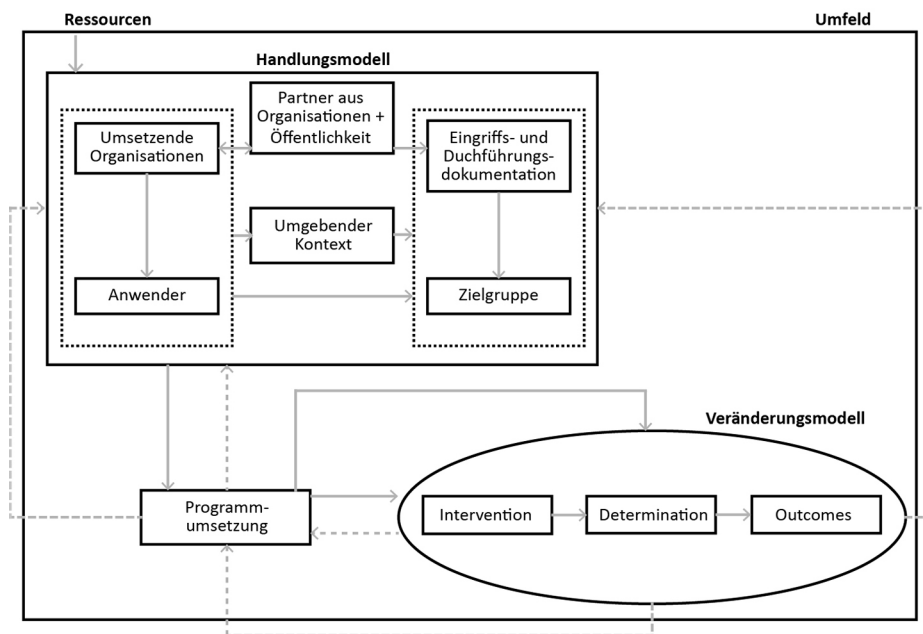
In einem zweiten Schritt geht es im Rahmen einer theoriegeleiteten Evaluation zu Wirkungen von Schulinspektion dann darum zu beschreiben, mithilfe welcher Mechanismen, auf welche Art und Weise die definierten Wirkungen durch ein Schulinspektionsverfahren generiert werden sollen. Von einfach-linearen Modellen wird abgeraten, da diese für die Analyse von Schulinspektionswirkungen zu unterkomplex sind: „The complexity and situatedness of much organizational activity begs for a style of theory that preserves some ambiguity.“ (vgl. Weber 2006, S. 120)

Komplexe programmtheoretische Modelle bestehen in der Regel (vgl. Chen 2006; Coryn et al. 2011) aus mindestens zwei Teilen (vgl. Abb. 3), die eng miteinander verknüpft sind und es u. a. ermöglichen, Kontextbedingungen und/oder rekursive Wirksamkeitsmechanismen zu inspizieren. Chen (2006) unterscheidet diesbezüglich zwischen einem Veränderungs- und einem Handlungsmodell:

„The action model and change model are closely related to each other and are essential for the success of a program. On the one hand, a change model is needed to justify the selection of an intervention for achieving the goals and/or outcomes and it provides a basis for developing the action model. On the other hand, the action model provides a blueprint to organize program activities and to activate and energize the change model for achieving program goals.“

Dabei umfasst das *Veränderungsmodell* drei Komponenten: a) die Intervention, die sich auf das Programm bezieht, mit dessen Hilfe Veränderungen bewirkt werden sollen, b) die Determination, die diejenigen Prozesse und Mechanismen umfasst, die den Input durch die Intervention, mit Blick auf mögliche Effekte, vermitteln, und c) die Outcomes, womit die antizipierten Effekte des Programms gemeint sind. Das Veränderungsmodell unterstellt entsprechend, dass die Implementation einer Intervention diejenigen Determinanten kausal beeinflusst, die ihrerseits zu Veränderungen in den Outputs führt.

Abbildung 3: Modell zur Evaluation von Inspektionswirkungen (nach Chen 2006)



Das *Handlungsmodell* hingegen beschreibt die Voraussetzungen und Annahmen der Intervention selber, stellt also dar, mit welchen Mitteln, unter welchen Voraus-

setzungen und bei welchen Adressaten die Intervention Wirkung erzielen soll. Das Handlungsmodell umfasst dabei sechs Bereiche:

- 1) *Umsetzende Organisationen*: Die Organisationen, die für die Koordination von Personal und die Ressourcenallokation bzgl. der Intervention verantwortlich sind.
- 2) *Anwender*: Die Personen, die für die konkrete Umsetzung resp. Implementierung des Programms verantwortlich sind.
- 3) *Partner aus Organisationen und Öffentlichkeit*: Die Organisationen oder Partner, die zusätzlich zur Primärorganisation notwendig sind, um die Intervention zielgerichtet durchzuführen.
- 4) *Umgebender Kontext*: Der ökologische Kontext, der direkt mit dem zu implementierenden Programm interagiert, z. B. in Form von Zielvorgaben, Normen, Gesetzen.
- 5) *Eingriffs- und Durchführungsdokumentation*: Der Umsetzungsplan, der die Intervention inkl. ihrer spezifischen Inhalte und Aktivitäten beschreibt und dabei die Schrittfolge definiert, mit der die Intervention im Feld umgesetzt werden soll.
- 6) *Zielgruppe*: Die Zielgruppe, für die das intervenierende Programm gedacht ist, insbesondere das Vorhandensein von klar definierten Ansprechpartnern und die Einbindung in das Programm.

4.3 Konsequenzen: Validierung mithilfe von *Interpretation/Use-Argumentationen*

Die Ergebnisse, die mit diesen Modellen generiert werden können, müssen in einem letzten Schritt sinnvoll zusammengeführt und validiert werden, um hieraus Konsequenzen auf Systemebene, aber auch auf Ebene der einzelnen Inspektion zu ziehen. Scriven (2012, S. 2 ff.) betont konsequenterweise, dass Validität im Rahmen von Evaluation das zentrale Kriterium ist:

„Validity: This is the key criterion – the matter of truth. [...] It requires some evidence of ‚real‘ value, which usually (not quite always) means visible or directly testable evidence somewhere along the line of implications of the evaluation.“

Hierbei geht es dann vor allem darum, die Annahmen zur Wirkungsweise von Inspektionen zu prüfen. Was aber meint Validierung hier konkret? In den Standards für Pädagogisches und Psychologisches Testen von 1999 und 2014, die fachlich-methodische und ethische Richtlinien unter anderem für Validität beinhalten, wird Validität als das Ausmaß definiert, in dem empirische Nachweise und Theorien

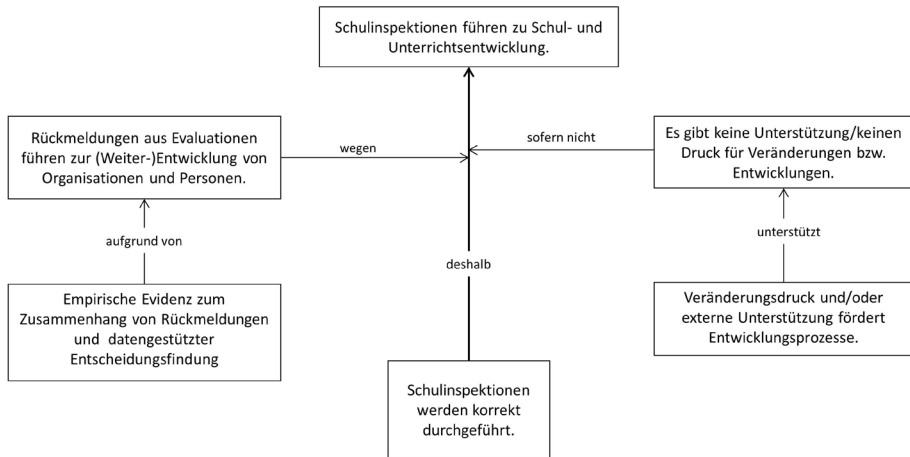
die Interpretation von Testergebnissen für eine bestimmte Nutzung stützen. Die Standards unterscheiden fünf Kategorien der Evidenz (Testinhalt, Antwortprozesse, Interne Struktur, Zusammenhänge mit anderen Variablen und Testkonsequenzen), die bei der Validierung evaluiert werden können. Jedoch wird Validität als ein einheitliches Konzept definiert (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999; S. 14):

„It is the degree to which all the accumulated evidence supports the intended interpretation of test scores for the proposed use.“

Zu validieren bedeutete in diesem Sinne, Argumente für und gegen die geplante Interpretation von Testergebnissen oder, im Falle von empirischen Untersuchungen zur Wirksamkeit von Schulinspektionen, von Studienergebnissen, zu konstruieren und zu evaluieren. Damit spiegeln die Standards den professionellen Konsens darüber wider, dass Validieren das Sammeln von Evidenz für eine wissenschaftlich solide Basis der Testwertinterpretationen und Schlussfolgerungen beinhaltet. In diesem Sinne ist Validität graduell. Für unterschiedliche Tests werden unterschiedliche Arten der Evidenz benötigt, wobei Validität ein einheitliches Konstrukt bleibt. Dabei ist Validität keine Eigenschaft des Tests oder der Untersuchung, sondern bezieht sich auf deren Interpretationen (vgl. Kane 2006; Shaw & Crisp 2012; Messick 1989; American Educational Research Association, American Psychological Association, National Council on Measurement in Education 1999 und 2014).

Neuere Validierungsansätze, die Validitätsnachweise in Form von Validitätsargumenten organisieren, wie der *Argument Based Approach* von Kane (2013), das *Evidence-Centered Design* von Mislevy (2003) und das *Assessment Use Argument* von Bachman (2005), basieren auf Toulmins Beschreibung von „Informellen Argumenten“ (1958, 2003). Die informellen Argumente werden in Form einer Argumentationskette aufgebaut, die für eine bestimmte Schlussfolgerung plädiert. In dieser Argumentationskette werden die Plausibilität von Behauptungen und Beobachtungen begründet und Zusammenhänge zwischen diesen bewiesen. In seinem Schema schließt der Argumentierende aufgrund von Daten und mithilfe einer Schlussregel auf die Konklusion. Dabei liefert eine Stützung Beweise für die Schlussfolgerung, Qualifikatoren bringen die Stärke der Behauptung zum Ausdruck und eine Ausnahme beschränkt oder verstärkt die Behauptung.

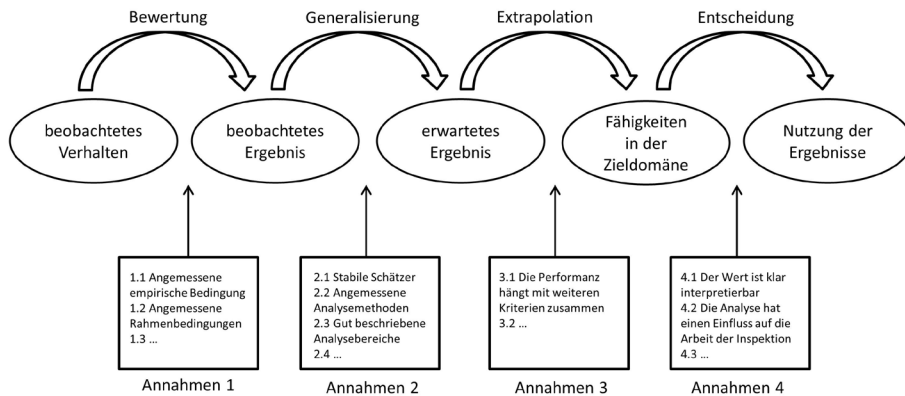
Abbildung 4: Modell zur argumentativen Validierung von Schulinspektionswirkungen (nach Toulmin 2003)



Ein Beispiel für so eine Behauptung wäre, dass Schulinspektion zu einer Verbesserung der Schul- und Unterrichtsqualität führt (siehe Abb. 4). Das Argument hierfür lautet, dass sie korrekt durchgeführt werden. Die Schlussregel ist eine Generalisierung, die genutzt wird, um die spezifischen Daten mit der spezifischen Behauptung zu verbinden. Im Falle des Beispiels in Abbildung 4 lautet die Schlussregel, dass Schulinspektionen, die korrekt durchgeführt werden, in der Regel zu einer Verbesserung der Schul- und Unterrichtsqualität an evaluierten Schulen führen. Die Stützung liefert Beweise für die Schlussfolgerung in Form von Theorien, Untersuchungen, Erfahrung etc.; alternative Erklärungen relativieren eine Behauptung und werden durch Ausnahmen beschränkt oder verstärkt.

Kane geht noch weiter als Toulmin (2003) bzw. Mislevy (2003) und argumentiert, dass für die Verbindung der Beobachtung mit der angestrebten Behauptung mehrere Arten von Schlussfolgerungen notwendig sind. In Abbildung 5 wird das sogenannte *Interpretation/Use Argument* (IUA) dargestellt, welches die multiplen Schlussfolgerungen und Ergebnismutzungen spezifiziert. Jede dieser Schlussfolgerungen basiert auf Annahmen, die Stützung benötigen. Die erste Schlussfolgerung „Bewertung“ geschieht bei der Ergebnisermittlung. Das beobachtete Verhalten bei oder infolge der Durchführung einer Inspektion wird in ein beobachtetes Ergebnis umgesetzt.

Abbildung 5: Interpretation/Use Argument zur Validierung von Schulinspektionswirkungen (nach Kane 2002)



Dabei wird beispielsweise angenommen, dass die Bewertungsprozeduren angemessen sind, korrekt angewendet wurden und frei von offenkundigen Fehlern sind. Die zweite Schlussfolgerung bezieht sich auf die Generalisierung der Ergebnisse aus einer ausgewählten Messung/Studie (mit bestimmten Kriterien, an einem bestimmten Tag, zu einer bestimmten Zeit) auf andere, vergleichbare Situationen. Dabei wird unter anderem davon ausgegangen, dass die Stichprobe der Umstände, unter denen gemessen wird, repräsentativ ist. Oft werden die Interpretationen auf weitere Bereiche wie z.B. die Realsituation erweitert. Die hierauf gründenden Schlussfolgerungen ermöglichen eine deskriptive Interpretation der Ergebnisse. Die Einbeziehung von Verwendungen und Entscheidungen führt schließlich zu einer entscheidungsbasierten Interpretation (Kane 2002). Bei der Schlussfolgerung „Entscheidung“ werden die Ergebnisse für das Treffen von Konsequenzen verwendet (Kane 2013).

5. Fazit

Vor rund zehn Jahren wurden die ersten Schulinspektorate in Deutschland eingeführt. Die Frage, welche Wirkungen Schulinspektionen nach sich ziehen, ist bislang jedoch nur in Ansätzen und darüber hinaus sehr unsystematisch erforscht. Befunde zur Wirksamkeit, also bezogen darauf, ob Schulinspektionen intendierte Wirkungen erzielen, liegen noch seltener vor. Problematisch ist darüber hinaus, dass die Befunde häufig widersprüchlich sind und keine Generalisierung ermöglichen. Evidenzbasierte Entscheidungen zum weiteren Umgang mit Inspektionen, zu

Veränderungs- und Verbesserungsmöglichkeiten, aber auch zur Einbettung in die länderspezifischen Bildungssysteme sind daher kaum möglich. Aktuelle Veränderungen basieren entsprechend eher auf anekdotischer Evidenz und auf Alltagserfahrungen, denn auf empirischer Evidenz.

Um die Wirksamkeit von Inspektionen jedoch nachhaltig und erfolgreich zu optimieren, sind Entscheidungsträger auf belastbare Befunde aus empirischen Untersuchungen und Evaluationen zu Wirkungen von Inspektion auf erwartete Zielkriterien wie Schul- und Unterrichtsqualität, aber auch Schülerleistungen angewiesen. Wir denken, dass der von uns vorgestellte konzeptionelle Rahmen zur Evaluation von Schulinspektionswirkungen dazu beitragen kann, derartige Informationen zu generieren, sofern er systematisch, im Sinne eines Forschungsprogramms angewendet wird.

Besonders wichtig erscheint es uns dabei, an die aktuelle Validitätsdebatte in der empirischen Bildungsforschung anzuknüpfen und das *Interpretation/Use Argument* von Kane im Rahmen einer theoriegeleiteten Evaluation von Schulinspektionswirkungen anzuwenden. Dieses Modell kann unserer Ansicht nach zur Evaluation der Wirksamkeit von Schulinspektionen genutzt werden und Programm-, Theorie- und Methodenfehler aufdecken bzw. vorbeugen. Dabei wird die Argumentationskette zweimal durchlaufen. In der ersten Phase wird die Maßnahme, in diesem Fall die Schulinspektion, evaluiert. Bei der Bewertung kann überprüft werden, ob mit den Instrumenten ermittelte Ergebnisse das Verhalten der Schule tatsächlich repräsentieren oder ob beispielsweise die Bewertung der Fragebögen, Interviews, Schulbegehungen etc. subjektiv oder fehlerbehaftet ist. Bei der Generalisierung kann untersucht werden, ob sich die Schulinspektion auf andere Facetten und Kontexte der Durchführung übertragen lässt. Hier stellt sich unter anderem die Frage, ob eine Inspektion mit anderen Inspektoren, zu einem anderen Zeitpunkt, in anderen Unterrichtsstunden oder bei anderen Lehrkräften ein gleichwertiges Ergebnis produziert hätte. Bei der Extrapolation ist es möglich, zu analysieren, ob das Konstrukt einer „guten Schule“ empirisch bestätigt werden kann und ob die Erfüllung dieser Normen tatsächlich eine „gute Schule“ in der realen Welt bedingt. Lassen sich keine Argumente oder sogar Gegenargumente dafür finden, dass das Ergebnis der Schulinspektion die reale Schulwirklichkeit beschreibt, kann dies keinen oder einen unerwünschten Prozess zur Folge haben. Gründe können ein Programm- oder ein Methodenfehler in der Schulinspektion sein. Können jedoch Argumente gefunden werden, dass das Ergebnis der Schulinspektion die reale Schulwirklichkeit beschreibt, so liegt eine solide Basis für die Schulentwicklung vor. Auf dieser Basis können dann Entscheidungen getroffen und ein gewünschter Prozess initiiert werden. Der Prozess führt wiederum zu einem eventuell veränderten beobachtbaren Verhalten der Schule. Dieses Verhalten kann erneut mithilfe von Instrumenten gemessen werden, womit das Modell von Kane ein zweites Mal durchlaufen werden kann. Bei der Bewertung kann

erneut untersucht werden, ob die mit den Instrumenten generierten Ergebnisse das Verhalten der Schule korrekt abbilden. Bei der Generalisierung ist es möglich, festzustellen, ob die Befunde aus verschiedenen Untersuchungen einheitlich und übertragbar sind. Bei der Extrapolation kann wiederum evaluiert werden, ob sich die Theorie über den Prozess durch die Testergebnisse bestätigen lässt und ob der Prozess zu erwünschten Effekten in der Wirklichkeit führt. Hierbei können Theorien und Ergebnisse aus dem Handlungsmodell und Veränderungsmodell (Chen 2006) integriert werden. Deuten die Argumente darauf, dass mit dem Instrument nicht die gewünschten Effekte des Prozesses in der Wirklichkeit repräsentiert werden, so kann es sich um einen Theoriefehler des Prozesses oder einen Methodenfehler in der Messung des Prozesses handeln. Bei der Entscheidung geht es wiederum um die Nutzung der Ergebnisse. In diesem Schritt stellt sich u. a. die Frage, ob die Maßnahme als erfolgreich gewertet werden kann, ob sie positive oder negative Konsequenzen für die Schule, die Lehrkräfte oder Schülerinnen und Schüler hat, etc. In diesem Schritt können auch die Aspekte aus dem Modell zur Kategorisierung von Inspektionswirkungen berücksichtigt werden, beispielsweise durch die Kategorisierung von beabsichtigten und unbeabsichtigten Prozessen und Ergebnissen über die Zeit.

Auf diese Weise bietet das Modell von Kane die Möglichkeit, Schulinspektion und ihre Konsequenzen umfassend zu evaluieren. Die Evaluation der einzelnen Schritte ist transparent und offenbart die evidenzbasierten Argumente, jedoch auch Argumente mit schwacher Evidenz, Gegenargumente und Ausnahmen. Auf diesem Wege wird es möglich, wissensbasierte Entscheidungen zur weiteren Ausgestaltung von Schulinspektionsverfahren zu treffen und das Verfahren so zu optimieren, dass es die beabsichtigten Wirkungen auch tatsächlich entfalten kann, also wirksam wird.

Literatur

- van Ackeren, I. & Klemm, K. (2009). *Entstehung, Struktur und Steuerung des deutschen Schulsystems*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Alkin, M. C. & Taut, S. M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, 29, 1–12.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

- Astbury, B. & Leeuw, F.L. (2010). Unpacking black boxes. Mechanisms and theory building in evaluation. *American Journal of Evaluation*, 31 (3), 363–381.
- Bachman, L.F. (2005). Building and Supporting a Case for Test Use. *Language Assessment Quarterly*, 2 (1), 1–34. doi:10.1207/s15434311laq0201_1
- Böttcher, W. & Kotthoff, H.-G. (2007). Gelingensbedingungen einer qualitätsoptimierenden Schulinspektion. In W. Böttcher & H.-G. Kotthoff (Hrsg.), *Schulinspektionen: Evaluation, Rechenschaftslegung und Qualitätsentwicklung* (S. 223–230). Münster: Waxmann.
- Böttcher, W. & Kotthoff, H.-G. (2010). Neue Formen der Schulinspektion: Wirkungshoffnungen und Wirksamkeit im Spiegel empirischer Bildungsforschung. In H. Altrichter & K. Maag Merki (Hrsg.), *Neue Steuerung im Schulwesen* (S. 295–325). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Brunsdon, V., Davies, M. & Shevlin, M. (2006). Anxiety and stress in educational professionals in relation to OFSTED. *Education Today*, 56 (1), 24–31.
- Chen, H.T. (2006). A theory-driven evaluation perspective on mixed methods research. *Research in the Schools*, 13 (1), 75–83.
- Coryn, C.L.S., Noakes, L.A., Westine, C.D. & Schröter, D.C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation*, 32 (2), 199–226.
- Cronbach, L.J., Ambron, S.R., Dornbusch, S.M., Hess, R.D., Hornik Phillips, D.C., Walker, D.F. & Weiner, S.S. (1980). *Toward reform of program evaluation*. San Francisco: Jossey-Bass.
- Dedering, K. (2012). Schulinspektion als wirksamer Weg der Systemsteuerung? *Zeitschrift für Pädagogik*, 58 (1), 69–88.
- de Wolf, I.F. & Janssens, F.J.G. (2007). Effects and side effects of inspection and accountability in education: on overview of empirical results. *Oxford Review of Education*, 33 (3), 379–396.
- Döbert, H., Rürup, M. & Dedering, K. (2008). Externe Evaluation von Schulen in Deutschland – die Konzepte der Bundesländer, ihre Gemeinsamkeiten und Unterschiede. In H. Döbert & K. Dedering (Hrsg.), *Externe Evaluation von Schulen. Historische, rechtliche und vergleichende Aspekte* (S. 63–152). Münster: Waxmann.
- Ehren, M.C.M., Leeuw, F.S. & Scheerens, J. (2005). On the impact of the Dutch educational supervision act: Analysing assumptions concerning the inspection of primary schools. *American Journal of Evaluation*, 26 (1), 60–76.
- Ehren, M.C.M. & Visscher, A.J. (2006). Towards a theory on the impact of school inspections. *British Journal of Educational Studies*, 54 (1), 51–72.
- Ehren, M.C.M. & Visscher, A.J. (2008). The relationships between school inspections, school characteristics and school improvement. *British Journal of Educational Studies*, 56 (2), 205–227.
- Herbert, J.L. (2014). Researching evaluation influence: A review of the literature. *Evaluation Review*, 38 (5), 388–419.

- Husfeldt, V. (2011). Wirkungen und Wirksamkeit der externen Schulevaluation: Überblick und Stand der Forschung. *Zeitschrift für Erziehungswissenschaft*, 14 (2), 259–283.
- Kane, M. T. (2002). Validating high-stakes testing programs. *Educational Measurement: Issues and Practice*, 21 (1), 31–41.
- Kane, M. T. (2006). In Praise of Pluralism. A Comment on Borsboom. *Psychometrika*, 71 (3), 441–445. Verfügbar unter: <http://doi.org/10.1007/s11336-006-1491-2> [Zugriff am 30.03.2015]
- Kane, M. T. (2013). Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement*, 50 (1), 1–73. Verfügbar unter: <http://doi.org/10.1111/jedm.12000> [Zugriff am 30.03.2015]
- King, J. A. & Pechman, E. M. (1984). Pinning a wave to the shore: Conceptualizing evaluation use in school systems. *Educational Evaluation and Policy Analysis*, 6 (3), 241–451.
- Kirkhart, K. E. (2000). Reconceptualising evaluation use: An integrated theory of influence. *New Directions for Evaluation*, 88, 5–23.
- Kluger, A. N. & DeNisi, A. S. (1996). The Effects of Feedback Interventions on Performance: Historical Review, a meta-analysis and a preliminary feedback intervention theory. *Psychological Bulletin*, 119 (2), 254–284.
- Köller, O. (2009). Evaluation pädagogisch-psychologischer Maßnahmen. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 331–351). Heidelberg: Springer.
- Luginbuhl, R., Webbink, D. & de Wolf, I. F. (2009). Do inspections improve primary school performance? *Educational Evaluation and Policy Analysis*, 31 (3), 221–237.
- Maritzen, N. (2009). Schulinspektion und Schulaufsicht. In H. Buchen & H.-G. Rolff (Hrsg.), *Professionswissen Schulleitung* (S. 1368–1392). Weinheim: Beltz.
- Messick, S. (1998). Test Validity: A Matter of Consequence. *Social Indicators Research*, 45 (1/3), 35–44. Verfügbar unter: <http://doi.org/10.1023/A:1006964925094> [Zugriff am 30.03.2015]
- Mislevy, R. J. (2003). Substance and structure in assessment arguments. *Law, Probability, and Risk*, 2, 237–258.
- Müller, S. (2010). Erste Effekte von Schulinspektion – eine Zwischenbilanz. In N. Berke-meyer, W. Bos, W., H. G. Holtappels, N. McElvany & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung Band 16. Daten, Beispiele und Perspektiven* (S. 289–308). Weinheim: Juventa.
- Nutley, S., Walter, I. & Davies, H. T. O. (2003). From Knowing to doing: A framework for understanding the evidence-into-practice agenda. *Evaluation*, 9 (2), 125–148.
- Pietsch, M., Janke, N. & Mohr, I. (2013). Führt Schulinspektion wirklich nicht zu besseren Schülerleistungen? Eine Einschätzung zur Belastbarkeit vorliegender Wirksamkeitsstudien aus programmtheoretischer Perspektive. In K. Schwippert, M. Bonsen & N. Berke-meyer (Hrsg.), *Schul- und Bildungsforschung. Diskussionen, Befunde und Perspektiven* (S. 167–185). Münster: Waxmann.
- Pietsch, M., Janke, N. & Mohr, I. (2014). Führt Schulinspektion zu besseren Schülerleistungen? Difference-in-Differences-Studien zu Effekten der Schulinspektion

- Hamburg auf Lernzuwächse und Leistungstrends. *Zeitschrift für Pädagogik*, 60 (3), 446–470.
- Pietsch, M., Schnack, J. & Schulze, P. (2009). Unterricht zielgerichtet entwickeln: Die Schulinspektion Hamburg entwickelt ein Stufenmodell für die Qualität von Unterricht. *Pädagogik*, 61 (2), 38–43.
- Pietsch, M., Schulze, P., Schnack, J. & Krause, M. (2011). Elaborierte Rückmeldungen zur Qualität von Unterricht. Über empirisch abgesicherte Bezugsnormen als Grundlage für die Weiterentwicklung von Unterricht und Schule. In S. Müller, M. Pietsch & W. Bos (Hrsg.), *Schulinspektionen in Deutschland – Eine Zwischenbilanz aus empirischer Sicht* (S. 193–216). Münster: Waxmann.
- Ramaprasad, A. (1983). On the Definition of Feedback. *Behavioral Science*, 28 (1), 4–13.
- Scheerens, J. (2007). The case of evaluation and accountability provisions in education as an area for the development of policy malleable system indicators. In H.-H. Krüger, T. Rauschenbach, U. Sander (Hrsg.), *Bildungs- und Sozialberichterstattung* (Zeitschrift für Erziehungswissenschaft: Sonderheft 6, S. 207–224). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Scriven, M. (2012). *Evaluating Evaluations: A Meta-Evaluation Checklist*. Verfügbar unter: http://michaelscriven.info/images/EVALUATING_EVALUATIONS_8.1.11.pdf [Zugriff am 30.03.2015]
- Shaw, S. & Crisp, V. (2012). An approach to validation: Developing and applying an approach for the validation of general qualifications. *Research Matters: A Cambridge Assessment Publication* (Special Issue 3). Verfügbar unter: <http://www.cambridgeassessment.org.uk/images/110003-research-matters-special-issue-3-an-approach-to-validation.pdf> [Zugriff am 30.03.2015]
- Stame, N. (2010). What doesn't work? Three failures, many answers. *Evaluation*, 16 (4), 371–387.
- Stamm, M. (2003). Evaluation im Spiegel ihrer Nutzung: Grande idée oder grande illusion des 21. Jahrhunderts? *Zeitschrift für Evaluation*, 2 (2), 183–200.
- Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.
- Toulmin, S. (2003). *The uses of argument* (aktualisierte Aufl.). Cambridge: Cambridge University Press.
- Visscher, A.J. & Coe, R. (2003). School Performance Feedback Systems: Conceptualisation, Analysis and Reflection. *School Effectiveness and School Improvement*, 14 (3), 321–349.
- Weber, K. (2006). From nuts and bolts to toolkits: theorizing with mechanisms. *Journal of Management Inquiry*, 15 (2), 119–123.
- Weiss, C.H. (1980). Knowledge creep and decision accretion. *Knowledge: Creation, Diffusion, Utilization*, 1 (3), 381–404.
- Weiss, C.H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19 (1), 21–33.

Expertenurteile – Achillesferse oder Trumpf der Schulinspektion?

Moritz G. Sowada

Zusammenfassung

Die Bewertung der Qualität von Schulen durch Inspektionsteams ist ein konstitutives Element von Schulinspektionsverfahren. Dieser Beitrag gibt einen Überblick über einschlägige Forschungsliteratur zu Qualitätsurteilen im Rahmen von Schulinspektion. Es wird dabei deutlich, dass erkenntnistheoretische Annahmen in die Forschung mit einfließen, die beeinflussen, ob Differenzen in den Urteilen von Inspektorinnen und Inspektoren per se als problematisch gelten. Anschließend wird gezeigt, welche innovativen Ansätze zur Erforschung von Urteilen, Entscheidungen und Bewertungen in der Psychologie und Soziologie verwandt werden und welchen Mehrwert deren Adaption für die Erforschung von Expertenurteilen in der Schulinspektion verspricht. Schließlich werden unterschiedliche Optionen der Professionalisierung von Expertenurteilen angedeutet.

1. Einleitung

Es ist die Aufgabe von Schulinspektionsteams, die Qualität von Schulen auf Basis von standardisierten Kriterien und Indikatoren zu bewerten. Für diese voraussetzungsvolle Aufgabe wird meist qualifiziertes Personal eingesetzt. So bestehen z. B. die Inspektionsteams in Hamburg normalerweise aus zwei hauptamtlichen Inspektoren bzw. Inspektorinnen, einem Leitungsmitglied einer anderen Schule der gleichen Schulform (sogenannte Schulformexperten bzw. Schulformexpertinnen) und – an beruflichen Schulen – noch einer weiteren Person als Vertretung der Wirtschaft (vgl. IfBQ 2011, S. 6; Döbert et al. 2008, S. 87–91).

Die fachliche Qualifikation und die berufliche Erfahrung der Inspektorinnen und Inspektoren sind bedeutsam. Sie stellen Ressourcen dar, die kompetente Expertenurteile ermöglichen. Die Berufserfahrung der Hamburger Inspektorinnen und Inspektoren wird wie folgt charakterisiert: „Von den 15 Personen, die Inspektionen hauptamtlich durchführten, haben neun ein Lehramtsstudium absolviert mit einer schulischen Unterrichtserfahrung zwischen 3 und 17 Jahren. Die übrigen Inspektorinnen und Inspektoren kommen aus geistes- und sozialwissenschaftlichen Berufen. Alle verbindet, dass sie über eine vertiefte Kenntnis sowohl von Schule als auch von Qualitätssicherung und Qualitätsentwicklung verfügen“ (IfBQ 2012,

S. 11). Sie erhalten zudem vor Übernahme von Teamleitungen eine etwa viermonatige Zusatzausbildung (vgl. IfBQ 2014). Hauptamtliche Inspektorinnen und Inspektoren erwerben darüber hinaus kontinuierlich berufspraktische Erfahrung in der Beurteilung von Schulen im Laufe ihrer Inspektionstätigkeit. Sie können daher als Expertinnen und Experten für die externe Evaluation von Schulen gelten.

Die Einschätzung und Bewertung von Schule und Unterricht anhand vorab definierter Qualitätsmerkmale ist der Kern der Expertise von Inspektorinnen und Inspektoren und steht damit im Zentrum ihrer beruflichen Tätigkeit. Trotz der zentralen Funktion des Bewertens im Prozess der Schulinspektion liegen vergleichsweise wenig gesicherte Erkenntnisse darüber vor, wie Inspektorinnen und Inspektoren Bewertungen vornehmen. Und dies, obwohl der bildungspolitischen Erwartung, dass Schulen aus den Inspektionsergebnissen Schul- und Unterrichtsentwicklungsmaßnahmen ableiten, diese durchführen und schließlich evaluieren sollen (vgl. Altrichter et al. 2014; Sommer 2011). So gibt es zwar präskriptive Vorgaben dazu, wie Schulinspektionen durchgeführt werden sollen, diese sind jedoch als Richtlinien zu verstehen (vgl. Nutley et al. 2012). Solche Bewertungsvorgaben sind unterdeterminiert, d. h. sie können nur in eingeschränktem Maße vorwegnehmen, wie Inspektorinnen und Inspektoren in der Praxis tatsächlich bewerten (vgl. Sadler 2009, 2014). Bewertungsrichtlinien lassen daher nur begrenzt Schlüsse auf tatsächliche Bewertungspraxen zu.

In der Regel vollziehen Inspektorinnen und Inspektoren ihre Bewertungsarbeit unter restriktiven Rahmenbedingungen. Sie müssen mit begrenzten Personalressourcen und unter Zeitdruck umfangreiche Daten erheben, analysieren und ggf. synthetisieren, um eine Vielzahl an Qualitätskriterien zeitnah zu bewerten (vgl. Gray & Wilcox 1995, S. 139). Wie Inspektorinnen und Inspektoren dies vollbringen, ist weitgehend ungeklärt. Dennoch wird die Expertise von Inspektorinnen und Inspektoren von Boyne, Day und Walker (2002) als zentrale Variable für die Kosten-Nutzen-Relation von Inspektionen postuliert – ohne Expertise jedoch inhaltlich zu bestimmen. Sie operieren mit folgender Annahme: Je höher die Expertise der Inspektorinnen und Inspektoren ist, desto größer ist der Nutzen von Inspektionen und desto geringer sind die damit verbundenen direkten und indirekten Kosten.

Anstatt den Prozess des Bewertens zu untersuchen, hat die empirische Bildungsforschung sich dem Forschungsfeld Schulinspektion bisher hauptsächlich über die Wirkungsfrage genähert (vgl. Dederling 2012; Husfeldt 2011; Kotthoff & Böttcher 2010). Es wird untersucht, was im Zusammenhang mit Schulinspektion an Schulen passiert. Dabei wird oft die Sicht der Schulleitung und manchmal auch der Lehrkräfte abgefragt (vgl. Lambrecht & Rürup 2012, S. 65). Teilweise wird zwischen intendierten und nicht intendierten sowie erwünschten und unerwünschten Wirkungen unterschieden (vgl. Ehren & Visscher 2006). Auch international gibt es kaum Wissen dazu, was eine „wirksame“ Schulinspektion ausmacht. Es ist jedoch bekannt, dass groß angelegte Reformmaßnahmen im Bildungswesen ca.

sieben Jahre bedürfen, bevor entsprechende Wirkungen verlässlich nachweisbar sind (vgl. Borman et al. 2003). Die vorgelagerte Frage, wie Schulinspektionsergebnisse zustande kommen, wird hingegen in der empirischen Forschung weitgehend ausgeblendet (vgl. Lambrecht & Rürup 2012, S. 72). Dieses Desiderat ist Gegenstand des vorliegenden Textes.

Es gilt grundsätzlich zu beachten, dass mit Schulinspektion unterschiedliche Funktionen erfüllt werden sollen (vgl. Maritzen in diesem Band). Diese Funktionen können in einem Spannungsverhältnis zueinander stehen, so z. B. die Kontroll- und Entwicklungsfunktion (vgl. Böttcher & Keune 2012). Dieses Spannungsverhältnis äußert sich für die Inspektorinnen und Inspektoren in einem „Rollenkonflikt zwischen objektivem Beobachter und Akteur, der über schulisches Wissen verfügt, aber nicht beraten darf“ (Preuß et al. 2015, S. 136). Je nachdem, ob die Kontroll- oder Entwicklungsfunktion hervorgehoben wird, sind unterschiedliche Entwicklungsstrategien zu favorisieren (vgl. hierzu Abschnitt 5).

Im Folgenden werden einschlägige Forschungsbefunde zu Bewertungen und Bewertungsprozessen von Inspektorinnen und Inspektoren vorgestellt. Zuerst werden kritische Anfragen hinsichtlich der Methodologie und Methoden von Schulinspektion referiert. Im Anschluss daran wird auf Ergebnisse aus Inter-Beurteiler-Studien eingegangen und diese daraufhin um Befunde qualitativer Forschung ergänzt. Im dritten Abschnitt werden Heuristiken als vielversprechende Option zur Erforschung von Schulinspektionsurteilen eingeführt. Kontrastierend dazu wird im vierten Abschnitt die Analyse von Wertordnungen thematisiert, die Bewertungsdifferenzen zugrunde liegen können. Im fünften und letzten Abschnitt werden die Potenziale der angerissenen Ansätze für zukünftige Bewertungsforschung im Kontext von Schulinspektion erörtert.

2. Urteile von Inspektorinnen und Inspektoren

An dieser Stelle erfolgt eine knappe Bestandsaufnahme der Literatur, die bisher die Bewertung von Inspektorinnen und Inspektoren explizit zum Gegenstand gemacht hat. Es gibt zum einen psychometrische Ansätze, in denen die Güte der vorgenommenen Bewertungen von Inspektorinnen und Inspektoren Analysegegenstand ist. Geprüft wird mithilfe statistischer Verfahren, inwieweit klassische wissenschaftliche Gütekriterien wie Objektivität, Reliabilität und Validität erfüllt werden. Entsprechende Erhebungen basieren meist auf der Bewertung ein und derselben Unterrichtssequenz durch mehrere Inspektorinnen und Inspektoren auf Basis eines standardisierten Unterrichtsbeobachtungsbogens. Die Annahme ist zumeist, dass je größer die Übereinstimmung zwischen den bewertenden Personen ist, desto verlässlicher sei das Bewertungsergebnis. Die zugrunde liegende erkenntnistheoretische Prämisse beinhaltet, dass es eine – und nur eine – korrekte Bewertung

gibt. Abweichungen von diesem hypothetisch „wahren“ Wert seien daher als problematisch einzustufen und sollten dementsprechend soweit wie möglich reduziert werden.

Andererseits gibt es qualitativ orientierte Forschungsberichte, die grundlegend infrage stellen, ob eine umfassende Standardisierung des Bewertungsverhaltens von Inspektorinnen und Inspektoren erstrebenswert oder gar möglich ist. Gerade bei so komplexen Phänomenen wie Schul- und Unterrichtsqualität gäbe es Ermessensspielräume, deren Vereindeutigung sachunangemessen sei. Zur Stützung dieser Position wird auf unterschiedliche Argumente zurückgegriffen, die unten weiter ausgeführt werden. Zuvor werden jedoch empirische Befunde aus dem psychometrischen Forschungsansatz angeführt.

2.1 Psychometrische Forschung

Wie erwähnt zielt psychometrische Forschung u. a. darauf ab, zu überprüfen, in welchem Maße die klassischen Gütekriterien seitens des Inspektionsverfahrens abgesichert und durch Inspektorinnen und Inspektoren erfüllt werden. Dies impliziert die Annahme, dass solche wissenschaftlichen Gütekriterien gegenstandsangemessen sind. Gärtner (2013) attestiert Schulinspektionssystemen in diesem Zusammenhang eine „positivistische [...] Grundorientierung“ (S. 698). Schulinspektion selbst wird – allerdings in unterschiedlichem Maße – an dem Ideal der evidenz- bzw. wissenschaftsbasierten Evaluation von Schulen ausgerichtet (vgl. Rürup 2008). Es sollte daher nicht verwundern, wenn Schulinspektion an diesen Ansprüchen gemessen und in der Folge dann kritisiert wird. Andererseits identifiziert Rürup (2008) neben einer wissenschaftlichen Orientierung von Schulinspektionssystemen auch eine konkurrierende, da den *objektiven* Vergleich erschwerende Einzelfallorientierung. Das Verhältnis von Standardisierung und Fallorientierung fällt dabei zwischen den Inspektionsverfahren in den einzelnen Bundesländern unterschiedlich aus.

2.1.1 Methodisch-methodologische Kritik an Schulinspektionsverfahren

In England sind mehrere Aufsätze veröffentlicht worden, die kritische Fragen bezüglich der methodischen bzw. methodologischen Güte von Schulinspektion formulieren (vgl. Fitz-Gibbon & Stephenson-Forster 1999; Gray & Wilcox 1995) und dabei z. T. Forschungsstandards als Referenzpunkt heranziehen. Ähnlich gelagerte Kritik wird auch gegenüber der Inspektion der ersten Phase der Lehrerbildung (*Initial Teacher Training*) an Hochschulen in England geäußert (vgl. Cale & Harris 2003; Campbell & Husbands 2000; Sinkinson & Jones 2001). Problematisiert wird

u. a. die Auswahl von zu beobachtenden Unterrichtssequenzen und damit deren Repräsentativität für die Unterrichtsqualität an einer Schule (vgl. Wilcox & Gray 1996, S. 70 f.). Ebenfalls wird angezweifelt, ob Schulinspektion die Reliabilität und Validität u. a. von Unterrichtsbewertungen gewährleisten kann (vgl. Fitz-Gibbon & Stephenson-Forster 1999).

Auch im deutschsprachigen Kontext werden die methodisch-methodologischen Herausforderungen, insbesondere im Zusammenhang mit Unterrichtsbewertungen, thematisiert. So wird etwa der mit 20 Minuten relativ kurze Beobachtungszeitraum für die Unterrichtsbewertungen problematisiert (vgl. Bos et al. 2006). Zu dieser begrenzten Beobachtungsdauer kommt hinzu, dass Inspektorinnen und Inspektoren mitunter relevante Kontextinformationen fehlen und die Unterrichtsbewertung selbst die Beobachtungssituation verfälschen könnte (vgl. Praetorius 2013). Nicht zuletzt werden Beurteilereffekte als Gefährdung der Aussagekraft des Ergebnisses sowie deren Nutzens angesehen (vgl. Pietsch & Tosana 2008).

2.1.2 Inter-Beurteiler-Studien zur Unterrichtsbewertung

Die Studie von Matthews, Holmes, Vickers und Corporaal (1998) kann als Reaktion auf die methodisch-methodologische Kritik an dem Verfahren des englischen Schulinspektorats Ofsted (*Office for Standards in Education*) gedeutet werden. Sie haben die Reliabilität und Validität von Unterrichtsbewertung durch jeweils zwei Inspektorinnen bzw. Inspektoren untersucht. Die freiwillig teilnehmenden Inspektorinnen und Inspektoren haben ihre Bewertung von 173 gemeinsam beobachteten Unterrichtssequenzen getrennt voneinander festgehalten. Es wurde für jede Unterrichtssequenz jeweils eine Bewertungskategorie von 1 bis 7 vergeben. In 66 Prozent aller Fälle wurde von dem jeweiligen Bewertungspaar die gleiche Bewertungskategorie gewählt (vgl. Matthews et al. 1998, S. 182). Somit besteht ein Bewertungsdissens in einem Drittel aller untersuchten Fälle, der allerdings im kritischen Grenzbereich von *satisfactory* zu *unsatisfactory* (Bewertungsstufen 4 und 5) höher ausfällt (vgl. ebd., S. 184).

Auch für Deutschland liegen erste Inter-Beurteiler-Studien mit Inspektorinnen und Inspektoren vor. Pietsch und Tosana (2008) konnten in ihrer Hamburger Studie feststellen, dass „nur geringe Urteilsfehler nachweisbar sind (rund 9 Prozent der Gesamtvariation), wobei diese vor allem auf die unterschiedliche Interpretation von Items durch die bewertenden Unterrichtsbeobachter zurückzuführen sind“ (S. 449). Sie verweisen aber auch darauf, dass die Inspektorinnen und Inspektoren bei der Unterrichtsbewertung unterschiedliche Anspruchsniveaus anlegen, d. h., einige urteilen strenger als andere. In einer Nachfolgestudie wurde jedoch herausgestellt, dass diese Strengeeffekte einen geringeren Anteil an der Beurteilervarianz haben als die variierende Interpretation der Items des Unterrichtsbewertungs-

bogens durch die Inspektorinnen und Inspektoren (vgl. Pietsch 2010, S. 128). In einer Vergleichsstudie haben Müller und Pietsch (2011) Inspektorinnen und Inspektoren aus sechs Bundesländern videografierte Unterrichtssequenzen mit verschiedenen Unterrichtsbewertungsbögen beurteilen lassen. Sie stellten dabei eine zum Teil sehr hohe Beurteilervarianz von bis zu 71 Prozent fest (vgl. Müller & Pietsch 2011, S. 45). Dennoch seien aggregierte Aussagen zur Unterrichtsqualität auf Schulebene, laut den Ergebnissen der Studie, statistisch zulässig (vgl. ebd., S. 51).

Die in der Literatur aufgeworfene methodisch-methodologische Kritik wurde also im Rahmen von Inter-Beurteiler-Studien adressiert. Diese empirischen Studien weisen Beurteilervarianz nach, wobei das Ausmaß der den Inspektorinnen und Inspektoren zugerechneten Varianz selbst eine große Spannbreite aufweist. Über die Gründe für die festgestellte Beurteilervarianz können die Studien hingegen keine Auskunft geben, da der Prozess der Beurteilung und die Gründe der Inspektorinnen und Inspektoren für die Wahl einer Bewertungskategorie methodisch ausgeblendet werden. An dieser Stelle können Interview-, *Think-Aloud*- und ethnographische Studien weiterführen. Auf entsprechende qualitative Forschungsstränge wird daher nun auch eingegangen.

2.2 Qualitativ-orientierte Forschung

Qualitative Studien, die die Bewertungsarbeit von Inspektorinnen und Inspektoren in den Blick nehmen, sind relativ dünn gesät. Entsprechende Studien haben zu meist ein dezidiert *nicht* positivistisches Verständnis zugrunde gelegt (vgl. Lincoln et al. 2011). Dies hat Auswirkungen darauf, wie Schulinspektion als Forschungsgegenstand verstanden wird. So ist das Forschungsinteresse nicht darauf gerichtet, wie etwa die „tatsächliche Qualität“ von Schule und Unterricht mit optimierten Instrumenten und Verfahren besser erfasst werden kann. Dies deswegen, weil in entsprechenden Paradigmen nicht von *einer* singulären, objektiv erfassbaren Qualität ausgegangen wird. Stattdessen liege Qualität im Auge des Betrachters, ohne jedoch deswegen beliebig zu sein (vgl. Heid 2013). Vorstellungen von guter Schule und gutem Unterricht sind in soziohistorische Kontexte eingebunden, die Qualitätszuschreibungen stabilisieren. Qualität ist in einer solchen Perspektive Aushandlungsgegenstand zwischen den Stakeholdern. Wenn es also nicht darum geht, der *einen* objektiven Qualität so nah wie möglich zu kommen, worum geht es dann in qualitativer Forschung zu Bewertungen im Kontext von Schulinspektion? Um hierauf eine Antwort zu geben, werden ausgewählte qualitative Studien nachfolgend skizziert.

Interviewstudien sind eine häufige Variante qualitativer Schulinspektionsforschung, da sie vergleichsweise einfach zu realisieren sind. In Interviewstudien ge-

ben Inspektorinnen und Inspektoren selbst Auskunft zu ihrer Bewertungstätigkeit. Sie können dabei als Expertinnen und Experten verstanden werden, die direkt und informativ zu ihren Bewertungspraxen relevante Informationen angeben können. Andererseits kann im Reden über Schulinspektion mehr mitgeteilt werden, als den Akteuren in der Situation unbedingt bewusst ist.

2.2.1 *Expertenurteile als Professional Judgements*

Nixon und Rudduck (1993) beschäftigt die Frage, wie Expertenurteile (*Professional Judgements*) aus Sicht von Inspektorinnen und Inspektoren zustande kommen. Die Einführung neuer Schulinspektionsverfahren unter Ofsted wurde als Wandel von impliziten zu expliziten Bewertungskriterien beschrieben. Für Nixon und Rudduck (ebd.) greift diese Charakterisierung zu kurz. Sie verweisen zum einen in einer philosophisch gestützten Argumentation darauf, dass die explizit benannten Kriterien den Bewertungsprozess zwar anleiten, aber nicht determinieren können. Die Kriterien haben vielmehr eine orientierende Funktion für einzelne Inspektorinnen und Inspektoren, Inspektionsteams und Schulen (vgl. ebd., S. 140 f.).

Ihre Interviewbefunde verweisen auf eine Differenz zwischen den Selbstaussagen der Inspektorinnen und Inspektoren zu ihren Urteilspraxen und einer kriterienbasierten Bewertungslogik des offiziellen Schulinspektionsverfahrens. Inspektorinnen und Inspektoren würden Such- und Bewertungsprozesse unter Zuhilfenahme ihres impliziten Wissens bewerkstelligen (vgl. ebd., S. 143). Dabei verschiebt sich der Fokus der Aufmerksamkeit situationsspezifisch, z. B. von der Erfüllung von binär kodierten Vorschriften hin zur Einschätzung von Qualität im Gesamtzusammenhang (vgl. ebd., S. 144). Hieran thematisch anschließend besteht für Meyer (2012, S. 46–49) die Inspektionstätigkeit nicht allein aus der Addition von beobachteten Fakten zu einer Gesamtbewertung, auch die Wertung einzelner Elemente basiere auf der Gesamtbetrachtung einer Schule. Nur so sei eine angemessene Bewertung von komplexen Qualitätsaspekten überhaupt möglich. Expertenurteile von Inspektorinnen und Inspektoren beruhen demnach nicht nur auf der Bewertung von voneinander isolierten Fakten, sondern gerade bei komplexeren Bewertungsobjekten auf einem hermeneutischen Bewertungsprozess.

Wilson (1996) verweist auf weitere Aspekte, die die Komplexität von Expertenurteilen verdeutlichen. Seine zehnmonatige Feldstudie aus dem Jahre 1992 fällt in eine Zeit, in der das englische Schulinspektionswesen tief greifend umstrukturiert wurde. Die Reform kann als eine Umstellung von einem *Connaissanceur*-Modell, das auf Expertise und Kennerschaft beruhte (vgl. Eisner 2013), hin zu einem sozialwissenschaftlichen Inspektionsverfahren, das auf Standardisierung und expliziten Kriterien basiert, verstanden werden (vgl. Scheerens et al. 2003, S. 44 f.). Gleichwohl, so Wilson (1996, S. 122), seien Expertenurteile auch weiterhin erfahrungs-

gebunden und abhängig von unbewussten kognitiven Prozessen. Die Gründe für ein Expertenurteil sind daher nur begrenzt explizierbar. Gerade das Praxiswissen (*Practitioner Knowledge*) von Inspektorinnen und Inspektoren ermögliche kompetente Expertenurteile. Kriterienkataloge sind zwar standardisiert, aber deren fallbezogene Anwendung erfordere ein gewisses Maß an Intuition und Expertise (vgl. Nutley et al. 2012, S. 878). Dass dies potenziell zu unterschiedlichen Urteilen führen könne, sei keine Schwäche, sondern eine Stärke von Schulinspektion, so Wilson (1996). In dem Verfahren ginge es nicht darum, absolute, sondern nur relationale Urteile zu fällen, die zur Auseinandersetzung mit ihnen innerhalb einer Schulgemeinschaft anregen sollten. Wilson (ebd.) betont primär die Entwicklungsfunktion von Schulinspektion.

Inspektorinnen und Inspektoren seien darum bemüht, ihre Qualitätseinschätzungen mit Evidenzen abzusichern und ihnen dadurch Gewicht zu verleihen. Das Verhältnis von Evidenz und Expertenurteil sei von Wechselwirkungen geprägt. Es werden demnach nicht erst Evidenzen systematisch gesammelt und dann abschließend beurteilt, sondern Bewertung und Datensammlung finden während der gesamten Schulinspektion parallel statt. So beeinflussen vorgängige Einschätzungen von Inspektorinnen und Inspektoren, welche Daten zu Evidenzen werden (vgl. Wilson 1996, S. 205). Dabei gehe es darum, die vorläufigen, aber dennoch aufmerksamkeitsstrukturierenden Einschätzungen und Annahmen mit Evidenzen abzusichern. Die Suche nach konträren Evidenzen kommt hingegen seltener zum Einsatz, da sie eine aufwendigere Prüfmethode darstellt (vgl. Nutley et al. 2012, S. 878).

Dass diese Erkenntnisse vor dem Hintergrund eines positivistischen Paradigmas problematisch sind, zeigen folgende Studien.

2.2.2 Expertenurteile zwischen Programmatik und Pragmatik

Millett und Johnson (1998) verweisen darauf, dass Ofsted von den Inspektorinnen und Inspektoren erwartet, dass sie die Effektivität von Unterricht bewerten und dabei ihre persönlichen Präferenzen für bestimmte Methoden und Unterrichtsstile außen vor lassen. Etwaige Präferenzen für einen bestimmten pädagogischen Stil werden dementsprechend abgewertet, wohingegen fachliche Expertise in z.B. Mathematik im Bewertungsprozess ausdrücklich gewünscht sei. Solche politisch motivierten Vorgaben operieren mit der Prämisse, dass es möglich sei, bestimmte Formen von Vorwissen zu exkludieren, andere hingegen zu inkludieren. Die interviewten Inspektorinnen und Inspektoren geben an, dass sie versuchen, solchen Vorgaben zu entsprechen. Gleichwohl scheint es, dass sie dies auf unterschiedliche Art und Weise tun. Solche abstrakten Vorgaben werden rekontextualisiert, d. h. mit unterschiedlichem Sinn versehen und darüber in das eigene Aufmerksamkeits- und Urteilsverhalten integriert (vgl. Fend 2008). Mathematisch stärker vorgebildete In-

spektorinnen und Inspektoren bewerten, so mutmaßen Millett und Johnson (1998), eher die fachliche Qualität des Unterrichts, wohingegen andere stärker die allgemeine Unterrichtsqualität in den Blick nehmen.

Auch Lindgren (2015) nimmt das Verhältnis von Bewertungsvorgaben und Urteilspraxen in den Blick. Er benutzt hierfür Goffmans konzeptuelle Unterscheidung zwischen Vorder- und Hinterbühne (Goffman 1959). Auf der Vorderbühne werde von Inspektionsteams angestrebt, einheitlich gegenüber der Schulöffentlichkeit aufzutreten. Es werde eine konsistente Ergebnisdarstellung angestrebt. Dabei gelte es, die formale Übereinstimmung mit den offiziellen Verfahrensvorgaben sicherzustellen. Zum anderen werde darauf hingewirkt, das Ergebnis einheitlich als Team zu vertreten. Bevor ein Schulinspektionsbericht als Produkt auf der Vorderbühne bekannt gegeben wird, wird auf der Hinterbühne auf vielfältige Art und Weise darauf hingearbeitet. Lindgren kann zeigen, dass auf der Hinterbühne durchaus methodologische Unsicherheiten, z. B. bzgl. der Herstellung von Urteilskonsistenz, vorherrschen. Solche Interna werden gegenüber Außenstehenden jedoch nicht kommuniziert. Divergente Einschätzungen von Qualitätsaspekten werden vielmehr in geschützten Gesprächen vereinheitlicht. Zudem werden negative Bewertungen eingehender geprüft und aufwendiger abgesichert als positive Urteile (siehe auch Nutley et al. 2012, S. 879). Dabei gehe es *nicht nur* um das richtige, da evidenzbasierte Expertenurteil, sondern auch um Fragen der antizipierten Ergebnisrezeption aufseiten der Schule.

2.3 Zwischenfazit

Die angeführten quantitativen und qualitativen Studien zeigen, dass Expertenurteile in sehr unterschiedlicher Weise zum Gegenstand von Forschung gemacht werden. Wo quantitative Studien eine Komplexitätsreduktion vornehmen, indem mittels Inter-Beurteiler-Studien Fragen der Konsistenz bzw. Variation von Bewertungen unter Absehung des Bewertungsprozesses untersucht werden, wird in qualitativen Studien oftmals Komplexität gesteigert. Expertenurteile werden dabei nicht allein als technisch-methodische, sondern als mikropolitisch machtbasierte Vorgänge konzeptualisiert. Durch diese andere Perspektive kommen Aspekte jenseits der Konsistenzfrage in den Blick.

Die bisherige Forschung zum Thema, auf welche Weise und wie konsistent Inspektorinnen und Inspektoren bewerten, ist erst in den Anfängen begriffen. Da evidenzbasierte Expertenurteile gleichwohl den Kern der Arbeit von Inspektorinnen und Inspektoren darstellen, werden nachfolgend vielversprechende Forschungsansätze in benachbarten Forschungsfeldern vorgestellt, die dazu geeignet sind, zukünftige Forschung zu Schulinspektionsurteilen in neue und innovative Richtungen zu führen.

3. Psychologie des Urteilens und Entscheidens

In den letzten Jahrzehnten sind bedeutsame Fortschritte in der psychologischen Urteils- und Entscheidungsforschung erzielt worden. Ausgangspunkt war die von Simon (2000) Mitte der 1950er popularisierte Einsicht, dass Menschen nur begrenzt rational (*Bounded Rationality*) handeln können (vgl. Gigerenzer & Selten 2002). Aufgrund begrenzter kognitiver Kapazitäten, endlicher Zeit und unzuverlässiger Informationen in der Umwelt ist es Menschen nicht möglich, optimale rationale Urteile und Entscheidungen zu treffen. Diese limitierenden Bedingungen gelten auch für die Bewertungsarbeit von Inspektorinnen und Inspektoren. Hierzu liegen jedoch bisher keine empirischen Forschungsergebnisse vor.

In der Psychologie gibt es hingegen eine Vielzahl an Forschungsbefunden, die zeigen, dass Menschen bei ihren Urteilen und Entscheidungen Heuristiken verwenden (vgl. Gigerenzer & Gaissmaier 2011). Heuristiken sind kognitive Prozesse, mit denen Menschen viable Urteile und Entscheidungen *unter Unsicherheit* treffen. Heuristiken können bewusst oder unbewusst ablaufen. Ein Vorteil von Heuristiken ist, dass sie in komplexen Situationen effiziente Urteils- und Entscheidungsfindung ermöglichen, indem gerade nicht alle verfügbaren Informationen berücksichtigt werden. Andererseits können diese Abkürzungsstrategien potenziell zu systematischen Urteilsverzerrungen führen (vgl. Tversky & Kahneman 1974). Es ist zu erwarten, dass Inspektorinnen und Inspektoren Heuristiken bereichsspezifisch einsetzen oder gar aufgabenspezifische Heuristiken verwenden.

Es gibt in der Psychologie des Urteilens und Entscheidens zwei Forschungsstränge, in denen Heuristiken bedeutsam sind, die jedoch mit unterschiedlichen Prämissen und Forschungsvorgehen arbeiten (vgl. Hammond 1996). Sie werden nachfolgend vorgestellt.

3.1 *Judgment and Decision-Making*

Der Ansatz des *Judgment and Decision-Making* (JDM) nutzt vermeintlich einfache Aufgabenstellungen zum Vergleich menschlicher Urteile und Entscheidungen mit „korrekten“ Urteilen und Entscheidungen, die mittels formaler Logik abgeleitet werden und somit dem Ideal der Rationalität entsprechen. Rationalität wird diesem Ansatz gemäß als Kohärenz zwischen logischen und empirischen Urteilen und Entscheidungen gefasst. In diesem Ansatz erscheinen menschliche Urteile und Entscheidungen als defizitär und imperfekt, da postulierte Rationalitätsmaßstäbe insofern verletzt werden, dass tatsächliches Urteils- und Entscheidungsverhalten unter bestimmten reproduzierbaren Umständen von den abgeleiteten „rationalen Urteilen“ abweichen. Es handelt sich hierbei daher um einen präskriptiv-norma-

tiven Ansatz, der einen Beitrag dazu leisten soll, Entscheidungen und Urteile auf eine rationalere Basis zu stellen, um sie auf diese Weise zu optimieren (vgl. Kundin 2010, S. 349).

Es wird mittlerweile in einer Vielzahl kognitions- und sozialpsychologischer Modelle davon ausgegangen, dass das menschliche Denken auf zwei separaten Prozessen beruht (vgl. Evans 2008). Zum einen wird angenommen, dass es schnelle und weitgehend unbewusste Kognitionen gibt, die u. a. auf implizitem Wissen und impliziten Erfahrungen beruhen (System 1). Zum anderen gebe es langsamere, dafür aber bewusster ablaufende Kognitionen, die auf explizitem Wissen basieren (System 2). Für Expertenurteile von Inspektorinnen und Inspektoren ist daraus abzuleiten, dass beide Systeme in die Bewertungsarbeit mit einfließen. Das Verhältnis von beiden Systemen bei der Auswahl von Bewertungskategorien ist jedoch ungeklärt. Es ist aber davon auszugehen, dass Inspektorinnen und Inspektoren, um effiziente Bewertungsarbeit zu leisten, Kognitionen des Systems 1 bedürfen. Dieser Form der Kognition wird im folgenden Ansatz eine größere funktionale Bedeutung zugesprochen, als dies im JDM-Ansatz der Fall ist.

3.2 *Naturalistic Decision-Making*

Dieser zweite Ansatz fokussiert u. a., welche Rolle Expertise in realen Situationen spielt. Dieser Ansatz wird daher *Naturalistic Decision-Making* (NDM) genannt (vgl. Klein 2008). Hier interessiert u. a., inwiefern Experten sich von Novizen in ihrem Urteils- und Entscheidungsverhalten unter natürlichen Bedingungen unterscheiden. In einer mittlerweile klassischen Studie wurde untersucht, wie Einsatzleiter bei der Feuerwehr ihre Entscheidungen treffen (vgl. Klein et al. 2010). Die Studie zeigte, dass im Ernstfall nicht mehrere Handlungsalternativen entworfen und gegeneinander abgewogen werden, um die beste Handlungsoption zu identifizieren. Stattdessen haben die Einsatzleiter nur einen Lösungsansatz entworfen und bei positiver Ad-hoc-Evaluation umgesetzt.

Der NDM-Ansatz ist für die Forschung zu Urteilen von Inspektorinnen und Inspektoren insofern interessant, als er danach fragt, wie Akteure unter Bedingungen begrenzter Rationalität dennoch situationsadäquate Urteile und Entscheidungen zu treffen vermögen. Ebenfalls von Interesse ist, welche impliziten Strategien dabei verwandt werden. Der Rückgriff auf Feldstudien kann dabei die ökologische Validität der Befunde gegenüber dem Berufsalltag enthobenen Experimenten des JDM-Ansatzes steigern.

Ein weiterer Aspekt mit Relevanz für Schulinspektion ist die Frage danach, ob Expertinnen und Experten andere Bewertungsstrategien als Novizen erwerben und verwenden. Tourmen (2009) zeigt etwa, dass Berufserfahrung einen Einfluss darauf hat, welche konzeptuellen Ressourcen Evaluatoreninnen und Evaluatoren bei ihrer

Bewertungsarbeit nutzen. Sie berichtet, dass Novizinnen und Novizen sich stark an methodischen Vorschriften orientieren, wohingegen Expertinnen und Experten eine stärkere Ergebnisorientierung aufweisen. Letztere würden auch stärker die soziale und politische Dimension von Evaluationen berücksichtigen und mögliche Evaluationsverläufe antizipieren. Sie empfiehlt, in der Ausbildung den Erwerb von konzeptuellem Denken stärker als bisher gegenüber Methodentraining zu fördern. Eine Übertragbarkeit entsprechender Befunde auf Inspektionen ist jedoch aufgrund von Unterschieden nur partiell möglich, denn es werden z. B. statt spezifischer Programme Schulen evaluiert. Darüber hinaus kommen bei der Schulinspektion standardisierte Instrumente zum Einsatz, wohingegen bei der Programmevaluation die Evaluation programmspezifisch konstruiert wird.

Die gesellschaftliche Einbettung von Expertenurteilen wird im folgenden Abschnitt thematisiert.

4. Soziologie der Bewertung

Es handelt sich hierbei um ein in der Entstehung begriffenes Forschungsfeld (vgl. Lamont 2012). Anders als die Psychologie des Urteilens und Entscheidens fokussiert es weniger auf Bewertung als psychisches, sondern als soziales und gesellschaftliches Phänomen. Entsprechende Zeitdiagnosen proklamieren, dass wir in einem Zeitalter der Rechenschaftspflicht (*Age of Accountability*) und in einer Audit-Gesellschaft bzw. Audit-Kultur leben (vgl. Power 1999; Strathern 2000). Schulinspektion kann daher als ein dem Zeitgeist entsprechendes Instrument der Rechenschaftslegung und Qualitätskontrolle verstanden werden.

Aus einer solchen soziologischen Perspektive ist in den Qualitätsrahmen und Kriterienkatalogen eine wertbasierte Hierarchie eingeschrieben und wird darüber zugleich festgeschrieben. Moderne Kommunikationstechnologien ermöglichen sodann die Dissemination solcher konzipierter Werthierarchien an die Schulen im Land und können dadurch quasi auf Distanz wirksam werden – noch bevor Inspektionsteams dort eintreffen (vgl. Prior 2003, S. 168).

Akteure und Akteursgruppen können zu dieser an sie herangetragenen Werthierarchie Stellung beziehen und werden dies unterschiedlich tun, je nachdem, ob sie damit übereinstimmende oder davon abweichende Wertvorstellungen haben. Konflikte im Rahmen von Schulinspektion können u. a. auf unterschiedlichen Werten beruhen. Ein für Schulinspektion charakteristisches Spannungsverhältnis besteht zwischen dessen Kontroll- und Entwicklungsfunktion (vgl. Böttcher & Keune 2012). Dieses Spannungsverhältnis impliziert konkurrierende Wertordnungen. Was ist wichtiger: Vergleichbarkeit zum Zwecke des System-Monitorings oder fallbezogene Förderung von Schulentwicklung? Nicht nur auf Verfahrensebene gibt es unterschiedliche Präferenzen, sondern auch bei den Inspektorinnen und Inspekto-

ren. Einige Inspektorinnen und Inspektoren wollen lediglich evaluieren, wohingegen andere die schulischen Akteure auch gerne beraten möchten.

Wenn es innerhalb eines Inspektionsteams zu unterschiedlichen Bewertungspräferenzen kommt, dann können die Teammitglieder versuchen, die anderen von ihren guten Gründen zu überzeugen (vgl. Sowada & Dederich 2014). Diese Rechtfertigungsarbeit (*Justification Work*) beinhaltet die Verhandlung des Werts einer Sache, wobei bestimmte gesellschaftlich präformierte Wertordnungen hervorgehoben und andere abgeblendet werden (vgl. Jagd 2011, S. 346). Expertenurteile sind demnach nicht neutral, sondern wertgebunden.

5. Diskussion und Ausblick

Lambrecht und Rürup (2012, S. 72) stellten fest, dass „der eigentliche Prozess des Inspizierens, aber auch der politische Kontext des Steuerungsinstruments merkwürdig unterbelichtet“ sei. Genau hier können Forschungsansätze aus den erziehungswissenschaftlichen Nachbardisziplinen neue Perspektiven jenseits der Wirkungs- und Wirksamkeitsfrage eröffnen.

Inspektorinnen und Inspektoren agieren – wie alle Menschen – unter Bedingungen begrenzter Rationalität. Die Konsequenzen für die Bewertungsarbeit von Inspektorinnen und Inspektoren sind bisher unzureichend erforscht. Gleichwohl wären Forschungsergebnisse zu den von Inspektorinnen und Inspektoren eingesetzten Heuristiken in vielerlei Hinsicht hilfreich. Sie könnten zur Weiterentwicklung von Schulinspektion beitragen, z.B. dadurch, dass entsprechendes Wissen ein aufgeklärteres Bewertungsverhalten fördert. Präskriptive Bewertungsvorgaben würden durch Wissen um Bewertungspraxen komplementiert. Entsprechende Erkenntnisse könnten dann in die Aus- und Weiterbildung von Inspektorinnen und Inspektoren integriert werden. Wenn bspw. Abkürzungsstrategien gefunden werden, die eine effizientere Bewertungsarbeit ermöglichen, ohne die erwünschte Qualität der Bewertung zu beeinträchtigen, dann würde durch deren Dissemination auch ein Beitrag zur Belastungsreduktion und zum Gesundheitsschutz geleistet werden.

Für ein besseres Verständnis der Bewertungspraxen von Inspektorinnen und Inspektoren sind interdisziplinäre Forschungsprojekte mit Fachleuten aus der Urteilsforschung angezeigt. Ziel sollte es sein, die eingesetzten Heuristiken zu explizieren. Dies kann in naturalistischen Settings, aber ggf. auch über den Einsatz von Fallvignetten im Rahmen von *Think-Aloud*-Studien geschehen. Erforderlich ist ein Wechsel von der bisher überwiegenden Wirksamkeits- und Wirkungsforschung hin zur Untersuchung von Bewertungsprozessen im Kontext von Schulinspektion. Für ein solches Forschungsprogramm kommt die Sekundärnutzung von bereits vorliegenden Inspektionsdaten leider nicht infrage, da der hier interessierende Prozess

der Generierung von Expertenurteilen darin nicht vorkommt. Wenn hingegen die offizielle Begründung von Expertenurteilen interessiert, so könnte die Analyse von Argumenten in Inspektionsberichten weiterführend sein.

Die Soziologie der Bewertung sensibilisiert dafür, dass das, was unter Qualität verstanden wird, historisch und kulturell kontingent ist. Daran anschließende Forschungsprojekte können z. B. die den Schulinspektionsverfahren zugrunde liegenden Wertordnungen explizieren. Dies kann helfen zu klären, ob wir inspizieren, was wir wertschätzen, oder wertschätzen, was wir inspizieren können (vgl. Biesta 2010). Forschung kann zudem untersuchen, wie unterschiedliche Akteure und Akteursgruppen darüber verhandeln oder gar darum ringen, was guter Unterricht und was gute Schule sei. Die in Qualitätsrahmen und Kriterienkatalogen kodifizierten Bestimmungsversuche sind nicht alternativlos und können hinterfragt werden – genauso wie die Ergebnisse einer Schulinspektion (vgl. Gruschka 2010, S. 78).

Schulinspektion als Steuerungsinstrument ist ebenso hinterfragbar. So zeigen die angeführten Inter-Beurteiler-Studien ein teilweise beachtliches Maß an Beurteilervarianz. Gerade wenn die Expertenurteile von Inspektorinnen und Inspektoren als problematisch und unzuverlässig gerahmt werden, gibt es zwei gegensätzliche Lösungsansätze. Einerseits kann man versuchen, das wahrgenommene Problem intern zu lösen. Hier böten sich z. B. weitere Beurteilerschulungen und Vereindeutigung der Inspektionsinstrumente an, um eine größere Urteilshomogenität zu erzielen. Diese Entwicklungsstrategie zielt darauf ab, soweit wie möglich einheitliche Bewertungen sicherzustellen. Hierdurch würde primär der Kontrollfunktion von Schulinspektion gedient.

Andererseits kann Schulinspektion aufgrund vermeintlicher Defizite für nicht reparabel erklärt werden. Die Kontrollfunktion kann auch über andere Monitoring-Verfahren realisiert werden. Schulinspektion könnte z. B. durch standardisierte Leistungstests ersetzt werden, da letztere über eine vermeintlich höhere wissenschaftliche Güte verfügen. So führt Schulinspektion in den USA bspw. ein Schattendasein im Vergleich zur dort dominanten psychometrischen Testbranche.

Eine andere Rahmung wird in den referierten qualitativen Studien vorgenommen. Expertenurteile werden, da die Bedeutung von Urteilshomogenität relativiert wird, nicht als Schwäche, sondern als eine potenzielle Stärke von Schulinspektionsverfahren gesehen. Gerade Expertenurteile vermögen so, komplexe Phänomene wie Schul- und Unterrichtsqualität besser zu fassen, als dies etwa mittels Leistungstests für Schülerinnen und Schüler möglich wäre. Inspektorinnen und Inspektoren können bei ihren Bewertungen auf ihr implizites Wissen zurückgreifen, das sie im Inspektionsalltag, aber ggf. auch als Lehrkraft und in Leitungsfunktionen erworben haben (vgl. Bromme 2004; Polanyi 1985).

Auf der „Hinterbühne“ darf die vollständige Expertise zur Geltung kommen, auch wenn sie nicht bis ins letzte Detail den expliziten Verfahrensschritten und Vorschriften entsprechen sollte. Eine dies unterstützende Entwicklungsstrategie

würde beinhalten, den fachlichen Austausch zwischen den Inspektorinnen und Inspektoren zu fördern. Ziel wäre es, eine *Community of Practice* bzw. professionelle Lerngemeinschaft entstehen zu lassen, in der Wissen geteilt sowie Probleme thematisiert und bearbeitet werden.

Auf der „Vorderbühne“ wird der Form halber weiterhin ein abgesichertes Ergebnis in einem standardisierten Format präsentiert werden müssen. Wie valide ein solches Ergebnis ist, ist dabei möglicherweise sekundär. Denn auch das valideste Ergebnis bleibt für Schul- und Unterrichtsentwicklung womöglich folgenlos. Wichtiger scheint, zumindest wenn die Entwicklungsfunktion von Schulinspektion im Vordergrund steht, ob es den Inspektorinnen und Inspektoren gelingt, den schulischen Akteuren zu neuen Einsichten zu verhelfen, für Veränderungsbedarf zu sensibilisieren und für Entwicklungsaufgaben zu motivieren.

Literatur

- Altrichter, H., Ehren, M. C. M., McNamara, G. & O'Hara, J. (2014). Wie will Schulinspektion wirken? Analyse von Annahmen über Wirkungsmechanismen von Schulinspektion in sechs europäischen Ländern. In M. Pfeifer (Hrsg.), *Schulqualität und Schulentwicklung. Theorien, Analysen und Potenziale* (S. 184–207). Münster: Waxmann.
- Biesta, G. (2010). Valuing what we measure or measuring what we value? On the need to engage with the question of purpose in educational evaluation, assessment and measurement. In W. Böttcher, J. N. Dicke & N. Hoglebe (Hrsg.), *Evaluation, Bildung und Gesellschaft. Steuerungsinstrumente zwischen Anspruch und Wirklichkeit* (S. 35–46). Münster: Waxmann.
- Borman, G. D., Hewes, G. M., Overman, L. T. & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, 73 (2), 125–230. Verfügbar unter: <http://doi.org/10.3102/00346543073002125> [Zugriff am 30.03.2014]
- Bos, W., Holtappels, H. G. & Rösner, E. (2006). Schulinspektion in den deutschen Bundesländern – eine Baustellenbeschreibung. In W. Bos, H. G. Holtappels, H. Pfeiffer, H.-G. Rolff & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung* (Band 14, S. 81–123). Weinheim: Juventa.
- Böttcher, W. & Keune, M. (2012). Externe Evaluation und die Steuerung der Einzelschule: Kontrolle oder Entwicklung? In M. Ratermann & S. Stöbe-Blossey (Hrsg.), *Governance von Schul- und Elementarbildung* (S. 63–80). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Boyne, G., Day, P. & Walker, R. (2002). The evaluation of public service inspection: A theoretical framework. *Urban Studies*, 39 (7), 1197–1212. Verfügbar unter: <http://doi.org/10.1080/00420980220135563> [Zugriff am 30.03.2014]

- Bromme, R. (2004). Das implizite Wissen des Experten. In B. Koch-Priewe, F.-U. Kolbe & J. Wildt (Hrsg.), *Grundlagenforschung und mikrodidaktische Reformansätze zur Lehrerbildung* (S. 22–48). Bad Heilbrunn: Klinkhardt.
- Cale, L. & Harris, J. (2003). OFSTED revisited: Re-living the experience of an initial teacher training PE inspection. *European Physical Education Review*, 9 (2), 135–161. <http://doi.org/10.1177/1356336X03009002002>
- Campbell, J. & Husbands, C. (2000). On the reliability of OFSTED inspection of initial teacher training: A case study. *British Educational Research Journal*, 26 (1), 39–48. Verfügbar unter: <http://doi.org/10.1080/014119200109507> [Zugriff am 30.03.2014]
- Dederling, K. (2012). Schulinspektion als wirksamer Weg der Systemsteuerung? *Zeitschrift für Pädagogik*, 58 (1), 69–88.
- Döbert, H., Rürup, M. & Dederling, K. (2008). Externe Evaluation von Schulen in Deutschland – die Konzepte der Bundesländer, ihre Gemeinsamkeiten und Unterschiede. In H. Döbert & K. Dederling (Hrsg.), *Externe Evaluation von Schulen. Historische, rechtliche und vergleichende Aspekte* (S. 63–151). Münster: Waxmann.
- Ehren, M. C. M. & Visscher, A. J. (2006). Towards a theory on the impact of school inspections. *British Journal of Educational Studies*, 54 (1), 51–72. Verfügbar unter: <http://doi.org/10.1111/j.1467-8527.2006.00333.x> [Zugriff am 30.03.2014]
- Eisner, E. W. (2013). The roots of connoisseurship and criticism: A personal journey. In M. C. Alkin (Hrsg.), *Evaluation roots: A wider perspective of theorists' views and influences* (2. Aufl., S. e-1–e-7). Los Angeles: Sage. Verfügbar unter: http://www.sagepub.com/upm-data/47742_alkin2e_ch31and32.pdf [Zugriff am 30.03.2014]
- Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59 (1), 255–278. Verfügbar unter: <http://doi.org/10.1146/annurev.psych.59.103006.093629> [Zugriff am 30.03.2014]
- Fend, H. (2008). *Schule gestalten. Systemsteuerung, Schulentwicklung und Unterrichtsqualität*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Fitz-Gibbon, C. T. & Stephenson-Forster, N. J. (1999). Is Ofsted helpful? An evaluation using social science criteria. In C. Cullingford (Hrsg.), *An inspector calls: Ofsted and its effect on school standards* (S. 97–118). London: Kogan Page.
- Gärtner, H. (2013). Praxis und Verhältnis interner und externer Evaluation im Schulsystem im internationalen Vergleich. *Zeitschrift für Erziehungswissenschaft*, 16 (4), 693–712. Verfügbar unter: <http://doi.org/10.1007/s11618-013-0444-7> [Zugriff am 30.03.2014]
- Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62 (1), 451–482. Verfügbar unter: <http://doi.org/10.1146/annurev-psych-120709-145346> [Zugriff am 30.03.2014]
- Gigerenzer, G. & Selten, R. (2002). Rethinking rationality. In G. Gigerenzer & R. Selten (Hrsg.), *Bounded rationality: The adaptive toolbox* (S. 1–12). Cambridge, MA: MIT Press.
- Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday Anchor.

- Gray, J. & Wilcox, B. (1995). The methodologies of school inspection: Issues and dilemmas. In T. Brighouse & B. Moon (Hrsg.), *School Inspection* (S. 127–142). London: Pitman.
- Gruschka, A. (2010). Die Schulinspektion war da und hinterließ einen Bericht. *Pädagogische Korrespondenz*, 23 (41), 75–92.
- Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. New York, NY: Oxford University Press.
- Heid, H. (2013). Logik, Struktur und Prozess der Qualitätsbeurteilung von Schule und Unterricht. *Zeitschrift für Erziehungswissenschaft*, 16 (2), 405–431. Verfügbar unter: <http://doi.org/10.1007/s11618-013-0363-7> [Zugriff am 30.03.2014]
- Husfeldt, V. (2011). Wirkungen und Wirksamkeit der externen Schulevaluation. *Zeitschrift für Erziehungswissenschaft*, 14 (2), 259–282. Verfügbar unter: <http://doi.org/10.1007/s11618-011-0204-5> [Zugriff am 30.03.2014]
- IfBQ (Hrsg.) (2011). *Jahresbericht der Schulinspektion Hamburg. 2009–2010*. Hamburg: Institut für Bildungsmonitoring und Qualitätsentwicklung (IfBQ). Verfügbar unter: <http://epub.sub.uni-hamburg.de/epub/volltexte/2011/10666/> [Zugriff am 30.03.2014]
- IfBQ (Hrsg.) (2012). *Jahresbericht der Schulinspektion, Schuljahr 2010/2011*. Hamburg: Institut für Bildungsmonitoring und Qualitätsentwicklung (IfBQ). Verfügbar unter: <http://nbn-resolving.de/urn:nbn:de:101:1-2013101511516> [Zugriff am 30.03.2014]
- IfBQ (2014). FAQ Schulinspektion. Verfügbar unter: <https://www.hamburg.de/bsb/faq/> [Zugriff am 30.03.2014]
- Jagd, S. (2011). Pragmatic sociology and competing orders of worth in organizations. *European Journal of Social Theory*, 14 (3), 343–359. Verfügbar unter: <http://doi.org/10.1177/1368431011412349> [Zugriff am 30.03.2014]
- Klein, G. (2008). Naturalistic decision making. *Human Factors*, 50 (3), 456–460. Verfügbar unter: <http://doi.org/10.1518/001872008X288385> [Zugriff am 30.03.2014]
- Klein, G., Calderwood, R. & Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. *Journal of Cognitive Engineering and Decision Making*, 4 (3), 186–209. Verfügbar unter: <http://doi.org/10.1518/155534310X12844000801203> [Zugriff am 30.03.2014]
- Kotthoff, H.-G. & Böttcher, W. (2010). Neue Formen der „Schulinspektion“: Wirkungshoffnungen und Wirksamkeit im Spiegel empirischer Bildungsforschung. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 15–39). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kundin, D. M. (2010). A conceptual framework for how evaluators make everyday practice decisions. *American Journal of Evaluation*, 31 (3), 347–362. Verfügbar unter: <http://doi.org/10.1177/1098214010366048> [Zugriff am 30.03.2014]
- Lambrecht, M. & Rürup, M. (2012). Bildungsforschung im Rahmen einer evidence based policy: Das Beispiel „Schulinspektion“. In A. Wacker, U. Maier & J. Wissinger (Hrsg.), *Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung* (S. 57–77). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Lamont, M. (2012). Toward a comparative sociology of valuation and evaluation. *Annual Review of Sociology*, 38 (1), 201–221. Verfügbar unter: <http://doi.org/10.1146/annurev-soc-070308-120022> [Zugriff am 30.03.2014]
- Lincoln, Y.S., Lynham, S.A. & Guba, E.G. (2011). Paradigmatic controversies, contradictions, and emerging confluences, revisited. In N.K. Denzin & Y.S. Lincoln (Hrsg.), *The Sage handbook of qualitative research* (4. Aufl., S. 97–128). Thousand Oaks, CA: Sage.
- Lindgren, J. (2015). The front and back stages of Swedish school inspection: Opening the black box of judgment. *Scandinavian Journal of Educational Research*, 59 (1), 58–76. Verfügbar unter: <http://doi.org/10.1080/00313831.2013.838803> [Zugriff am 30.03.2014]
- Matthews, P., Holmes, J.R., Vickers, P. & Corporaal, B. (1998). Aspects of the reliability and validity of school inspection judgements of teaching quality. *Educational Research and Evaluation*, 4 (2), 167–188. Verfügbar unter: <http://doi.org/10.1076/edre.4.2.167.6959> [Zugriff am 30.03.2014]
- Meyer, H. (2012). Theoretische Grundlagen der Inspektionstätigkeit. In I. Oldenburg (Hrsg.), *Schule und Inspektion. 9 kritische Studien* (S. 17–64). Baltmannsweiler: Schneider-Verlag Hohengehren.
- Millett, A. & Johnson, D.C. (1998). Expertise or ‘baggage’? What helps inspectors to inspect primary mathematics? *British Educational Research Journal*, 24 (5), 503–518. Verfügbar unter: <http://doi.org/10.1080/0141192980240502> [Zugriff am 30.03.2014]
- Müller, S. & Pietsch, M. (2011). Was wir messen, wenn wir Unterrichtsqualität messen. Inter-Beurteilerübereinstimmungen und -Reliabilität bei Unterrichtsbeobachtungen im Rahmen von Schulinspektion. In S. Müller, M. Pietsch & W. Bos (Hrsg.), *Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht* (S. 33–55). Münster: Waxmann.
- Nixon, J. & Rudduck, J. (1993). The role of professional judgment in the local inspection of schools: A study of six local education authorities. *Research Papers in Education*, 8 (2), 135–148. Verfügbar unter: <http://doi.org/10.1080/0267152930080202> [Zugriff am 30.03.2014]
- Nutley, S., Levitt, R., Solesbury, W. & Martin, S. (2012). Scrutinizing performance: How assessors reach judgements about public services. *Public Administration*, 90 (4), 869–885. Verfügbar unter: <http://doi.org/10.1111/j.1467-9299.2011.02022.x> [Zugriff am 30.03.2014]
- Pietsch, M. (2010). Evaluation von Unterrichtsstandards. *Zeitschrift für Erziehungswissenschaft*, 13 (1), 121–148.
- Pietsch, M. & Tosana, S. (2008). Beurteilereffekte bei der Messung von Unterrichtsqualität. *Zeitschrift für Erziehungswissenschaft*, 11 (3), 430–452. Verfügbar unter: <http://doi.org/10.1007/s11618-008-0021-7> [Zugriff am 30.03.2014]
- Polanyi, M. (1985). *Implizites Wissen*. Frankfurt a. M.: Suhrkamp.
- Power, M. (1999). *The audit society: Rituals of verification*. Oxford: Oxford University Press.

- Praetorius, A.-K. (2013). Einschätzung von Unterrichtsqualität durch externe Beobachterinnen und Beobachter – Eine kritische Betrachtung der aktuellen Vorgehensweise in der Schulpraxis. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 31 (2), 174–185.
- Preuß, B., Wissinger, J. & Brüsemeister, T. (2015). Einführung der Schulinspektion: Struktur und Wandel regionaler Governance im Schulsystem. In H.J. Abs, T. Brüsemeister, M. Schemmann & J. Wissinger (Hrsg.), *Governance im Bildungssystem* (S. 117–142). Wiesbaden: Springer VS Verlag für Sozialwissenschaften.
- Prior, L. (2003). *Using documents in social research*. London: Sage.
- Rürup, M. (2008). Typen der Schulinspektion in den deutschen Bundesländern. *Die Deutsche Schule*, 100 (4), 467–477.
- Sadler, D.R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34 (2), 159–179. Verfügbar unter: <http://doi.org/10.1080/02602930801956059> [Zugriff am 30.03.2014]
- Sadler, D.R. (2014). The futility of attempting to codify academic achievement standards. *Higher Education*, 67 (3), 273–288. Verfügbar unter: <http://doi.org/10.1007/s10734-013-9649-1> [Zugriff am 30.03.2014]
- Scheerens, J., Glas, C. & Thomas, S. (2003). *Educational evaluation, assessment, and monitoring: A systemic approach*. Lisse: Swets & Zeitlinger.
- Simon, H.A. (2000). Bounded rationality in social science: Today and tomorrow. *Mind & Society*, 1 (1), 25–39. Verfügbar unter: <http://doi.org/10.1007/BF02512227> [Zugriff am 30.03.2014]
- Sinkinson, A. & Jones, K. (2001). The validity and reliability of OFSTED: Judgements of the quality of secondary mathematics initial teacher education courses. *Cambridge Journal of Education*, 31 (2), 221. Verfügbar unter: <http://doi.org/10.1080/0305764012006155> [Zugriff am 30.03.2014]
- Sommer, N. (2011). Wie beurteilen schulische Gruppen die erlebte Schulinspektion? Ergebnisse einer Befragung. In S. Müller, M. Pietsch & W. Bos (Hrsg.), *Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht* (S. 137–164). Münster: Waxmann.
- Sowada, M.G. & Dederig, K. (2014). Ermessensspielräume in der Bewertungsarbeit von Schulinspektor/innen. *Zeitschrift für Bildungsforschung*, 4 (2), 119–135. Verfügbar unter: <http://doi.org/10.1007/s35834-014-0091-y> [Zugriff am 30.03.2014]
- Strathern, M. (Hrsg.) (2000). *Audit cultures: Anthropological studies in accountability, ethics and the academy*. London: Routledge.
- Tourmen, C. (2009). Evaluators' decision making: The relationship between theory, practice, and experience. *American Journal of Evaluation*, 30 (1), 7–30. Verfügbar unter: <http://doi.org/10.1177/1098214008327602> [Zugriff am 30.03.2014]
- Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185 (4157), 1124–1131. Verfügbar unter: <http://doi.org/10.1126/science.185.4157.1124> [Zugriff am 30.03.2014]

- Wilcox, B. & Gray, J. (1996). The methodological problems of inspection. *Inspecting schools: Holding schools to account and helping schools to improve* (S. 65–80). Buckingham: Open University Press.
- Wilson, T.A. (1996). *Reaching for a better standard: English school inspection and the dilemma of accountability for American public schools*. New York, NY: Teachers College Press.

Daten für die Schulentwicklung – auf die Perspektive kommt es an

Interpretationen und Fehlinterpretationen von Daten aus der Schulinspektion

Knut Schwippert

Zusammenfassung

Dem Schulinspektionsbericht von 2009/2010 ist zu entnehmen, dass neben der Aufgabe, den Einzelschulen Rückmeldung über den Stand ihrer Qualitätsentwicklung zu geben, die Schulinspektion auch den Auftrag zum Bildungsmonitoring hat. Das damit betraute Institut stellt Daten über die Qualität des Hamburger Schulsystems zur Verfügung, die vor dem Hintergrund unterschiedlicher Fragestellungen ausgewertet werden können. Neben dem Impuls, direkte Informationen an Schulen und für die Unterrichtsentwicklung vor Ort bereitzustellen, können die Daten auch im Sinne eines Bildungsmonitorings genutzt werden. Werden Bedingungen, unter denen die Daten erhoben wurden, bei der rechnerischen Auswertung nicht berücksichtigt, kann es jedoch zu Fehlinterpretationen kommen. Cronbach hat schon 1976 die Auffassung vertreten, dass die (bis dato) durchgeführte Unterrichtsforschung mehr verschleiert, denn zum Vorschein gebracht habe. Dieser Beitrag möchte für die jeweiligen Tragweiten unterschiedlicher Analyseansätze sensibilisieren und dies anhand der Schulinspektionsdaten durch vergleichend gegenübergestellte Mehrebenenanalysen verdeutlichen.

Dem Schulinspektionsbericht von 2009/2010 ist zu entnehmen (Institut für Bildungsmonitoring 2011, S. 6), dass neben der Aufgabe, den Einzelschulen Rückmeldung über den Stand ihrer Qualitätsentwicklung zu geben, die Schulinspektion auch den Auftrag zum Bildungsmonitoring hat. Hierbei stellt sie übergreifende Erkenntnisse über die Qualität des Hamburger Schulsystems im Rahmen ihrer Jahresberichte zur Verfügung. Die im Rahmen der Inspektionen erhobenen Daten können dann vor dem Hintergrund ganz unterschiedlicher Fragestellungen ausgewertet werden. Neben dem Impuls, direkte Informationen an Schulen und für die Unterrichtsentwicklung vor Ort bereitzustellen, können diese auch im Sinne eines Bildungsmonitorings für Fragen, die für die Bildungsadministration bzw. behördliche Bildungsplanung relevant sind, ausgewertet werden. Und schließlich bieten die Daten auch eine mögliche Basis zur Beantwortung wissenschaftlicher Frage-

stellungen, die nicht unmittelbar Wirkungen im Unterricht bzw. in der Bildungssteuerung entfalten. Je nachdem, welche Fragestellungen mit den Inspektionsdaten beantwortet werden sollen, ist sicher zu stellen, dass der besonderen Struktur der Daten bei den empirischen Analysen Rechnung getragen wird. So sind bei Analysen einzelner Klassen bzw. Schulen ggf. andere Verfahren anzuwenden als bei solchen, die das gesamte Bildungssystem in den Blick nehmen. Dabei ist die Zuordnung von Fragestellung, der Identifizierung der angemessenen Analyseebene (Individuum, Klasse, Schule, Bezirk, System) und der jeweils angemessenen Analysemethode nicht beliebig austauschbar. Dieser Beitrag möchte anhand von Re-Analysen der Hamburger Schulinspektionsdaten für die jeweiligen Tragweiten unterschiedlicher Analyseansätze sensibilisieren. Auch wenn dieser Beitrag insbesondere mit der Perspektive einer wissenschaftlich orientierten Sekundäranalyse geschrieben wurde, so sind die methodologischen Anmerkungen auch für weiterführende Analysen im Rahmen der Hamburger Schulinspektion sowie für Analysen anderer externer Evaluationen in Deutschland relevant – wenn spezifische Perspektiven für die Auswertung eingenommen werden.

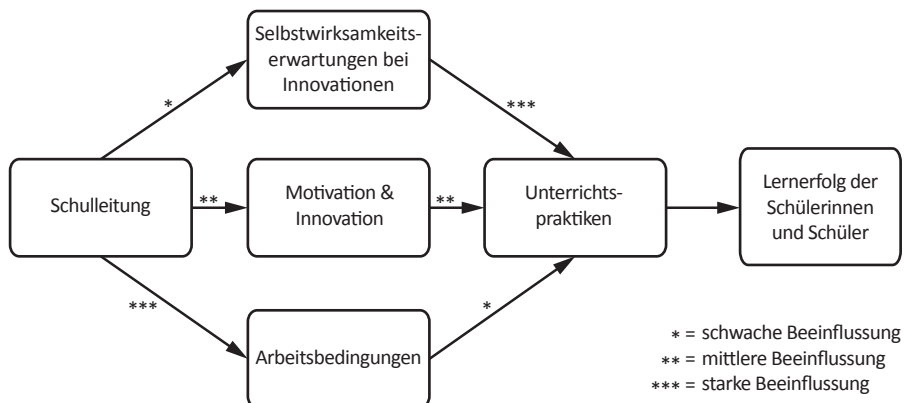
Aktuelle Untersuchungen zur Nutzung von Daten aus der Schulinspektion können auf eine Reihe von Daten zurückgreifen, die eine empirische Absicherung vermuteter bzw. in Einzelfällen beobachteter Zusammenhänge erlauben. Daten aus Befragungen und Unterrichtsbeobachtungen im Rahmen von Schulinspektionen basieren auf einem komplexen Stichproben- und Testdesign (vgl. Adams & Wu 2000; Institut für Bildungsmonitoring 2009, 2011; Institut für Bildungsmonitoring und Qualitätsentwicklung 2012; Martin et al. 2007; Martin et al. 1999). Werden Bedingungen, unter denen die Daten erhoben wurden, bei der rechnerischen Auswertung nicht berücksichtigt, kann es zu Fehlinterpretationen kommen. Cronbach hat schon 1976 in seiner Veröffentlichung *Research on Classrooms and Schools* (Cronbach 1976) die Auffassung vertreten, dass die (bis dato) durchgeführte Unterrichtsforschung mehr verschleiert, denn zum Vorschein gebracht habe.

Cronbachs Kritik richtet sich insbesondere darauf, dass bei Analysen von Daten, in denen Schülerinnen und Schüler ganzer Klassen bzw. Schulen getestet wurden, nicht angemessen berücksichtigt würde, dass sich Schüler einer Schulklasse (oder einer Schule) untereinander ähnlicher sind als die gleiche Anzahl von Schülern, die zufällig aus verschiedenen Klassen stammen. Um eine angemessene Berücksichtigung dieser sogenannten Klumpungseffekte geht es in diesem Beitrag (vgl. z. B. Cochran 1972; Kish 1995). Aber auch wenn die Daten auf zu hoher Ebene zusammengefasst werden, kann sich ein Problem ergeben: Bei der Betrachtung von solchen aggregierten Daten kann es zu einer Übergeneralisierung kommen, wenn die Ergebnisse gleichförmig auf alle beteiligten Bildungseinheiten, also z. B. Schulen oder Klassen, übertragen werden. Solche Form der Interpretation (im Sinne eines: ‚Eigentlich sind wir alle gleich.‘) kann dazu führen, dass problematische, aber auch vorbildliche Situationen übersehen werden. Der Kern der zuvor skizzierten Kritik

ist, dass es bei Untersuchungen im Bereich der Schul- und Unterrichtsforschung in Abhängigkeit von der jeweils zu beantwortenden Fragestellung zur fehlerhaften Berechnung z. B. von Schul- bzw. Unterrichtseffekten kommen kann, wenn die Gruppierung von Individuen in Gruppen unberücksichtigt bleibt (vgl. Ditton 1998; Janke 2006; Schwippert 2002) und die zu beantwortende Frage nicht mit der verwendeten Analyseverfahren korrespondiert. Dies hieße für die Schulinspektion, dass dann, wenn generelle Effekte auf Bildungssystemebene zu untersuchen sind, auch diese Ebene in den Blick genommen werden muss. Wenn es aber um das Wirken oder das Zusammenspiel von Qualitätsmerkmalen zwischen Schulen geht, müssen die Beobachtungen auch für die Beschreibung von Unterschieden zwischen den Schulen geeignet sein. Schließlich muss es analog gelten, Qualitätsmerkmale im Unterricht zu erheben, wenn Unterschiede des Unterrichts zwischen verschiedenen Klassen betrachtet werden sollen. Im letzten Fall wäre eine differenzielle Diagnostik bzw. ein differenzieller Blick auf verschiedene Merkmale des Unterrichts zu richten, während bei der Beurteilung der Bildungssystem- bzw. Schulqualität eine umfassendere und für die einzelnen Schulen bzw. Unterrichte spezifischere Diagnostik (die dann über die verschiedenen Bildungseinrichtungen hinweg ihre Breite belegen) methodisch angemessen ist.

Für die nachfolgend präsentierten Betrachtungen auf verschiedenen Ebenen des Bildungssystems wird hier auf das Modell (siehe Abbildung 1) des Schulleitungshandelns von Leithwood et al. (2006) zurückgegriffen (zitiert nach Pietsch 2014).

Abbildung 1: Modell des Einflusses von Schulleitungen auf die Unterrichtsgestaltung von Lehrkräften (adaptierte Übersetzung von Pietsch 2014)

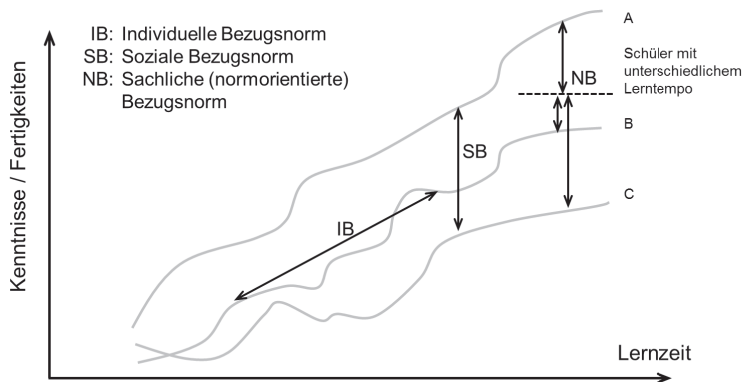


In diesem Modell ist die empirisch abgesicherte Wirkungskette des Schulleitungshandelns auf „Selbstwirksamkeitserwartungen bei Innovationen“, auf „Motivation und Identifikation“ und auf „Arbeitsbedingungen“ und deren Wirkungen auf die

„Unterrichtspraktiken“, die wiederum direkte Effekte auf den Lernerfolg von Schülerinnen und Schülern haben, abgebildet. Aus diesem Wirkungsmodell sei hier der Pfad (die Wirkung) des Schulleitungshandelns auf die „Motivation und Identifikation“ von Lehrkräften herausgegriffen. Anhand des gewählten Beispiels wird verdeutlicht, zu welchen unterschiedlichen Einschätzungen man bezüglich des Zusammenhangs zwischen der Zufriedenheit der Lehrkräfte und dem Schulleitungshandeln in Bezug auf den Aspekt der Führung kommt, wenn mit Daten auf unterschiedlichen Ebenen des Bildungssystems argumentiert wird.

Doch warum sollten bei der Analyse des Zusammenhangs des Führungshandelns von Schulleitungen mit der Zufriedenheit der Lehrkräfte unterschiedliche Ebenen des Bildungssystems betrachtet werden? Ein Blick auf drei mögliche Bezugsnormorientierungen bei Leistungsbewertungen von Schülerinnen und Schülern durch Lehrkräfte soll als Analogiebeispiel zur Beantwortung dieser Frage herangezogen werden. Nach Rheinberg (2001) können individuelle, soziale und sachliche Beurteilungsmaßstäbe unterschieden werden. (Alternativ werden diese auch intraindividuelle, interindividuelle und kriteriale Bezugsnormen genannt, vgl. Krapp & Weidenmann 2001.) Jede dieser Bewertungsnormen hat ihre spezifischen Vor- und Nachteile (Rheinberg bezeichnet Letztere als „blinde Flecke“), sodass keine der einzelnen Normen pauschal als besser oder weniger geeignet bewertet werden kann.

Abbildung 2: Bezugsnormorientierungen nach Rheinberg (2001, S. 62) (eigene Darstellung)



In Abbildung 2 sind die (fiktiven) Lernverläufe von drei sich unterschiedlich entwickelnden Kindern abgetragen. Bei der Bewertung des Lernstandes lässt sich zunächst die *individuelle Bezugsnormorientierung* (IB) anlegen, die hier

die Veränderungen einer einzelnen Person über die Zeit in den Blick nimmt und hierbei – orientiert an einem Ausgangspunkt – den Lernzuwachs bewertet. Bei der *sozialen Bezugsnormorientierung* (SB) wird der relative Kenntnisstand eines Kindes im Vergleich mit – in der Regel – den anderen Kindern in derselben Schulklasse bewertet. Diese Bezugsnormorientierung ist im deutschen Schulsystem eine sehr oft genutzte Bezugsnorm. Schließlich gibt es die *sachliche (normorientierte) Bezugsnorm* (NB), bei der, bezogen auf ein vorgegebenes Niveau, die Leistungen der Kinder beurteilt werden. Für Letzteres sind das Zentralabitur oder die Führerscheinprüfung Beispiele.

Der blinde Fleck bei der *individuellen Bezugsnormorientierung* ist, dass – trotz detailliertem Wissen über die individuellen Veränderungen von Schülerinnen und Schülern – keine Information vorliegt, ob der Schüler oder die Schülerin altersgemäß eher gut oder eher schwach erscheint, es also nicht beurteilt werden kann, ob besondere Förderung notwendig ist oder nicht. Der blinde Fleck der *sozialen Bezugsnormorientierung* ist die fehlende Sensibilität gegenüber der generellen Entwicklung der Gruppe. Anhand der *sozialen Bezugsnormorientierung* kann nicht festgestellt werden, ob sich die Gruppenmitglieder weiterentwickelt haben oder nicht. Hieraus ergibt sich der Nachteil, dass schwache Schülerinnen oder Schüler einer Lerngruppe, obwohl sie wie alle anderen in der Gruppe einen bedeutenden Lernzuwachs (beurteilt nach der *individuellen Bezugsnorm*) realisieren konnten, auch weiterhin als schwach bewertet werden. Die *sachliche (normorientierte) Bezugsnorm* hat insbesondere dann einen blinden Fleck, wenn als Bewertungsergebnis nur der Befund „bestanden“ oder „nicht bestanden“ zurückgemeldet wird. In beiden Fällen ist nicht deutlich, ob das geforderte Lernziel knapp oder deutlich erreicht bzw. verfehlt wurde.

Lehrkräfte wünschen sich in der Regel die Möglichkeit der *individuellen Bezugsnormorientierung*, sind aber gehalten, in der Regel die *soziale Bezugsnorm* anzuwenden. (Diese ist auch die, die die Schülerinnen und Schüler untereinander anlegen, wenn sie Noten oder Zeugnisse untereinander vergleichen.) Schließlich wäre die *sachliche (normorientierte) Bezugsnorm* der fairste Ansatz, wenn es um die Zuweisung von Bildungsressourcen bzw. -zertifikaten geht, da hierbei die zufällige Zusammensetzung der Gruppe, aus der eine Person kommt, keine Rolle spielt. Mit der Frage nach der Wahl der Bezugsnorm ist auch die Frage verbunden, was im Fokus der Bewertung steht (individuelle Entwicklung, Status in einer Gruppe, Normerreicherung).

Übertragen auf die Betrachtung der Daten aus der Schulinspektion hieße dies, zunächst die Frage zu klären, mit welcher Perspektive bzw. auf welcher Ebene die Informationen aus der Schulinspektion genutzt werden sollen. Ist die Frage genereller Zusammenhänge im Hamburger Schulsystem zu klären, steht eine spezifische Schulform im Fokus der Betrachtung oder sind die Einzelschulen oder sogar der Unterricht der Lehrkräfte bei der Inspektion im Zentrum der Betrachtung?

Im Folgenden wird gezeigt, dass es – bei gleichen Datenlagen – zu unterschiedlichen und differenzierten Befunden führen kann, wenn verschiedene Betrachtungsperspektiven eingenommen werden. Hierbei wird verdeutlicht, dass keine Perspektive „die richtige“ ist – vielmehr wird argumentiert, dass es auf die Fragestellung ankommt, welche Perspektive eingenommen werden sollte oder sogar muss.

1. Empirische Annäherung

1.1 Datenlage

Für die Analysen stehen die Daten der Hamburger Schulinspektion der ersten Inspektionswelle zur Verfügung. Hierbei wurden im Zeitraum von 2007 bis 2013 alle Hamburger Schulen einmal besucht und anhand der im Orientierungsrahmen für die Qualitätsentwicklung an Hamburger Schulen (Behörde für Bildung und Sport 2006) genannten Qualitätsaspekte inspiziert. Zur Konzeption, Organisation, Durchführung und Auswertung von Daten aus der Schulinspektion sei auf die seit der Einführung der Hamburger Schulinspektion regelmäßig erscheinenden Jahresberichte verwiesen (vgl. Institut für Bildungsmonitoring 2009, 2011; Institut für Bildungsmonitoring und Qualitätsentwicklung 2012; Beiträge von Diedrich in diesem Band).

Tabelle 1: Anzahl von befragten Lehrkräften der ersten Inspektionskohorte nach Schulform und Schuljahr

Schulform Schuljahr	Grund- schule	GHR- Schule	Gesamt- schule	Gym- nasium	Sonder-/ Förder- schule	beruf- liche Schule	Stadtteil- schule	Gesamt
2006/2007	99	116	13	27	0	26	0	281
2007/2008	364	194	246	426	38	369	0	1637
2008/2009	186	122	313	436	230	369	0	1656
2009/2010	445	82	241	485	164	448	0	1865
2010/2011	598	0	0	539	94	229	573	2033
2011/2012	610	0	0	224	21	168	152	1179
2012/2013	290	0	0	0	118	66	132	606
Gesamt	2592	514	813	2137	665	1675	857	9253

Um die Betrachtungen zu Effekten unterschiedlicher Gruppierungen von Daten nicht zu komplex und damit unübersichtlich zu gestalten, wird bei dem Vergleich zwischen Schulen ausschließlich auf die Schulinspektion der Grundschulen zurückgegriffen. Dies geschieht aus zwei Gründen: Die Schulform „Grundschule“ hat im Laufe des ersten Schulinspektionszyklus keine wesentlichen Strukturänderung

im Rahmen von Schulstrukturdebatten erfahren. Die Grundschule ist – da sie als Gesamtschulsystem angelegt ist – auch die zahlenmäßig stärkste Bildungseinrichtung mit $N=163$ Schulen und $N=2.592$ Lehrkräften. In Tabelle 1 ist die Verteilung der befragten Lehrkräfte nach Schulform und Schuljahr aufgeführt. In Tabelle 2 ist die Verteilung der inspizierten Schulen nach Schulform und Schuljahr eingetragen.

Tabelle 2: Anzahl von inspizierten Schulen der ersten Inspektionskohorte nach Schulform und Schuljahr

Schulform Schuljahr	Grund- schule	GHR- Schule	Gesamt- schule	Gym- nasium	Sonder-/ Förder- schule	beruf- liche Schule	Stadtteil- schule	Gesamt
2006/2007	8	5	1	1	0	1	0	16
2007/2008	24	9	7	14	1	10	0	65
2008/2009	14	7	8	13	11	10	0	63
2009/2010	31	4	7	14	4	13	0	73
2010/2011	32	0	0	18	4	6	14	74
2011/2012	39	0	0	6	1	4	5	55
2012/2013	15	0	0	0	2	1	2	20
Gesamt	163	25	23	66	23	45	21	366

1.2 Betrachtung der hierarchischen Datenstruktur

Wenn im Umgang mit Daten aus der Schulinspektion unberücksichtigt bleibt, dass Lehrkräfte in Schulen und diese wiederum in Schulformen organisiert sind, kann es zu fehlerhaften Schuleffektmessungen kommen (vgl. Treiber 1980). Aufgrund dieser sogenannten ‚geschachtelten‘ (hierarchischen) Datenstrukturen sind bei der Untersuchung z. B. der wahrgenommenen Führung durch die Schulleitungen feststellbare Unterschiede nicht nur zwischen den Lehrkräften, sondern auch zwischen den Schulen (Schulformen etc.) zu erwarten. Die Notwendigkeit einer getrennten Betrachtung von Innerhalb- (intra-) und Zwischen- (inter-)Schuleffekten belegt Treiber analog anhand der Betrachtung von Schülerinnen und Schülern in Klassen und in Schulen (vgl. ebd., S. 212). Seiner Auffassung nach sind „Analysen auf Schülerebene [...] nur dann korrekt, wenn die Gruppierung von Schülern in ihren Schulklassen ohne nennenswerten Einfluß“ auf das untersuchte Merkmal bleibt. Analoges gilt für die Gruppierung von Lehrkräften in Schulen. Überträgt man somit dieses Argument auf die Schulinspektion, gilt dies für die Gruppierung der Lehrkräfte (in Schulen) und damit möglicherweise auch für eine fehlerhafte Ermittlung der im Fokus stehenden (Schul-)Effekte. Bei den entsprechenden Analysen ist daher die Schule als Analyseeinheit zu berücksichtigen.

Der Effekt, dass sich Mitglieder einer realen Gruppe untereinander ähnlicher sind als eine gleich große, aber zufällig ausgewählte Gruppe von Personen, wird in der Empirie (Stichprobentheorie) Klumpungseffekt genannt. Er besagt, dass eine Gruppen von Personen, die gleiche Erfahrungen, gleiche Motivation und z. B. auch vergleichbare Bildungswege bzw. -erfahrungen haben, sich auch in anderen Merkmale (z.B. Interessen, Fähigkeiten, Fertigkeiten) ähnlicher sind als eine gleich große Gruppe von Personen, die entsprechende gemeinsame Erfahrung nicht teilen. Klumpungseffekte können als Effekt der Eingangsselektivität für eine Gruppe aber auch – wie skizziert – durch gemeinsame (Lern-)erfahrungen entstehen. Als empirisches Maß des Klumpungseffekts wird die *Intra-Class-Correlation* (ICC) herangezogen. Sie kann Werte von 0 (es gibt keine Klumpungseffekte) bis 1 (es gibt keine Unterschiede zwischen den Mitgliedern einer Gruppe, alle Unterschiede sind nur zwischen den Gruppen zu beobachten) annehmen. In Bezug auf Kompetenzen von Schülerinnen und Schülern in Deutschland sind ICCs von ca. 0,12 bis 0,20 erwartbare Größen.

Im Rahmen der bei den Schulinspektionen ermittelten Informationen z. B. zur Zufriedenheit der Lehrkräfte und auch der durch die Lehrkräfte wahrgenommenen Führung durch die Schulleitung sind ebenfalls bedeutende *Intra-Class-Correlations* zu erwarten. Diese können durch die Zusammensetzung von Kollegien z. B. hinsichtlich ihrer Ausbildung (Studienort und Referendariat) oder auch ihrer Motivation, an eine bestimmte Schule zu gehen, entstehen. Gerade Letzteres ist ein Effekt, der mit der Veröffentlichung von Ergebnissen von Leistungsdaten der Schulen in Zusammenhang zu bringen ist. Während Eltern insbesondere aus bildungsfernen Familien kaum die Möglichkeit sehen bzw. nutzen, ihre Kinder an spezifisch ausgewählten Schulen anzumelden, ist die Entscheidung der Lehrkräfte, an eine bestimmte Schule zu gehen, durchaus einerseits von (typischen) Motiven und Interessen der Lehrkräfte und andererseits durch die zunehmende Veröffentlichung von Daten über die Schulen beeinflusst (vgl. für die Vereinigten Staaten von Amerika z. B. Frankenberg 2006).

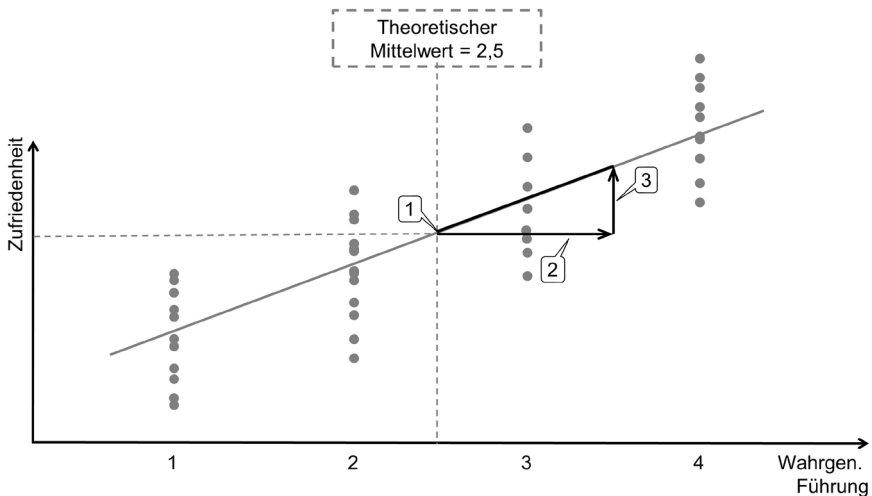
1.3 Die Methode der Regressionsanalyse

Um Hypothesen über den Zusammenhang von zwei Merkmalen zu überprüfen, bieten sich bei empirischen Untersuchungen verschiedene Methoden an. Eine sehr oft angewandte Methode ist die Regressionsanalyse, wenn Aussagen über den Einfluss einzelner oder mehrerer Variablen auf ein betrachtetes Kriterium (=eine abhängige Variable) untersucht werden sollen.

In Abbildung 3 ist das Prinzip der Regressionsanalyse abgebildet. Von befragten Lehrkräften liegen in diesem Beispiel jeweils paarweise Informationen über ihre berufliche Zufriedenheit und die von ihnen wahrgenommene Führung

der Schulleitung vor. Ziel des Regressionsverfahrens ist es, anhand eines Modells (lineares Gleichungssystem) die Unterschiede (Varianz) einer zu beschreibenden Variable (hier Zufriedenheit der Lehrkräfte) anhand einer anderen Variable (hier Führungshandeln der Schulleitungen) vorherzusagen. Die Regressionssteigung gibt dabei ein Maß der Abhängigkeit der Zufriedenheit der Lehrkräfte vom durch die Lehrkräfte wahrgenommenen Führungshandeln der Schulleitungen wieder. Beim theoretischen Skalenmittelwert von 2,5 der wahrgenommenen Führung gibt die Regressionsgerade (s. Abbildung 3) an der gestrichelten Linie ① den Grad der damit im Zusammenhang stehenden Zufriedenheit der Lehrkräfte wieder. Nimmt der Grad der wahrgenommenen Führung um eine Skaleneinheit auf 3,5 zu ②, verändert sich die Zufriedenheit in diesem Beispiel positiv ③. In den folgenden Abbildungen, die anhand der realen Inspektionsdaten berechnet wurden, wird nunmehr nur der in Abbildung 3 dunkel dargestellte Teil der Regressionsgeraden wiedergegeben, da dieser alle relevanten Informationen über den Zusammenhang der von den Lehrkräften wahrgenommenen Führung der Schulleitungen und ihrer beruflichen Zufriedenheit enthält.

Abbildung 3: Visualisierung der Analysemethode „Regressionsanalyse“



In Abhängigkeit davon, wie eine zu untersuchende Stichprobe bei den Analysen gruppiert wird, lassen sich unterschiedliche Effekte beobachten. In diesem Zusammenhang verweist Achtenhagen (1981, S. 321) auf eine Studie von Robinson, in der dieser zeigt, wie es bei Untersuchungen aufgrund von verschiedenen Gruppierungen einer Stichprobe dazu kommt, „daß der Korrelationskoeffizient zwischen

0,20 und 0,95 schwankt, je nachdem ob die Daten für Individuen, einzelne Staaten oder Gruppierungen von Staaten aufbereitet werden“. Bei der Bewertung, welcher Befund richtig ist, kommt es zunächst auf die im Fokus stehende Frage an, in der sich unterschiedliche Voraussetzungen bzw. Perspektiven widerspiegeln. Wichtig ist, dass die jeweils zur Forschungsfrage passende Analysemethode ausgewählt wird. Die Entscheidung für oder wider eine dieser Methoden – und darauf sei hier ausdrücklich hingewiesen – ist nicht beliebig. Die Auswahl folgt einer Abwägung von fachlich und methodisch zu berücksichtigenden Faktoren.

Augenfällig wird die besondere Datenstruktur, wenn neben den Lehrkräften auch die Schulleitungen in den Blick genommen werden sollen. Bei der Zusammenführung bestehen im Wesentlichen zwei Möglichkeiten: Zum einen können die Daten der Lehrkräfte je Schule gemittelt und mit den Daten der Schulleitungen zusammengespielt werden. Oder es können die Daten der Schulleitungen jeweils an die Daten der Lehrerinnen und Lehrer kopiert werden. Mit diesen Vorgehensweisen sind jeweils wesentliche Nachteile verbunden: Eine Zusammenfassung der Eigenschaften von Lehrkräften oder Merkmalen innerhalb einer Schule zu einem einzigen Wert zieht bedeutende Datenverluste nach sich, da alle individuellen Unterschiede zwischen den Lehrkräften innerhalb der Schulen verloren gehen. Zudem wird in den anschließenden Analysen nur mit ‚der durchschnittlichen Lehrkraft‘ gerechnet. Bei der Vervielfältigung von Schulleitungs- bzw. Strukturmerkmalen der Schule muss implizit vorausgesetzt werden, dass diese Merkmale auf jede Lehrkraft gleichermaßen wirken. In Bezug auf die Lehrkraft-Schulleitungs-Interaktion würde dies bedeuten, dass das von der Schulleitung gezeigte Verhalten auf jede Lehrkraft in der Schule den gleichen Einfluss ausübt. Eine Voraussetzung, die bisherigen Erfahrungen widerspricht. Und schließlich kommt es durch die Vervielfältigung der Schulleitungsdaten systematisch zu fehlerhaften Signifikanzschätzungen bei nachfolgenden Analysen.

Die in der Literatur übereinstimmende Forderung ist die, dass bei der Untersuchung von (hierarchisch) geschachtelten Daten deutlich zwischen den Ebenen zu unterscheiden ist. Bei bestimmten Fragestellungen ist danach zu berücksichtigen, dass z. B. Lehrkräfte in Schulen oder Schulen in Schulformen etc. organisiert sind. Unterbleibt eine solche analytische Trennung, kann es bei der Analyse z. B. von Zusammenhangshypothesen zu Fehlinterpretationen kommen.

Die hier vorgestellten Analysen der Schulinspektion lenken schrittweise den Blick auf verschiedene Ebenen der Schulinspektion. Als anspruchsvolles Standardverfahren bei der Analyse von Schulvergleichsuntersuchungen hat sich die sogenannte Mehrebenenanalyse bewährt, auf deren Darstellung hier jedoch verzichtet wird. Ein Anwendungsfall im Rahmen der Schulforschung ist bei Schwippert (2008) beschrieben.

1.4 Die verwendeten Skalen

Um die Aufmerksamkeit auf einen relevanten Aspekt im Hamburger Schulsystem zu lenken, wird auf das oben vorgestellte Modell des Einflusses von Schulleitungen auf den Lernerfolg von Schülerinnen und Schüler zurückgegriffen, in dem der Einfluss des Führungshandelns von Schulleitungen auf die Motivation von Lehrkräften eine wichtige Rolle spielt. Exemplarisch wird somit aus den Daten der Hamburger Schulinspektion die Skala des Schulleitungshandelns in Fragen der Führung und die Skala zur Zufriedenheit der Lehrkräfte (aus der sich Einflüsse auf motivationale Merkmale ableiten lassen) genutzt.

Das Führungshandeln von Schulleitungen wird anhand einer Reihe von Fragen, die diesbezüglich an Lehrkräfte gestellt wurden, operationalisiert und in eine Skala zusammengefasst.¹ Tabelle 3 gibt Auskunft über die in die Skala zum Führungshandeln eingehenden Items. Alle Fragen konnten auf einer vierstufigen Skala beantwortet werden, wobei der niedrigste Wert für Ablehnung bzw. Nicht-Zustimmung und der höchste für Akzeptanz bzw. Zustimmung stehen. Die über Mittelwertanalysen ermittelten Werte sind entsprechend zu interpretieren. Die Skala „Wahrgenommene Führung“ weist eine Reliabilität von $r_{tt(\text{Führung})} = 0,846$ (11 Items) auf.

Tabelle 3: Variablen der Skala „Wahrgenommene Führung im Schulleitungshandeln“

Variablenname	Frageformulierung im Lehrkräftefragebogen
LF010340	Die Schulleiterin/der Schulleiter sorgt für einen guten Informationsfluss an der Schule.
LF010350	Die Schulleiterin/der Schulleiter hat den Überblick über das Geschehen an der Schule.
LF010360	Die Schulleiterin/der Schulleiter vertritt ein klares pädagogisches Konzept.
LF010380	Die Schulleiterin/der Schulleiter repräsentiert die Schule gut nach außen.
LF010400	Die Schulleiterin/der Schulleiter ist bei wichtigen Fragen ansprechbar.
LF010021	An unserer Schule haben wir Entwicklungsprioritäten erarbeitet.
LF010441	An unserer Schule werden die Lehrerinnen und Lehrer bei wichtigen Entscheidungen systematisch beteiligt.
LF010131	An unserer Schule sind Zuständigkeiten für Entwicklungsmaßnahmen klar festgelegt.
LF010030	An unserer Schule gibt es ein systematisches Qualitätsmanagement, für das klar geregelt ist, wer was macht.
LF010230	Ich bin der Meinung, dass wir an unserer Schule momentan die zentralen Ziele und Probleme angehen.
LF010250u	Ich bin der Meinung, dass die Planung von Innovationen an unserer Schule ziemlich unsystematisch geschieht (umgepoltes Item).

1 Die Fragebögen sind im Internet einsehbar: www.schulinspektion.hamburg.de/index.php/article/detail/1006 [11.8.2014]

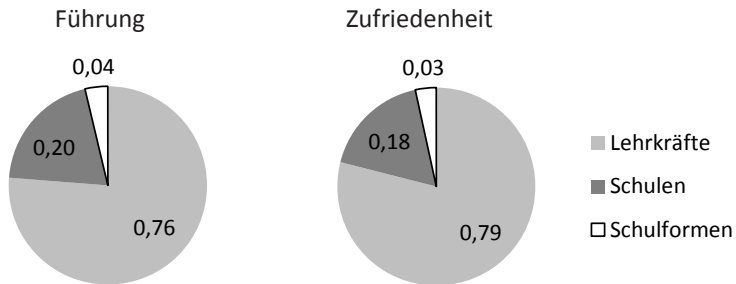
Tabelle 4: Variablen der Skala „Zufriedenheit der Lehrkräfte“

Variablenname	Frageformulierung im Lehrkräftefragebogen
LF010310	Ich bin der Meinung, dass die Räumlichkeiten der Schule und ihr Gelände einladend und freundlich sind.
LF011010	Ich bin zufrieden mit den allgemeinen Arbeitsbedingungen an der Schule.
LF011020	Ich bin zufrieden mit dem Arbeitsklima an meiner Schule.
LF011030	Ich bin zufrieden mit den Kommunikationsstrukturen an der Schule.
LF011040	Ich bin zufrieden mit meinen Möglichkeiten, Ideen zur Entwicklung der Schule einzubringen.
LF011050	Ich bin zufrieden mit meinen Möglichkeiten, berufliche Ziele, die sich auf Schule und Unterricht beziehen, verwirklichen zu können.
LF011060	Ich bin zufrieden mit der fachlichen Unterstützung durch die Kolleginnen und Kollegen.
LF011080	Ich bin zufrieden mit der technischen Ausstattung der Fach- und Unterrichtsräume.
LF011090	Ich arbeite gern an dieser Schule.

Auch die Fragen aus Tabelle 4 konnten auf einer vierstufigen Skala beantwortet werden, wobei der niedrigste Wert wieder für Ablehnung bzw. Nicht-Zustimmung und der höchste für Akzeptanz bzw. Zustimmung stehen. Die über Mittelwertanalysen ermittelten Werte sind entsprechend zu interpretieren. Die Skala zur Zufriedenheit der Lehrkräfte weist eine Reliabilität von $r_{tt(\text{Zufriedenheit})} = 0,859$ (9 Items) auf.

Die Fragen, ob sich die Kollegien in der Wahrnehmung der Führung durch die Schulleitung und in Bezug auf die Zufriedenheit der Lehrkräfte systematisch zwischen den Schulen und darüber hinaus auch noch zwischen den Schulformen unterscheiden, lässt sich anhand der oben vorgestellten *Intra-Class-Correlation* bestimmen. Hier zeigt sich, dass – neben den individuellen Unterschieden in der Wahrnehmung des Führungshandeln und der Zufriedenheit innerhalb der Schulen – auch ein substanzieller Anteil systematisch (signifikant) zwischen den Schulen zu beobachten ist ($ICC_{\text{Führung-Schule}} = 0,20 / ICC_{\text{Zufriedenheit-Schule}} = 0,18$). Wenn auch geringer, so lassen sich auch systematische (signifikante) Effekte zwischen den in der Hamburger Schulinspektion berücksichtigten Schulformen ($ICC_{\text{Führung-Schulform}} = 0,04 / ICC_{\text{Zufriedenheit-Schulform}} = 0,03$) feststellen. Die auf den verschiedenen Ebenen zu beobachtenden Varianzanteile der durch die Lehrkräfte wahrgenommenen Führung und der Zufriedenheit der Lehrkräfte sind Abbildung 4 zu entnehmen.

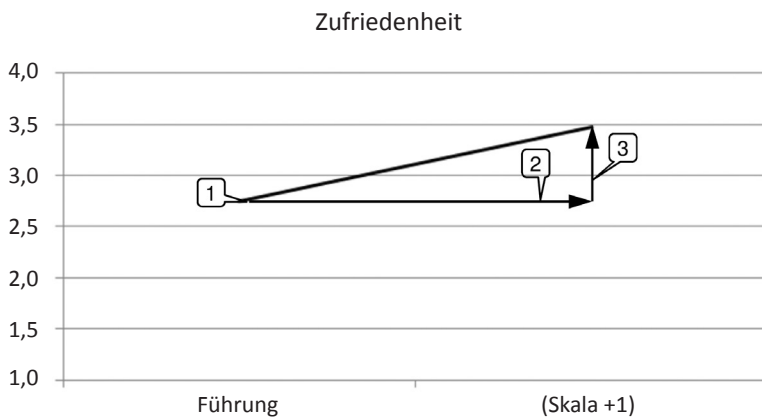
Abbildung 4: Varianzanteile (ICCs) der wahrgenommenen Führung und der Zufriedenheit der Lehrkräfte innerhalb von Schulen, zwischen Schulen und zwischen Schulformen



2. Befunde

In Abbildung 5 ist der generelle Zusammenhang zwischen der von den Lehrkräften wahrgenommenen Führung (Abszisse) und ihrer Zufriedenheit (Ordinate) visualisiert.

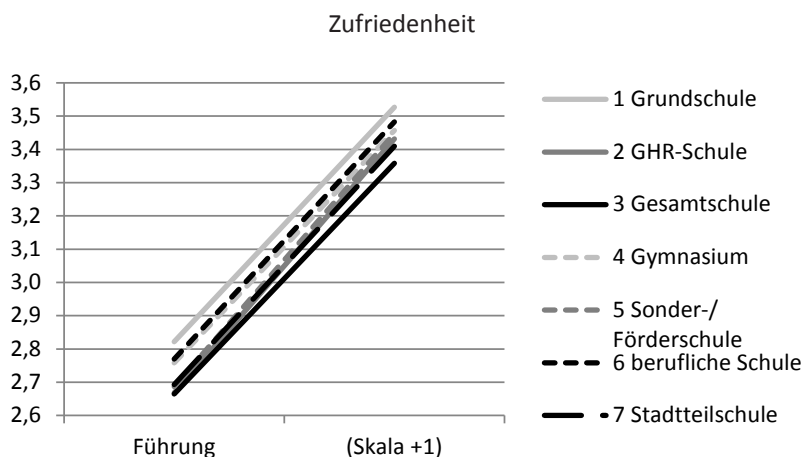
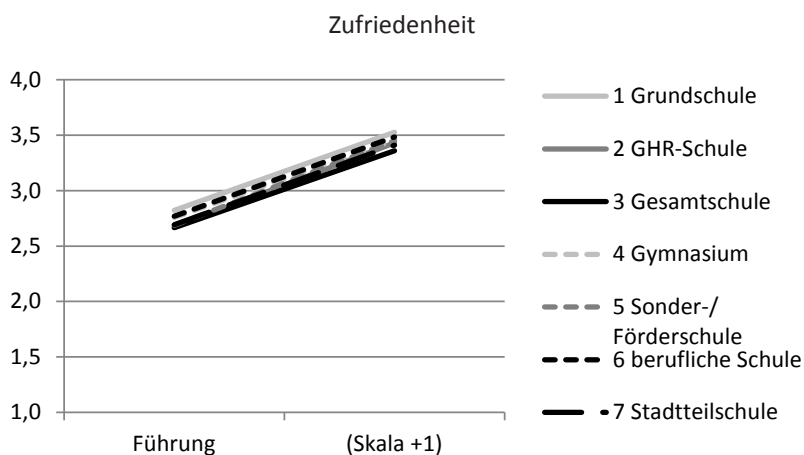
Abbildung 5: Zusammenhang von wahrgenommenem Führungshandeln und Zufriedenheit von Lehrkräften



Bei Lehrkräften, die auf der Skala für die wahrgenommene Führung einen theoretischen Skalenmittelwert aufweisen (2,5), liegt der erwartete Wert für ihre Zufriedenheit bei 2,75 (vgl. ①) und damit etwas über dem erwarteten Skalenmittelwert.

Für Lehrkräfte, für die ein um 1 höherer Skalenmittelwert von 3,5 (vgl. ②) für die Führung ermittelt wird, liegt die erwartete Zufriedenheit mit der wahrgenommenen Führung um 0,72 höher (vgl. ③) – und somit bei insgesamt 3,47 Skalenpunkten. Der im Mittel von den Lehrkräften angegebene Wert für die wahrgenommene Führung liegt bei 3,00 und damit ebenfalls leicht über dem erwarteten, theoretischen Skalenmittelwert – was auf eine tendenziell positive Einschätzung der Führung durch die Schulleitungen hindeutet.

Abbildung 6: Zusammenhang von wahrgenommenem Führungshandeln und Zufriedenheit von Lehrkräften nach Schulformen (Originalskala und Skalenausschnitt)



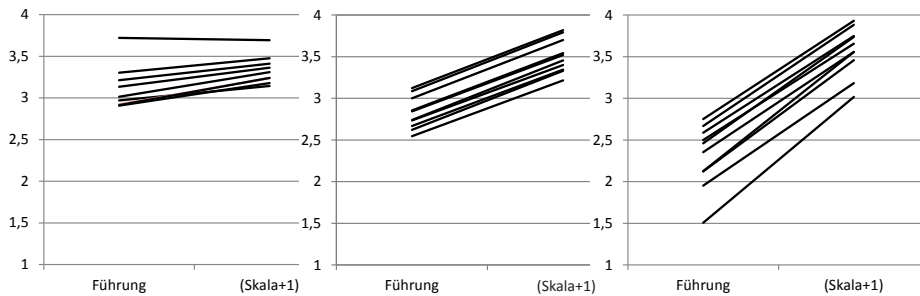
Der Abbildung 5 ist zu entnehmen, dass mit zunehmend positiver Wahrnehmung von Führung durch die Lehrkräfte auch die Zufriedenheit der Lehrkräfte an der Schule zunimmt.

Der generelle Zusammenhang von Führung und Zufriedenheit kann auch auf kleinere Betrachtungseinheiten (Schulformen) heruntergebrochen werden. Diese Differenzierung ist in Abbildung 6 dargestellt. Jede der kleineren Linien stellt hier eine Untergruppe (Schulform) der gesamten dargestellten Daten (hier aus der Hamburger Schulinspektion) dar. Bei diesen und den nachfolgenden Analysen wurden von den 163 Grundschulen 22 Schulen ausgeklammert, wenn weniger als 10 Lehrkräfte den Fragebogen pro Schule beantwortet haben, da es hierbei durch einzelne extreme Antworten zu verzerrten Schätzungen des Schuleffektes kommen kann. In Bezug auf die Lehrkräfte bedeutet dies, dass von den 2.592 befragten Grundschullehrkräften 162 nicht berücksichtigt werden konnten.

In Abbildung 6 (oben) sind wiederum die für die einzelnen Gruppen bestimmten Regressionsgeraden eingezeichnet (siehe den unteren Teil der Abbildung 6, in der die Steigungen vergrößert dargestellt sind). In dem hier gewählten Beispiel verlaufen alle Regressionsgeraden für jede der Schulformen von unten links nach oben rechts und zeigen damit einen (rechnerisch) positiven Zusammenhang zwischen wahrgenommenem Führungshandeln und Zufriedenheit an, auch wenn sich diese zwischen den verschiedenen Schulformen leicht unterscheidet. Je nach Schulform nimmt die Zufriedenheit der Lehrkräfte zwischen 0,69 und 0,76 Punkten zu, wenn sich im Kollegium die wahrgenommene Führung um einen Skalenpunkt erhöht.

Um einen Eindruck zu vermitteln, wie stark sich der Zusammenhang zwischen der wahrgenommenen Führung und der Zufriedenheit der Lehrkräfte jeweils zwischen den Einzelschulen unterscheidet, wurden in Abbildung 7 jeweils die 10 Schulen dargestellt, die den niedrigsten Zusammenhang aufweisen (flacher Verlauf: von $-0,03$ bis $0,34$), die einen mittleren Zusammenhang aufweisen ($0,67$ bis $0,72$) und schließlich die, die den höchsten Zusammenhang (steilster Verlauf: von $1,16$ bis $1,51$) aufweisen.

Abbildung 7: Zusammenhang von wahrgenommenem Führungshandeln und Zufriedenheit von Lehrkräften nach Grundschulen in Gruppen mit niedrigem, mittlerem, starkem Zusammenhang



Wird der globale Eindruck des Zusammenhangs (Abbildung 5) von wahrgenommener Führung und Zufriedenheit den Zusammenhängen je Schulform (Abbildung 6) und schließlich je Grundschule (Abbildung 7) gegenübergestellt, wird deutlich, wie wichtig es ist, verschiedene Analyseebenen zu differenzieren. Denn obwohl die Aussage einmal lautet, der Zusammenhang von wahrgenommenem Führungshandeln und Zufriedenheit ist positiv, lautet sie bei anderer Betrachtung, dass es durchaus Bildungseinrichtungen gibt, in denen dieser Zusammenhang nicht so (deutlich) ausgeprägt ist. Trotz dieses Widerspruchs: Beide Aussagen sind richtig – sie beziehen sich nur auf unterschiedliche Ebenen des Bildungssystems und haben damit auch unterschiedliche Reichweiten bei der Interpretation.

3. Zusammenfassende Betrachtung

Die vorgestellten Befunde zum Zusammenhang der von den Lehrkräften wahrgenommenen Führung durch die Schulleitung und ihrer Zufriedenheit lassen sich analog den vorgestellten Bezugsnormorientierungen interpretieren. Wird ein globaler Zusammenhang (über alle Schulen und Lehrkräfte hinweg) ermittelt, stellt sich die normative Frage, wie stark die Zufriedenheit der Lehrkräfte mit dem Führungshandeln der Schulleitungen zusammenhängen darf oder soll. Ist ein ausgeprägter Zusammenhang gut oder ist er eher nachteilig? Was bedeutet der festgestellte Zusammenhang für die Aus- und Weiterbildung von Schulleitungen in Bezug auf Fragen der Führung? Und noch globaler gefragt: Welchen Stellenwert nimmt die Zufriedenheit der Lehrkräfte im gesamten Kanon von Zielgrößen innerhalb der Schule ein? Ist die Zufriedenheit bereits Selbstzweck oder Mittel zum Zweck? Das Modell von Leithwood zeigt, dass die Zufriedenheit der Lehrkräfte sich, vermittelt

über ihr unterrichtliches Handeln, auf die Entwicklung der Schülerinnen und Schüler auswirkt – und somit auf *das* zentrale Ziel für alle Akteure im Schulsystem zielt.

Der Vergleich zwischen den Schulformen zeigt, dass es kaum Unterschiede zwischen den Schulformen in der Stärke des Zusammenhangs zwischen Schulleitungshandeln und Zufriedenheit der Lehrkräfte gibt; auch das Durchschnittsniveau der Zufriedenheit der Lehrkräfte unterscheidet sich zwischen den Schulformen nur um maximal 0,28 Punkte (also um rund einen Viertel-Skalenpunkt).

Lenkt man den Blick auf die individuelle Bezugsnorm, rückt im Analogieschluss die Einzelschule in das Zentrum der Betrachtung. Bei dieser Betrachtung bekommen die innerschulischen Differenzen im Zusammenhang zwischen Führung und Zufriedenheit eine ganz andere Bedeutung. Zum einen ist bei der Betrachtung zu hinterfragen, warum die Wahrnehmung der von der Schulleitung praktizierten Führung durch die Lehrkräfte sich im Kollegium (deutlich) unterscheidet und warum sich die wahrgenommene Führung in einigen Schulen stark auf die Zufriedenheit der Lehrkräfte auswirkt und in anderen (gar) nicht. Hier helfen Vergleiche (im Sinne einer sozialen Bezugsnorm) mit anderen Schulen, da bei der Erklärung dieser Merkmale eine Reihe von weiteren Merkmalen betrachtet werden muss, die bei dieser Analyse (noch) nicht berücksichtigt sind.

Die zentrale Erkenntnis dieses Beitrags ist, dass es keine bessere oder schlechtere Methode bei der Analyse von Daten aus den Schulinspektionen gibt. Vielmehr ist darauf zu achten, dass die der Fragestellung angemessene Methode auszuwählen ist. Hierbei ist, wie im vorgestellten Datenbeispiel abschließend angemerkt, nicht nur die eigentliche Fragestellung für sich in den Blick zu nehmen, sondern auch die Strukturmerkmale, die potenziell einen Einfluss auf den analysierten Befund nehmen können bzw. mit diesem in einem inhaltlichen Zusammenhang stehen.

Welche Implikationen ergeben sich hierdurch für die zukünftige Erstellung von Schulinspektionsberichten und den zukünftigen Umgang mit Schulinspektionen?

Abhängig von der Zielvariablen muss im Vorwege überlegt werden, ob die beobachteten Merkmale an einer normativen (kriterialen) Vorgabe gemessen werden, ob sie dazu herangezogen werden sollen, im Vergleich mit anderen Bildungseinrichtungen relative Stärken und auch Schwächen zu erkennen, oder ob es schließlich darum gehen soll, innerschulische Zusammenhänge zu analysieren. Die Frage, welche Analysemethode für die jeweils im Fokus stehende Fragestellung angemessen ist, gilt unabhängig von der Bezugsnormorientierung sowohl für eine Inspektions-interne Auswertung als auch für Auswertungen im Rahmen von (wissenschaftlichen) Re-Analysen der Hamburger Schulinspektionsdaten.

Kriteriale bzw. Normvorstellungen für das Bildungssystem sind gesellschaftlich geprägt. Möchte man an dieser Stelle analytisch bei der Schulinspektion anset-

zen, sind konsensual Ziele bzw. (z. B. Verhaltens-)Standards zu setzen, die dann als Zielgrößen für das gesamte Bildungssystem anzuerkennen sind. Bei dieser Form der Auseinandersetzung mit dem Bildungssystem spielen die für das Hamburger Schulsystem verfassten Bildungsberichte eine wichtige Rolle, da sie zum einen Bestandsanalysen vorlegen, aber auch Entwicklungen, die durch gesellschaftliche Erwartungen und auch durch sich wandelnde Voraussetzungen der verschiedenen Akteure im Bildungssystem beeinflusst sind, aufgreifen und somit einer konstruktiven Diskussion zuführen. Darüber hinaus könnten Rückmeldungen an Schulen genutzt werden, um im Vergleich mit anderen Schulen eine Bestandsanalyse zu machen. Hierbei ist es für Schulleitungen, Lehrkräfte, Eltern, Schülerinnen und Schüler hilfreich, Informationen der eigenen Schule im Vergleich mit anderen (ähnlichen) Schulen zu erhalten. So könnten z. B. anhand der Analyse von „erfolgreichen“ Schulen Hinweise auf gute Praxis herausgearbeitet werden, die ggf. für die Entwicklung der eigenen Schule genutzt werden können. Intellektuell problematisch, wenn auch analytisch sinnvoll ist es, auch einen vergleichenden Blick auf die Schulen zu werfen, die nicht erfolgreich sind. Problematisch wird deren Identifizierung jedoch, wenn damit eine Stigmatisierung einhergeht. Dennoch erscheint die Identifizierung auch „problematischer“ Schulen notwendig, um überhaupt die Chance zu haben, diesen Unterstützung seitens anderer Schulen oder der Behörde zukommen zu lassen. Auch wenn die Identifizierung nicht erfolgreicher Schulen für einzelne Personen oder Personengruppen belastend sein dürfte, so ist doch das gesellschaftlich vereinbarte Ziel, dass es zentral um das Wohl und die positive Entwicklung von Schülerinnen und Schülern geht, in den Vordergrund zu stellen. Gerade im Zusammenhang mit solchen Schulen erscheint es sinnvoll, diese zu identifizieren und mit Unterstützung durch geschulte Moderatoren den Beratungsprozess begleiten zu lassen, damit (komplexe) statistisch ermittelte Daten nicht falsch oder überinterpretiert werden.

Von einer Rückmeldung oder Besprechung von individuellen Antworten im schulischen Kontext ist abzuraten. Durch das Wissen, dass die im Rahmen der Inspektion gestellten Fragen individuell ausgewertet werden, kann es zur Orientierung an sozialen Normen und damit zur Beantwortung der Fragen gemäß der sozialen Erwünschtheit kommen. Dies kann nicht im Interesse der Schulinspektionen sein. Auch ohne die Anonymität aufzuheben, kann der Fundus an Informationen, der in den Einzelinformationen liegt, genutzt werden. Hierbei können geschulte Moderatoren helfen, die Heterogenität im Kollegium zu analysieren, ohne die Anonymität der Einzelnen aufzuheben. So könnten wichtige Informationen (z. B.: Warum wird Führung durch die Schulleitung so unterschiedlich wahrgenommen?) genutzt werden, um aktive Schulentwicklung zu betreiben.

Ein weites Forschungsfeld öffnet sich, wenn der Frage nachgegangen wird, wie Schulleitungen auf das Schulleitungshandeln vorzubereiten sind und – noch globaler angesetzt – was überhaupt Erwartungen an „gutes“ Schulleitungshandeln

ausmacht. Der Blick auf internationale Schulvergleichsuntersuchungen zeigt, dass es hier durchaus ganz verschiedene Vorstellungen betreffend des Schulleitungshandelns gibt und dass mit diesen unmittelbar auch die jeweilige Aus- bzw. Weiterbildung verknüpft ist. Hier reicht die Bandbreite von Lehrkräften mit erweiterten Verwaltungsaufgaben bis hin zu Managern von Bildungseinrichtungen (vgl. z. B. Lankes & Carstensen 2007).

Ein bisher durch Schul- und Unterrichtsforschung wenig systematisch untersuchtes Forschungsdesiderat besteht in der Auseinandersetzung mit der Frage, was die Erwartungen der verschiedenen Akteure im Bildungssystem in Bezug auf die Inspektion sind. Diese Frage erscheint für die Schulinspektionen substanziell, da sie den Blick auf die wichtigen und vor allem richtigen Stellen im Bildungssystem lenkt. Neben sowohl normativ als auch bildungspolitisch zu führenden Diskussionen bieten die bisher erhobenen Daten aus Schulinspektionen einen Fundus für eine ganze Reihe von Antworten – wobei die Herausforderung darin besteht, die richtigen Fragen zu stellen.

Literatur

- Achtenhagen, F. (1981). Mehrebenenanalyse in der Unterrichtsforschung. *Unterrichtswissenschaft*, 4, 319–336.
- Adams, R. J. & Wu, M. L. (2000). *PISA 2000 Technical Report*. Verfügbar unter: www.pisa.oecd.org [09.08.2014]
- Behörde für Bildung und Sport (2006). *Orientierungsrahmen – Qualitätsentwicklung an Hamburger Schulen*. Hamburg: Behörde für Bildung und Sport.
- Cochran, W. G. (1972). *Stichprobenverfahren*. Berlin: de Gruyter.
- Cronbach, L. J. (1976). *Research on classrooms and schools. Formulation of questions, design, and analysis*. Stanford: (Occasional Paper).
- Ditton, H. (1998). *Mehrebenenanalyse*. Weinheim: Juventa.
- Frankenberg, E. (2006). *The Segregation of American Teachers*. Verfügbar unter: <http://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/the-segregation-of-american-teachers/frankenberg-orfield-segregation-american-teachers.pdf> [09.08.2014]
- Institut für Bildungsmonitoring (2009). *Jahresbericht der Schulinspektion, Hamburg 2008*. Hamburg: Behörde für Schule und Berufsbildung.
- Institut für Bildungsmonitoring (2011). *Jahresbericht der Schulinspektion, Hamburg 2009–2010*. Hamburg: Behörde für Schule und Berufsbildung.
- Institut für Bildungsmonitoring und Qualitätsentwicklung (2012). *Jahresbericht der Schulinspektion, Hamburg 2010–2011*. Hamburg: Behörde für Schule und Berufsbildung.
- Janke, N. (2006). *Soziales Klima an Schulen aus Lehrer-, Schulleiter- und Schülerperspektive*. Münster: Waxmann.

- Kish, L. (1995). *Survey sampling*. New York: John Wiley & Sons (Nachdruck).
- Krapp, A. & Weidenmann, B. (Hrsg.) (2001). *Pädagogische Psychologie* (4. Aufl.). Weinheim: Beltz PVU.
- Lankes, E.-M. & Carstensen, C. H. (2007). Der Leseunterricht aus der Sicht der Lehrkräfte. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 161–193). Münster: Waxmann.
- Martin, M. O., Mullis, I. V. S. & Kennedy, A. M. (2007). *PIRLS 2006 Technical Report*. Chestnut Hill, MA: Boston College.
- Martin, M. O., Rust, K. & Adams, R. J. (Hrsg.) (1999). *Technical standards for IEA studies*. Amsterdam: International Association for the Evaluation of Educational Achievement.
- Pietsch, M. (2014). Wirksame Schulleitungen. Annahmen, Befunde und der Blick der Schulinspektion. *Hamburg macht Schule*, 1 (26), 6–11.
- Rheinberg, F. (2001). Bezugsnormen und schulische Leistungsbeurteilung. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 59–71). Weinheim: Beltz.
- Schwippert, K. (2002). *Optimalklassen: Mehrebenenanalytische Untersuchungen*. Münster: Waxmann.
- Schwippert, K. (2008). Interpretationen und Fehlinterpretationen von Daten: Mehrebenenanalysen in der Schul- und Unterrichtsforschung. In F. Hellmich (Hrsg.), *Lehr-Lernforschung und Grundschulpädagogik* (S. 157–167). Bad Heilbrunn: Klinkhardt.
- Treiber, B. (1980). *Qualifizierung und Chancenausgleich in Schulklassen. Teil 1: Theorien, Methoden, Ergebnisse*. Frankfurt a. M.: Lang.

Wie Schulinspektor/innen fragen und Grundschüler/innen antworten

Barbara Scholand unter Mitarbeit von Stephanie Graw

Zusammenfassung

In der empirischen Bildungsforschung finden Interviews mit Schülerinnen und Schülern, die im Rahmen von Schulinspektionen zur Datengewinnung über Schule und Unterricht genutzt werden, bislang wenig Beachtung. Für diesen Beitrag wurde eine Stichprobe von sechs Interviews, die an Grundschulen durchgeführt wurden, aus dem Datensatz des 1. Zyklus der Hamburger Schulinspektion gezogen. Es werden – unter Einbezug der Antworten der Schülerinnen und Schüler – die Einstiegs- und die Abschlussfragen im Hinblick darauf untersucht, welche inspektorenseitigen sprachlichen Äußerungen im Kontext der Grundschule förderlich bzw. weniger förderlich sind, um die Ziele und Zwecke des Interviews – die Evaluation von Schulqualität – unter der Bedingung eines relativ engen zeitlichen Rahmens zu erreichen. In methodischer Hinsicht orientiert sich die Studie an der funktional-pragmatischen Diskursanalyse (Ehlich & Rehbein 1986), da diese im besonderen Maße geeignet scheint, um die Umsetzung institutioneller Zwecke in sprachliche Handlungen zu analysieren.

„There is a growing recognition that young people have a right to be heard and have something worthwhile to say about their school experiences. [...] How to listen and learn, as well as to teach and lead, is the challenge for teachers, schools and their communities.“ (Teaching and Learning Research Briefing No. 5, University of Cambridge, June 2003)

Das obige Zitat ließe sich in Bezug auf das Instrument „Interview“ der Schulinspektion ergänzen: „*How to ask and listen is the challenge for school inspectors.*“ In der empirischen Bildungsforschung finden Interviews mit Schülerinnen und Schülern, die im Rahmen von Schulinspektionen zur Datengewinnung über Schule und Unterricht genutzt werden, bislang wenig Beachtung. Für diesen Beitrag wurde eine Stichprobe von sechs Interviews, die an Grundschulen durchgeführt wurden, aus dem Datensatz des 1. Zyklus der Hamburger Schulinspektion gezogen: Grundschulen deshalb, weil Interviews in diesem Bereich für Schulinspektorinnen und -inspektoren insofern herausfordernd sein können, als sie sich auf kindliche Sprach- und Vorstellungswelten sowie ein geringeres Abstraktionsniveau einlassen müssen.

Es wird jeweils – unter Einbezug der Antworten der Schülerinnen und Schüler – die Einstiegs- und die Abschlussfrage im Hinblick darauf untersucht, welche inspektorenseitigen sprachlichen Äußerungen im Kontext der Grundschule förderlich bzw. weniger förderlich sind, um die Ziele und Zwecke des Interviews – die Evaluation von Schulqualität – unter der Bedingung eines relativ engen zeitlichen Rahmens zu erreichen. Ein „Nebenprodukt“ sind Erkenntnisse darüber, welche Themen Grundschulkinder beschäftigen und wie sie diese zur Sprache bringen. In methodischer Hinsicht orientiert sich die Studie an der funktional-pragmatischen Diskursanalyse (Ehlich & Rehbein 1986), denn „[i]m Mittelpunkt des Untersuchungsinteresses der Funktionalen Pragmatik steht die Frage nach dem Zusammenhang von sprachlichem Handeln, sprachlicher Form und gesellschaftlichen, vor allem institutionellen Strukturen und Zwecken.“ (Weber & Becker-Mrotzek 2012, S. 2). Vor der Analyse der Fallbeispiele werden zunächst Begründungszusammenhänge von Schülerinterviews aufgezeigt und die Einbettung der Interviews im Verfahren der Schulinspektion erläutert. Abschließend erfolgt eine Diskussion der Ergebnisse vor dem Hintergrund der Funktion von Schulinspektion.

1. Interviews mit Schüler/innen: Begründungszusammenhänge

Im Rahmen der Schulinspektion als externer Evaluation stehen unterschiedliche Gruppen im Fokus der Datenerhebung. Die Schülerinnen und Schüler einer Schule bilden eine dieser Gruppen. Es war von Beginn an selbstverständlich, dass alle Schülerinnen und Schüler ab der dritten Klasse die Möglichkeit haben sollen, an einer standardisierten Online-Befragung teilzunehmen und zusätzlich eine Auswahl an Schüler/innen (s. u.) mittels Interview befragt wird. Diese Selbstverständlichkeit hat unterschiedliche Bezüge:

1. Normative Grundlage ist die UN-Kinderrechtskonvention von 1989, durch die Kinder als Akteure eigenen Rechts etabliert wurden. Dies hat seinen Niederschlag in den Schulgesetzen der Länder gefunden. Damit hat Schule nicht nur die Aufgabe, Wissen über Kinderrechte zu vermitteln, sondern die Verpflichtung, die Partizipation von Schülerinnen und Schülern sowie deren Recht auf Chancengleichheit in der Bildung konsequent umzusetzen (vgl. Hamburgisches Schulgesetz 2014, insbes. § 3). Entsprechend hat die Schulinspektion die Verpflichtung übernommen, Schülerinnen und Schüler zu bestimmten Aspekten von Schulqualität zu befragen bzw. anzuhören.
2. Theoretische und empirische Begründungen dafür, dass (Grund-)Schülerinnen und Schüler in die (externe) Evaluation ihrer Schule eingebunden werden sollten, finden sich vermehrt in der Folge des „PISA-Schocks“ im Jahr 2000 in den deutschsprachigen Publikationen zu Schulqualität, Schulentwicklung

und Schuleffektivität. Ditton und Arnoldt (2004, S. 116) untersuchten bspw. die Wirkung von *Feedbackinterventionen*, d. h. die systematische Einholung von Schülerfeedback und dessen Rückmeldung durch Dritte an die Lehrkräfte. Theoretisch schließen der Autor und die Autorin an Modelle der Handlungsregulation an. Im Ergebnis stellen sie eine bedingte Brauchbarkeit des Verfahrens fest.¹ Seitens der Schülerinnen und Schüler liegt der Vorteil einer solchen Feedbackintervention – wie sie auch durch die Schulinspektion gegeben ist – gegenüber einem direkten Feedback an die jeweilige Lehrkraft in der Anonymität und Nicht-Sanktionierbarkeit begründet.

3. Aus einer praxisbezogenen Perspektive plädieren z. B. Nietzsche und Vieluf (2006) dafür, „Schritt für Schritt eine Evaluationskultur aufzubauen“ (S. 34), die sowohl Schülerfeedback als auch Lernstandserhebungen für die Weiterentwicklung der pädagogischen Arbeit nutzt. Sie zeigen am Beispiel der damaligen Gesamtschule Bergedorf, dass die schriftliche Befragung von Schülerinnen und Schülern zwar „zu vorläufigen Interpretationen“ taugt, sich aber „viele [...] Fragen [...] erst in Gesprächen mit den Schülerinnen und Schülern klären“ lassen (ebd., S. 32). Dieser Hinweis stützt das Verfahren der Schulinspektion, auf Grundlage der schriftlichen Schülerbefragung zu vorläufigen Einschätzungen zu kommen und diese dann anhand der Interviews zu überprüfen und vor allem zu ergänzen.

Im Qualitätsleitbild der Schulinspektion (s. Beitrag von Ulrich in diesem Band) finden sich diese drei Perspektiven vereint. Dort verpflichtet sich die Schulinspektion zuallererst der *Verbesserung der Lernchancen von Schülerinnen und Schülern* – und folgt damit ihrem gesetzlichen Auftrag, durch eine sowohl kontroll- als auch entwicklungsorientierte Evaluation ihren Beitrag zur schulischen Qualitätsentwicklung und -sicherung zu leisten. Die Instrumente Schülerbefragung und Schülerinterview liefern Erkenntnisse zur Binnenperspektive der jeweiligen Schule. Diese sind nicht nur eine notwendige Ergänzung zu den Sichtweisen von Schulleitung, Lehrkräften und Eltern, sondern unterfüttern sowohl die standardisierte Schülerbefragung als auch die Unterrichtsbeobachtungen.² Kinder im Alter von acht bis zehn Jahren können als ernstzunehmende Akteure in Evaluationsverfahren betrachtet werden, denn sie sind in der Lage, schulische Geschehnisse differenziert wahrzunehmen sowie Fragen dazu zu beantworten.³ Wie das Instrument „Schüler-

1 Ditton und Arnoldt stellen – nicht unerwartet – fest: „Durchgängig besteht eine bedeutsame Korrelation zwischen der Akzeptanz der [Feedback-]Erhebung sowie des Rückmeldeverfahrens und dem erreichten Ergebnis.“ (Ditton & Arnoldt 2004, S. 135)

2 Im Verfahren der Schulinspektion wird triangulierend vorgegangen, d. h., es werden mehrere Datenquellen und verschiedene Perspektiven einbezogen (vgl. IfBQ 2013, S. 22–23).

3 Aus Studien an Grundschulen weiß man, dass sich bereits in diesem Alter z. B. „Einstellungstypen“ zu Schule und Lernen (vgl. Christen 2004) sowie Interessen (vgl. Pruisken 2005) herausbilden.

interviews“ im ersten Zyklus der Hamburger Schulinspektion praktisch angewandt wurde, soll zunächst anhand seiner Einbettung im Verfahren erläutert werden.

2. Das Schülerinterview im Verfahren der Hamburger Schulinspektion

Im Ablauf der Inspektion erfolgt das Interview mit einer Gruppe von Schülerinnen und Schülern gegen Ende des zwei- bis dreitägigen Schulbesuchs (s. Diedrich in diesem Band). Zuvor haben bereits Analysen schulischer Dokumente, Unterrichtsbeobachtungen sowie Interviews mit der Schulleitung und mit Vertretern der Eltern- und der Lehrerschaft stattgefunden.

In der Regel findet das Interview in einem von der Schule zugewiesenen Raum statt, meist ein Klassenzimmer. Die teilnehmenden Schülerinnen und Schüler kommen aus verschiedenen Klassen und Jahrgangsstufen. An Grundschulen gilt, dass jeweils acht bis maximal zwölf Schülerinnen und Schüler der dritten und vierten Klassen von der Schulleitung für das Interview ausgewählt werden. Die Schülerinnen und Schüler und deren Eltern werden vorab über den Zweck informiert und gebeten, der Teilnahme am Interview schriftlich zuzustimmen. Dafür erhält die Schule von der Schulinspektion entsprechende Musteranschreiben. Die Schülerinnen und Schüler wissen also von der Anwesenheit und der Bedeutung der Schulinspektorinnen und -inspektoren als Prüfungsinstanz. Sie haben die Inspektionspersonen (IP)⁴ meistens bereits vor dem Interview im Unterricht oder im Gebäude gesehen.

Der Leitfaden des Schülerinterviews orientiert sich prinzipiell an den Vorgaben des Orientierungsrahmens für Qualitätsentwicklung an Hamburger Schulen von 2006 (s. den Beitrag von Diedrich zum zweiten Zyklus der Schulinspektion in diesem Band). Die Schülerinnen und Schüler werden vor allem zu den Bereichen Bildung und Erziehung befragt.⁵

4 Zur Sicherung der Anonymität wird hier und im Folgenden von Inspektionsperson/en die Rede sein. Dementsprechend werden weibliche Pronomen verwendet – die jedoch keinen Rückschluss auf das Geschlecht der IP erlaubt.

5 Zur Funktionsbestimmung von Interviews in der Schulinspektion vgl. Tosana et al. 2011, S. 58 ff.

Beispiel einer Intervieweröffnung

Die folgende Abbildung zeigt einen exemplarischen Intervieweinstieg⁶, wie er von einer IP zur Vorbereitung auf das Interview entwickelt wurde:

Abbildung 1: Beispiel einer Eröffnungssequenz im Schülerinterview aus dem 1. Zyklus der Hamburger Schulinspektion (2007)

Setting

Dauer: 45 Minuten

Raum: Raum des Inspektionsteams an der Schule

Teilnehmende: Inspektoren; max. 12 Schülerinnen und Schüler (Schülerrat)

Rollen Inspektoren: Interviewführer, Protokollant;
zusätzlich kann auch vereinbart werden, dass neben dem Interviewführer auch andere nachfragen

Warm-up

- Begrüßung
- Dank für Bereitschaft/Bedeutung des Interviews
 - Ihr seid die, um die es hier geht. Schule ist für Schüler da.
 - Wir sind dafür da, dass Schule besser werden kann.
 - Kennt das sicher selbst: Manchmal kann ein anderer viel besser sehen, ob ich etwas gut hinkriege oder nicht. „Mensch, schau mal, da vergisst du immer beim T den Strich zu machen“ und plötzlich sieht man es auch. Genauso gucken wir die Schule an, um zu sehen, was vielleicht nicht so gut läuft.
 - Aber wir wollen auch hören, was an dieser Schule so richtig gut ist.
- Vertraulichkeit: keine Einzelaussagen im Bericht
- Ziele/Themen:
 - Was ihr gut findet an eurer Schule, was nicht.
 - Wie euer Unterricht abläuft.
 - Wir werden euch zu eurer Meinung fragen. Es gibt kein Richtig und Falsch. Jeder hat seine eigene Meinung.
- Dauer 45 Minuten
- Funktionen der Interviewpartner
- Art des Interviews: Informationsgewinn, keine Zeit für Diskussionen

6 Diese erste Begrüßung der Schülerinnen und Schüler wird durch die Audioaufnahmen regelhaft nicht dokumentiert, daher nutzen wir Unterlagen einer IP und danken ihr an dieser Stelle für die freundliche Überlassung.

Wie die Abbildung zeigt, unterstützen Layout und Gliederung die schnelle Erfassung der Inhalte. Es werden unter „Setting“ zunächst einige Rahmenbedingungen festgehalten: „Dauer“, „Raum“, „Teilnehmende“ und „Rollen“. Die Überschrift „Warm-up“ deutet an, dass es sich zu Beginn des Interviews zunächst um eine „kühle“ Situation handelt, die durch die nachfolgend dokumentierten Ansagen „erwärmt“ werden soll. Es folgen teils Stichworte, teils Ausformulierungen zur sprachlichen Gestaltung des Intervieweinstiegs. „Begrüßung“ und „Dank“ organisieren den Kommunikationsaspekt „Höflichkeit“ und sichern die Einhaltung sozialer Gesprächsnormen. Damit wird einer Forderung Helfferichs (2011, S. 177) entsprochen, die schreibt: „In [der] Eingangsphase ist eine offene und freundliche Atmosphäre herzustellen und die Bereitschaft, an einem Interview teilzunehmen, ist zu würdigen.“ Auf die „Bedeutung des Interviews“ wird in Form einfacher (Haupt-)Sätze verwiesen, die das hörerseitige Verstehen fördern: „Ihr seid die, um die es hier geht“ und „Schule ist für Schüler da“. Inhaltlich wird damit den anwesenden Schülerinnen und Schülern ein Anspruch darauf zugesichert, dass Schule sich ihnen nähert und anpasst. Als Zweck kann hier das Werben um Kooperation gesehen werden. Die Ansagen zielen auf die Etablierung eines guten Gesprächsklimas und die Auskunftsbereitschaft der Schülerinnen und Schüler; sie entkräften eventuelle Annahme seitens der Schülerinnen und Schüler, es handle sich um eine Situation, in der *sie* geprüft würden. Die folgende Proposition „Wir sind dafür da, dass die Schule besser werden kann“ sowie das konkrete, an Erfahrungen der Schülerinnen und Schüler anknüpfende Beispiel mit dem fehlerhaften „T“ und die daran anschließenden Sätze umreißen die Funktion der Schulinspektion: Sie ist „ein anderer[, der] viel besser sehen kann“, nämlich „was vielleicht nicht so gut läuft“, aber auch, „was [...] so richtig gut ist.“ Die Zusicherung von Vertraulichkeit („keine Einzelheiten im Bericht“) und Meinungsfreiheit („es gibt kein Richtig und Falsch, jeder hat seine eigene Meinung“), die Benennung der Ziele, Themen („was ihr gut findet an eurer Schule, was nicht; wie euer Unterricht abläuft“) sowie die Angabe der Dauer des Interviews konkretisieren den Rahmen und sorgen für Transparenz. Zugleich orientieren diese Stichpunkte die Interviewten wie die Interviewenden selbst in der gegebenen Konstellation⁷ und stimmen die Schülerinnen und Schüler im Hinblick auf das, was sie erwartet bzw. was von ihnen erwartet wird, ein. Der Punkt „Funktionen der Interviewpartner“ verweist auf einen wichtigen Schritt des Kennenlernens, nämlich die Vorstellung der Schülerinnen und Schüler, z. B. als Klassensprecherinnen oder Klassensprecher. Erst ganz am Ende erfolgt die notwendige Einschränkung „keine Zeit für Diskussionen“ und die zu Inter-

7 Mit „Konstellation“ sind in der Funktionalen Pragmatik (institutionell) wiederkehrende Kommunikationssituationen gemeint, deren Handlungsbedingungen durch soziale Strukturen und gesellschaftliche Zwecke vorstrukturiert sind. Siehe dazu den nachfolgenden Abschnitt zu Bedingungen und möglichen Schwierigkeiten von Schülerinterviews.

viewenden werden auf den Interviewzweck „Informationsgewinn“ fokussiert. Die Warm-up-Phase ist im Hinblick auf den weiteren Verlauf und das Gelingen eines Interviews entscheidend (vgl. Heintzel 2010, S. 711) – die vorliegende detaillierte Ausarbeitung und schriftliche Niederlegung tragen dem Rechnung und unterstützen die Entwicklung einer Bewältigungsroutine für die Eröffnung des Interviews.

Bedingungen und mögliche Schwierigkeiten von Schülerinterviews

Für die Beantwortung der Frage dieses Beitrags – Welche sprachlichen Äußerungen eignen sich im Kontext von Grundschulen für den Inspektionszweck? – müssen hier auch die Bedingungen und möglichen Schwierigkeiten aufgezeigt werden, denen das Schülerinterview unterliegt. Erst ihre Berücksichtigung ermöglicht es, eine angemessene Analyse und Kritik der empirisch vorgefundenen Einstiegs- und Abschlussfragen zu leisten. Sie sollen hier stichwortartig benannt werden:

- Ort und Zeitpunkt können weder von den interviewten Schulkindern noch von den interviewenden Schulinspektorinnen und -inspektoren frei gewählt werden, die Phase zum Aufbau einer gelingenden Kommunikation ist sehr kurz.
- Das Interview ist zeitlich und inhaltlich begrenzt. Die IP stellen ihre Fragen in evaluativer Absicht, d. h. sie hören die Antworten der Kinder unter der Perspektive, was diese für die Bewertung eines bestimmten Kriteriums von Schulqualität bedeuten und stellen entsprechende Nachfragen. Dabei entgehen ihnen eventuell Antworten, die nicht in das Leitfadenschema passen, aber dennoch den Gesamteindruck von der Schule bereichern können (vgl. Tosana et al. 2011, S. 63).
- Das Interview findet in der Regel in einem Klassenraum statt, der typische schulische Ordnungs- und Handlungsmuster aufruft: Hier befragen üblicherweise erwachsene Lehrkräfte die Schülerinnen und Schüler mit der Absicht, zu einer Beurteilung ihrer Leistungen zu gelangen. Dies kann dazu führen, dass Schülerinnen und Schüler Antworten geben, die sie für „richtig“ im Sinne sozialer Erwünschtheit halten.
- Im Interview werden bestimmte Sachverhalte abgefragt, es ähnelt daher einer Testsituation. Die Schülerinnen und Schüler wissen, dass das Verfahren der Schulinspektion für ihre Schule eine Prüf- bzw. Kontrollsituation darstellt. Situations- und/oder interessenbedingt können ihre Aussagen einen positiven oder negativen Bias aufweisen.⁸

⁸ Der erste Jahresbericht der Schulinspektion (BSB 2008, S. 20) stellt hierzu fest: „So spielt beispielsweise das Alter der befragten Schülerinnen und Schüler bei der Einschätzung der einzelnen Schule eine große Rolle. Es ist bekannt, dass jüngere Schülerinnen und Schüler generell positiver antworten als ältere Schülerinnen und Schüler.“

- Auch die soziale Kontrolle durch die Peergroup kann in der Situation des Gruppeninterviews Anpassungen im Antwortverhalten seitens der Schülerinnen und Schüler hervorrufen.
- Die Inspektorinnen und Inspektoren gehen mit einem erheblichen Vorwissen in die Interviewsituation hinein – d. h. sie stehen vor der Herausforderung, dieses Wissen zwar für die Formulierung von Fragen nutzbar zu machen, dabei jedoch die notwendige Offenheit zu wahren.
- Die Interessen der Fragenden und die der Antwortenden können divergieren: Was die Schulinspektorinnen und -inspektoren erfahren möchten, entspricht nicht unbedingt dem, was die Schülerinnen und Schüler möglicherweise erzählen möchten.
- Inspektor/innen wie Schüler/innen sind von der Kooperation des Gegenübers abhängig: Es handelt sich um ein Gruppeninterview, in dem die oder der Fragende darauf angewiesen ist, dass Antwortwillige sich melden; die Befragten wiederum sind darauf angewiesen, „drangenommen“ zu werden, wenn sie sich äußern möchten.

Methodischer Zugang, Fragestellung und Analysekriterien

Im Anschluss an Ehlich und Rehbein (1986) kann das Schülerinterview als Sonderfall institutioneller, schulischer Kommunikation angesehen werden. Die von eben diesen Autoren entwickelte funktional-pragmatische Diskursanalyse richtet ihr Augenmerk auf die Zwecke, die institutioneller Kommunikation zugrunde liegen (vgl. Galinski 2004, S. 16 ff.; Weber & Becker-Mrotzek 2012, S. 2 ff.).⁹ Schulinspektion dient dem Zweck, Schul- und Unterrichtsqualität festzustellen. Dazu bedient sie sich u. a. des Instruments „Interview“. In Interviews der Schulinspektion sind regelhaft wiederkehrende sprachliche Handlungsmuster aufzufinden, die jedoch nicht mit einem standardisierten Ablauf gleichzusetzen sind; vielmehr steht der Begriff „Handlungsmuster“ in unmittelbarem Zusammenhang mit dem

9 Zum Verständnis: „Die funktional-pragmatische Diskursanalyse bedient sich bei ihren Untersuchungen qualitativer Methoden und geht empirisch-induktiv vor. Das heißt, sie unterzieht Transkripte, die auf der Grundlage von akustischen oder audiovisuellen Gesprächsaufzeichnungen angefertigt wurden, einer hermeneutischen materialgeleiteten Analyse. Man kann dieses Vorgehen als reflektierte Empirie ohne starre Analyseschritte bezeichnen, bei der sich empirische Untersuchung und Theoriebildung über sprachliches Handeln gegenseitig bedingen; die Analyse authentischer sprachlicher Interaktion, die Entwicklung theoretischer Erkenntnisse und die erneute Überprüfung des gewonnenen Wissens am Datenmaterial befördern sich wechselseitig [...]“ (Weber & Becker-Mrotzek 2012, S. 3) Eine detaillierte, linguistischen Anforderungen entsprechende Transkription und Analyse kann und soll in diesem Beitrag jedoch nicht geleistet werden.

übergeordneten Kommunikationszweck bzw. den untergeordneten Teilzwecken.¹⁰ In den nachfolgenden Analysen geht es um die von den Inspektorinnen und Inspektoren genutzten sprachlichen Gestaltungsspielräume und die Wahl der konkreten sprachlichen Mittel zur Realisierung des Handlungsmusters „Einstiegsfrage“ bzw. „Abschlussfrage“. Leitend für die Analysen sind die folgenden Fragen: Was wird wie gefragt und welche Antworten folgen darauf seitens der Schülerinnen und Schüler? „Funktionieren“ die Fragen als Interviewfragen, die an Grundschülerinnen und Grundschüler gerichtet sind, und im Hinblick auf den Evaluationszweck, nämlich die Schulqualität zu erfassen? Ergänzend werden die von Tosana et al. (2011, S. 62 ff.) benannten Gütekriterien von Interviews (z. B. Nicht-Beeinflussung, Reichweite)¹¹ mit einbezogen. Für die Analyse gelten außerdem die von Brünner benannten Kriterien für diskursanalytische Studien im Kontext der funktionalen Pragmatik: Angabe der sprachlichen Indizien und ihrer Bedeutung sowie Verzicht auf psychologische Deutungen (vgl. Brünner 2009, S. 63).

3. Fallbeispiele von Schülerinterviews

Das im Folgenden analysierte Material entstammt in etwa der Mitte des ersten Zyklus (Laufzeit 2006–2013). Zu diesem Zeitpunkt war das Inspektionsverfahren bereits etabliert, die Instrumente waren erprobt und man kann davon ausgehen, dass sich Inspektionsroutinen entwickelt hatten. Aus dem Datensatz von rund 80 Schulen, die im Jahr 2009 inspiziert worden waren, wurden unter dem Gesichtspunkt der Kontrastierung (vgl. Merrens 2007) sechs Fälle ausgewählt, die zeitlich nahe beieinander liegen¹². Die Kontrastierung bezieht sich zum einen auf die Schulen – diese sind jeweils auf einer der sechs Stufen des damaligen KESS-Index¹³ (heute: Sozialindex) zu finden, zum anderen sollten sechs¹⁴ Varianten der Interviewführung miteinander verglichen werden, um eine Bandbreite an Äußerungsformen und sprachlichen Vorgehensweisen zu erfassen. In der Analyse der

10 Wie oben anhand der Analyse des „Warm-up“ gezeigt.

11 Tosana et al. (2011) beziehen sich auf Merton & Kendall (1993) sowie Hopf (1978) und führen insgesamt fünf Gütekriterien auf: „Nicht-Beeinflussung“ meint den Verzicht auf suggestive Fragen, „Spezifität“ die Berücksichtigung individueller Sichtweisen, „Tiefe“ die Einschätzung der unterschiedlichen Dimensionen einer Antwort, „Reichweite“ die Erhebung aller relevanten Themen einschließlich neu auftauchender Gesichtspunkte, „personaler Kontext“ die Berücksichtigung des individuellen Hintergrunds.

12 Die Interviews verteilen sich über einen Zeitraum von sechs Wochen.

13 Stufen des KESS-Index: Stufe 1 bedeutet eine stark belastete Lage der Schülerschaft, Stufe 6 eine bevorzugte soziale Lage der Schülerschaft (vgl. Pietsch et al. 2006)

14 In zwei Fällen wurde das Interview nicht von der Teamleitung, sondern der zweiten IP geführt. Es handelt sich daher nur um fünf interviewende Inspektorinnen und Inspektoren, jedoch wählte eine IP in zwei Interviews unterschiedliche Einstiegsfragen.

Fallbeispiele konzentrieren wir uns exemplarisch auf die Einstiegs- und die Abschlussfragen. Eine zusammenfassende Präsentation und Diskussion der Ergebnisse erfolgt in Abschnitt 4.

3.1 Vergleich der Einstiegsfragen

Als „Einstiegsfrage“ haben wir die Frage bestimmt, die nach Beginn der Aufnahme und der Dokumentationszwecken dienenden Ansage „Schule N, Schülerinterview“ formuliert wird. Die folgende Tabelle von Einstiegsfragen zeigt auf den ersten Blick die Vielfalt an (sprachlichen) Vorgehensweisen.¹⁵

Tabelle 1: Einstiegsfragen

IP1 GS A Index 1	Und dann würd' ich euch bitten, euch vorzustellen mit dem Namen, mit der Klasse, in die ihr geht, und, ähm, dass ihr mir sagt, eine Sache oder vielleicht zwei Sachen, die ihr an der Schule besonders toll findet. Möch'st Du anfangen, NS1w?
IP2 GS B Index 2	Dann wüsste ich, äääh, ganz gerne doch noch (.) eine Frage hab ich noch. Es gab ja hier so etwas, äh, wie einen SCHÜlerrat. Wer (.) is' jemand von euch in dem Schülerrat?
IP3 GS C Index 3	So. (--) Ähm, als erstes möchte ich gerne von Euch wissen, äh, wie kommt es dazu, dass grade IHR jetzt (.) hier (.) heute zu diesem Interview gekommen seid. (--) Wer hat das bestimmt und (--) wer hat das ausgesucht und wie habt ihr überhaupt davon erfahren.
IP4 GS D Index 4	Als erstes möchte ich gerne wissen, ihr habt (.) ich hab gehört, dass ihr ne (.) Kinderkonferenz (.) hier habt (.) an der Schule. Seid ihr da schon mal dran beteiligt gewesen?
IP5 GS E Index 5	Dann möchte ich beginnen mit der Frage, und ich bitte euch dann, euch zu melden, was findet ihr denn an der Schule GANZ besonders TOLL. NSw1, du warst die Erste.
IP6 GS F Index 6	Dann würd ich jetzt ganz gerne als Erstes mal die Kinder der KINDERkonferenz fragen, ob ihr mir kurz erzählen könnt, womit ihr euch im LETZTEN Schuljahr befasst habt. Was war'n die Dinge, über die ihr gesprochen habt.

IP = Inspektionsperson, GS = Grundschule, Index = Stufe des KESS-Index

15 Die Transkription erfolgt in Anlehnung an Lamnek 2010: (.) = ganz kurze Pause, neues Ansetzen; (--) = mittlere Pause (0,5 Sek.); (---) = lange Pause (1 Sek.); SCHÜ = Betonung; [Text Sprecher/-in A. Text Sprecher/ in B.] = gleichzeitiges Sprechen; die Angabe S3? bedeutet, dass das Geschlecht des sprechenden Schülers/der Schülerin sich nicht ermitteln ließ. Alle weiteren Transkriptionsmerkmale sind selbsterklärend.

Zunächst fallen drei unterschiedliche Eröffnungen auf: IP1 wählt die Form der Bitte („würde ich Euch bitten“). Demgegenüber entscheiden sich IP2, 3 und 4 für auf „wissen wollen“ bezogene Formulierungen: „wüsste ich ganz gerne doch noch“ bzw. „möchte ich gerne wissen“. IP5 schließlich „möchte beginnen mit der Frage“ und schließt daran die Bitte um Meldung an, IP6 „würde ganz gerne fragen“. Die von IP1 und IP5 geäußerte Bitte hat zwar Aufforderungscharakter (vgl. Bührig & Redder 2013, S. 98 f.), vermittelt jedoch die Angewiesenheit auf Kooperation. Die Fragen von IP2, 3 und 4 etablieren demgegenüber tendenziell das unterrichtstypische Setting der Wissensabfrage und -kontrolle. Dies gilt ähnlich auch für die Fragen von IP5 und 6.

Anhand des Kriteriums „Inhalt“ lassen sich drei verschiedene Typen von Einstiegsfragen unterscheiden:

- I. Bitte um Vorstellung bzw. Meldeaufforderung in Verbindung mit der Frage nach positiven Eindrücken von der Schule (IP1, IP5).
- II. Frage nach Schülerrat bzw. Kinderkonferenz und Teilnahme daran (IP2, IP4) sowie Frage nach Teilnehmenden der Kinderkonferenz und Themen der Konferenz (IP6).
- III. Frage nach dem Zustandekommen der Interviewteilnahme (wie) und nach Entscheidern (wer) (IP3).

Zunächst werden die beiden Fälle vom Typ I betrachtet:

Der Einstieg von IP1 – „Und dann würd’ ich euch bitten, euch vorzustellen mit dem Namen, mit der Klasse, in die ihr geht, und, ähm, dass ihr mir sagt, eine Sache oder vielleicht zwei Sachen, die ihr an der Schule besonders toll findet. Möch’st Du anfangen, NS1w?“ – kann aus mehreren Gründen als besonders gelungen angesehen werden: Es wird eine konkrete Bitte bzw. Aufforderung – Vorstellung mit Äußerung des Namens, der Klasse und von ein oder zwei positiven Eindrücken zur Schule – formuliert, die durch den Konjunktiv „würde ich euch bitten“ eine Höflichkeitssteigerung erfährt, welche zum einen kooperationsförderlichen Respekt bezeugt, zum anderen die real vorhandene Distanz markiert. Der Einstieg über eine Vorstellungsrunde verschafft nicht nur der IP1 einen Überblick über die Gruppe, sondern berücksichtigt auch, dass die Schülerinnen und Schüler sich möglicherweise nicht kennen, weil sie aus verschiedenen Klassen kommen. Die Bitte, „dass ihr mir sagt, eine Sache oder vielleicht zwei Sachen, die ihr an der Schule besonders toll findet“, bezieht sich unmittelbar auf den Inspektionszweck; sie ist insbesondere im Hinblick auf die begrenzte Zeit des Interviews funktional, weil die schülerseitigen Eindrücke, die zur Sprache kommen sollen, auf „eine Sache oder vielleicht zwei“ begrenzt werden. Zugleich kommuniziert die IP damit Interesse an den subjektiven Erfahrungen der Schülerinnen und Schüler, d. h. das Gütekriterium

„Spezifität“ (s. Fußnote 9) wird berücksichtigt. Hörerseitig wird durch die Bitte der IP im individuellen mentalen Bereich¹⁶ das Aufrufen von positiven Erlebnissen in Gang gesetzt. Dies kann die Befragungssituation entspannen, da mit „toll finden“ angenehme Gefühlsqualitäten angesprochen werden. Die abschließende namentliche Sprechaufforderung an NS1w erspart den Prozess des „Meldens und Drannehmens“ und zeigt zudem, dass eine erste Kennenlernrunde bereits stattgefunden hat. Diese zweite Namensrunde dient somit offensichtlich der Dokumentation.

IP1-seitig liegt dieser sprachlich präzisen Äußerung ein mentaler Fokussierungsprozess zugrunde. Die vorliegende Versprachlichung kann nicht allein durch langjährige Routine, sondern nur dadurch hervorgebracht werden, dass die IP im Moment der Eröffnung des Interviews voll präsent, aufmerksam und konzentriert ist. Was zeigt sich nun in den Äußerungen der Schülerinnen und Schüler, wie antworten sie?

Tabelle 2: Schülerantworten zur Einstiegsfrage von IP1

IP1: Und dann würd' ich euch bitten, euch vorzustellen mit dem Namen, mit der Klasse, in die ihr geht, und, ähm, dass ihr mir sagt, eine Sache oder vielleicht zwei Sachen, die ihr an der Schule besonders toll findet. Möch'st Du anfangen, NS1w?
S1w: Ja. Ich bin NS1w.
IP1: Mhm.
S1w: und geh in die 3B.
IP1: in die 3B. Und was findest du toll an der Schule?
S1w: Dass sie groß ist.
IP1: Dass sie groß ist. Super. NS2w?
S2w: Äh, ich bin NS2w und ähm, ich geh inne 2D und ich finde hier das schön, dass (--) wir so schöne Lehrer haben.
IP1: Dass Ihr so schöne Lehrer habt. Okay. Dankeschön, NS2w.
S3w: Ich heiße NS3w, ich geh in die 4A und ich finde hier an der Schule schön, dass es hier so groß ist und, ähm, (--) dass alle Kinder hier nett sind, und die Lehrer auch.
IP1: Okay. (--) NS4m?
S4m: Isch bin NS4m und isch bin 'n die/ isch bin in die 4D, isch finde die Schule schön, weil, (.) hier macht es Spaß und da gibt es, äh, (.) nette Lehrerinnen, das macht richtig Spaß.
IP1: Mhm, Dankeschön.
S5w: Ich bin NS5w, geh in die Klasse 4B, mir gefällt die Schule, weil sie groß ist und, äh, jeder, äh, jede nett ist.
IP1: Mhm. Dankeschön.

16 Der Begriff des mentalen Bereichs geht auf Ehlich & Rehbein (1986) zurück. Er umfasst die Repräsentation der außersprachlichen Wirklichkeit, also das Wissen im „im Kopf“ des Hörers/der Hörerin einschließlich kognitiv-emotionaler Aktivitäten.

S6m: Ich heiß' NS6m, geh in die 3A, und ich find' (--) den Sportunterricht cool.
IP1: Okay! Danke!
S7w: Ich bin NS7w, ähm, ähm, ich geh in die 4C, ähm, und ich finde, ähm, an der Schule sehr toll, dass, ähm, wir so 'n schönen großen Schulhof haben (---)
IP1: Dankeschön.
S8m: Ich bin NS8m, ich (.) mir macht das Spaß in der Schule, dass wir da lernen.
IP1: Mhm. Okay. Und in welcher Klasse bist Du?
S8m: 2B.
IP1: In der 2B bist Du. (.) Okay. (.) Wunderbar. [Min. 02:12]

Die Schülerantworten kommen ohne Zögern und weisen nur wenige Sprechverzögerungen auf: Als „toll“ wird dreimal die Größe der Schule bzw. einmal des Schulhofs benannt, dreimal, dass die Kinder bzw. Lehrkräfte „nett“ seien – „schöne Lehrer“ meint wahrscheinlich ebenfalls die Qualität „nett“ – zwei Antworten verweisen auf „Spaß“ am Lernen, lediglich eine Antwort weicht vom Muster der generalisierenden Antwort ab und spezifiziert den Sportunterricht als „cool“. IP1 setzt Elemente der „Sprecher-Hörer-Steuerung“ (Kameyama 2004, S. 87 f.) ein, wie z. B. „mhm“, Wiederholungen (*Echoing*), positive Kommentierungen („super“) und Dank. Dieses aktive Zuhören „steuert“ insofern, als es dazu geeignet ist, auf die Auskunftsbereitschaft der Schülerinnen und Schüler positiv einzuwirken. Betrachtet man die Redeanteile, so sind von insgesamt rund 270 Wörtern ein Drittel der IP1 und zwei Drittel den Schülerinnen und Schülern zuzuordnen.¹⁷ Auch dies ist wiederum ein Hinweis auf die Funktionalität bzw. Klarheit der von IP1 gestellten Frage: Es sind keine korrigierenden oder re-fokussierenden Eingriffe erforderlich, lediglich in zwei Fällen erfragt IP1 eine in der Schülerantwort fehlende Information.

Die inhaltlich ähnliche Einstiegsfrage von IP5 führt – ungeachtet der hier fehlenden Begrenzung auf „ein oder zwei Sachen“ – ebenfalls zu relativ kurzen Antworten:

17 Es geht hier nicht um eine genaue Auszählung, sondern um ein Abschätzen der Redeanteile. Die Auszählung wurde im Word-Programm mittels Markierung vorgenommen.

Tabelle 3: Schülerantworten zur Einstiegsfrage von IP5

IP5: Dann möchte ich beginnen mit der Frage, und ich bitte euch dann, euch zu melden, was findet ihr denn an der Schule GANZ besonders TOLL. NSw1, du warst die Erste.
S1w: Das Rechnen.
IP5: Mhm. NS2m.
S2m: Ich find das Deutsch ganz gut, weil man da auch immer so (.) und Sachkunde weil ähm, bei beiden Fächern kann man ganz viel lernen auch. Das macht [dann auch Spaß.
IP5: Mhm. Okay.] Mhm.
S3w: Die Lehrer.
IP5: Mhm. NS4w.
S4w: Den Schulhof.
IP5: Mhm. NS5m.
S5m: Ähm, dass man auch andere Sachen macht, sondern auch mal in' Stadtpark geht oder, ähm, Klassenfahrt [auf Klassenfahrt geht.
IP5: Mhm, Mhm] ((Sehr leise, freundliche Anweisungen, die Reihenfolge der Antwortenden betreffend.))
S6m: Also, ich finde ganz besonders Sachunterricht gut, weil da kommen immer beson (.) also, für mich kommen da immer besonders spannende Themen, weil.*
IP5: Kannst du ein Beispiel nennen für so ein Thema, was (.) [besonders
S6m: Ähm, (..)] Feuer.
IP5: Mhm. (.) Danke.
S7m: Ich finde Religion eigentlich ganz gut.
IP5: Mhm. (...) Mhm. Danke.
S8(?): Ich find' den Sportunterricht toll.
IP5: Mhm. Okay, klasse. [Min. 01:14]

* Vermutlich Schüler mit Deutsch als Zweitsprache: Das „weil“ mit abfallendem Tonfall ist hier wohl anstatt „darium“ eingesetzt.

IP5, die ebenfalls Signale aktiven Zuhörens nutzt, erhält im Vergleich zu IP1 eher Aussagen über individuelle Vorlieben von Schülerinnen und Schüler. Es finden sich weniger generalisierende, sondern mehr spezifizierte, z. B. auf bestimmte Fächer bezogene Antworten – ein Unterschied, der sich jedoch nicht auf die Art und Weise des Fragens zurückführen lässt, sondern möglicherweise auf das vorherige Warm-up. Die Redeanteile betragen hier rund 70 (IP5) zu 90 (SuS) Wörtern. Die Sequenz ist insgesamt von kürzerer Dauer als bei IP1, da die Schülerinnen und Schüler nicht nach Name und Klasse gefragt werden. Außerdem verzichtet IP5 auf ein *Echoing* und drei Schülerinnen und Schüler beschränken sich auf Ein-Wort-Sätze.

In beiden Fällen kann die Einstiegsfrage vom Typ I auch als erweitertes Warm-up angesehen werden, insofern sie der Etablierung eines kooperativen Gesprächsklimas dient – der Evaluationszweck der Schulinspektion scheint zugunsten eines

wohlwollenden Interesses¹⁸ in den Hintergrund zu treten. Dieses wertschätzende Interesse ist in seiner Bedeutung für eine *offene Haltung der IP* im Hinblick auf den Evaluationszweck keinesfalls zu unterschätzen, denn: „Es ist nicht leicht, über die Praxis anders als negativ zu reden“ (Bourdieu 1993, S. 147) – und umso leichter kann es geschehen, dass das Fehlen einer Äußerung über eine Sache mit dem Fehlen der Sache selbst verwechselt und kritisiert wird. Die auf positive Äußerungen abzielende Einstiegsfrage beugt dem vor und folgt der klassischen Feedback-Regel, zuerst das Positive *zu hören*.

Nun zu den drei Fällen vom Typ II:

Dem Typ II sind die Frage nach Schülerrat bzw. Kinderkonferenz und Teilnahme daran (IP2, IP4) sowie die Frage nach Teilnehmenden der Kinderkonferenz und Themen der Konferenz (IP6) zugeordnet. Dieser Fragentyp führt aus verschiedenen Gründen zu Schwierigkeiten, wie die drei folgenden Analysen zeigen.

Tabelle 4: Schülerantworten zur Einstiegsfrage von IP2

IP2: Dann wüsste ich, ääh, ganz gerne doch noch (.) eine Frage hab ich noch. Es gab ja hier so etwas, äh, wie einen SCHÜlerrat. Wer (.) is' jemand von euch in dem Schülerrat?
IP2: Eins, zwei, drei. Wunderbar.
S1w: Wir drei.
IP2: Ähm, NS1w und / ich schreib das noch mal einmal auf. (.) Wer is' im Schülerrat? (--) NS2w und NS3w: Mhm.
S1w: Und NS4w war mal.
S4w: Ja, ich WAR mal.
IP2: und NS4w. Okay, vielleicht kannst du später (.)
S5w: Ich bin Vertreterin.
IP2: [und NS5w, okay. Vielleicht könnt ihr euch ja dann an etwas erinnern.
SuS: ((Gespräche im Hintergrund))]
IP2: Ich wüsste nämlich ganz gerne, als ihr euch das letzte Mal getroffen habt, um welche Themen ging es denn da? (2 Sek.) NS2w. (.) Heißt du NS2w oder NS2w-Nw? ((Doppelname))
Zwei Sw: NS2w-Nw.
IP2: NS2w-Nw.
S2w: Ähm, wir haben (.) weil die Fahrzeuge ja verbrannt sind, da haben wir darüber gesprochen.
IP2: Mhm.

18 Dies entspricht in der Grafik von Ehren und Visscher zu *Communication Styles* dem Segment „Helping/Friendly“, das u. a. mit „Showing Interest“ und „Giving Trust“ verbunden ist (vgl. Ehren & Visscher 2006, S. 55)

S2w: Mit dem Schuppen da, wo die Fahrzeuge verbrannt sind.
IP2: Mhm. Und kannst du mir mal erzählen, wie das genau abläuft, wenn ihr euch trifft, wer ist dabei und worum geht es dann.
S2w: Ääämh, da sind (.) da war'n zwei Lehrer dabei.
IP2: Mhm.
S2w-N: Uuund (5 Sek.) [Frau N1 und (.) Frau (.) und Frau N2, genau.
SuS: ((Namensnennungen im Hintergrund))]
S2w: Denn haben wir das aber besprochen.
IP2: Jaaa? Und was ist dabei rausgekommen? (3 Sek.)
S2w: ((leise:)) Weiß ich nicht mehr. [Min. 01:34]
... [aus Platzgründen gekürzt; Dauer der Sequenz insgesamt bis Min. 03:10]

Die Aufnahme beginnt mit einer Formulierung, die vermuten lässt, dass das Aufnahmegerät mitten in der noch zum Warm-up gehörenden Phase „Funktionen der Interviewpartner“ (s.o.) eingeschaltet wurde. Der anschließende Satz fokussiert auf das Thema „Schülerrat“, die dann folgende Frage fordert diejenigen Schülerinnen und Schüler, die dem Schülerrat angehören, dazu auf, sich zu erkennen zu geben. Es schließen sich Namensklärungen an und IP2 wiederholt die Frage nach den Mitgliedern des Schülerrats, weitere Namen werden von ihr notiert und zwar in Verbindung mit zwei Äußerungen, die mit „vielleicht“ beginnen: Die erste wird abgebrochen, die zweite ist die vage gehaltene Vermutung, die Schülerinnen und Schüler könnten sich erinnern: „Vielleicht“, „ja dann“ und „an etwas“ bieten den Schülerinnen und Schülern jedoch keinen Anhaltspunkt, sie führen ihre Gespräche untereinander weiter. Mit „Ich wüsste nämlich ganz gerne“ leitet IP2 die nächste Frage ein, in der sie einen zeitlichen („das letzte Mal ...“) und einen inhaltlichen („welche Themen“) Fokus setzt. Es folgt eine 2-sekündige Pause, in der die Angehörigen des Schülerrats ihre Erinnerung aktivieren. Nach dem Aufrufen der Schülerin S2 folgt zunächst mit der Namensklärung eine kurzfristige Unterbrechung der Fokussierung. Die dann zügig erfolgende Antwort der Schülerin zeigt jedoch, dass sie mental bei ihrem Thema geblieben ist: Sie berichtet von einem außergewöhnlichen Ereignis – „... weil die Fahrzeuge ja verbrannt sind“ – und präzisiert nach dem „mhm“ von IP2, wo das Ereignis stattfand („Schuppen da“). Mit der nächsten Frage zielt IP2 mit den Fragewörtern „wie“, „wer“ und „worum“ auf den generellen Ablauf der Schülerratssitzung („wie das genau abläuft, wenn ihr euch trifft“) und versucht damit, eine stärker abstrahierende Antwort hervorzulocken, was jedoch misslingt. Die Schülerin bricht den Versuch der Umorganisation ihrer Antwort ab – „Ääämh, da sind“ – und kehrt mit dem Tempuswechsel – „da war'n“ – zu ihrem Thema zurück. Dabei wird sie auf der Suche nach den Namen der Lehrkräfte von anderen Schülerinnen und Schülern unterstützt. IP2 will dann das Ergebnis wissen, was die Schülerin aber nicht mehr weiß – das vermutlich mit

Emotionen wie Aufregung verbundene Ereignis überlagert die Erinnerung an die nachfolgende Schülerratsbesprechung.¹⁹

Der eigentliche Beginn des Interviews wird von IP2 nicht deutlich kenntlich gemacht, das Ansprechen thematischer Aspekte überlappt sich mit Namensklärungen. Der mehrfache Fokuswechsel erklärt den hohen Redeanteil von IP2, der hier zwei Drittel einnimmt. Auch im nächsten Fall von IP4, die zwar die Einstiegsfrage eindeutig markiert – „Als erstes möchte ich ...“ – führt die Frage nach der Kinderkonferenz zu weiterem Klärungsbedarf und einem entsprechend höheren Redenanteil von IP4:

Tabelle 5: Schülerantworten zur Einstiegsfrage von IP4

IP4: Als erstes möchte ich gerne wissen, ihr habt (.) ich hab gehört, dass ihr ne (.) Kinderkonferenz (.) hier habt (.) an der Schule. Seid ihr da schon mal dran beteiligt gewesen?
S1m: [Also ICH nicht.
Mehrere SuS: Nein. Nein. Ich auch nicht.]
IP4: Gar niemand von euch? Und wart ihr Klassensprecher im letzten Jahr?
Mehrere SuS: Nein.
IP4: Alles andere, ja? Warst DU schon Klassensprecher im letzten Jahr? (--) Ja, NS2m? Warst DU denn dann schon mal bei 'ner Kinderkonferenz mit dem Schulleiter (.) zusammen?
S2m: Mmh.
IP4: Was habt ihr denn dann besprochen? (3 Sek.) Worum ging's da so? Was habt ihr da besprochen? Welche Themen?
S3?: ((flüsternd:)) Das mit den Kaninchen.
IP4: Ja, sag mal, wenn du das weißt.
S3?: Ähm, wegen unserm Schulzoo in [klein
IP4: Mhm]
S3?: Haben die denn das besprochen. [Min. 00:56]
... [aus Platzgründen gekürzt; Dauer der Sequenz insgesamt bis Min. 01:56]

Über die gesamte Länge der Einstiegssequenz gerechnet beträgt das Verhältnis der Redeanteile rund 160 Wörter (IP4) zu etwa 120 Wörtern (SuS). Die Einstiegsfrage enthält eine Umorganisation, IP4 unterbricht sich selbst („ihr habt (.) ich hab gehört“). Da sich die Frage nach der „Kinderkonferenz“ als unergiebig erweist, fragt IP4 nach „Klassensprecher im letzten Jahr“ – d. h. sie fordert dazu auf, sich an das vergangene Schuljahr zu erinnern, das jedoch zum Interviewzeitpunkt bereits ein paar Monate zurückliegt. Insofern ist es nicht verwunderlich, dass sich S2, der sich

19 Das Thema „Fahrzeuge verbrannt“ beansprucht weitere 30 Sekunden, das Thema „Schülerrat“ insgesamt 3:02 Minuten. Für den Zweck dieses Beitrags ist die analysierte Sequenz jedoch ausreichend.

als ehemaliger Klassensprecher meldet, nicht an die Themen der Kinderkonferenz erinnern kann. IP4 fragt S2 daraufhin, unterbrochen von einer dreisekündigen Pause, insgesamt viermal nach den Themen der Kinderkonferenz – ohne Ergebnis. Das „Abarbeiten“ des Interviewleitfadens führt in diesem Fall zum „Nachbohren“. S3 springt schließlich S2 flüsternd bei: Ihm oder ihr ist „das mit den Kaninchen“ eingefallen – und damit wiederum eine Episode, die vermutlich mit Emotionen verbunden ist und sich S3 deshalb besonders gut eingeprägt hat.

Im nächsten Fall rekurriert IP6 ebenfalls auf das vorangegangene Schuljahr. Es folgt unmittelbar auf die Einstiegsfrage eine Antwort von S1, der sich offensichtlich gleich gemeldet hatte, wie sich aus der Aufforderung „Mhm, NS1m?“ schließen lässt. S1 thematisiert „wie das jetzt ... mit der neuen Schule wird“.

Tabelle 6: Schülerantworten zur Einstiegsfrage von IP6

IP6: Dann würd ich jetzt ganz gerne als Erstes mal die Kinder der KINDERkonferenz fragen, ob ihr mir kurz erzählen könnt, womit ihr euch im LETZTEN Schuljahr befasst habt. Was war'n die Dinge, über die ihr gesprochen habt. Mhm, NS1m?
S1m: Also wir haben immer darüber gesprochen, wie das jetzt mit der neuen Klasse w (.), ähm, mit der neuen Schule wird (--) und (--)
IP6: Also, ihr habt über die SCHULreform und, und das, was das für die Schule heißt, gesprochen, ah-ha?
S1m: Mmh. Und wir haben dann eben noch (.) ob sie vielleicht irgendwie eine (.) zusammen eine Sch (.) Schülerzeitung haben oder so.
IP6: Ja.
S1m: Und darüber sprechen wir jetzt ganz viel, über die (.) neue Schule.
IP6: Mhm. Meinst du zusammen mit der Schülerzeitung, zusammen mit der anderen Schule, mit der ihr vielleicht zusammengehen sollt? [Oder?
S1m: Mmh.]
IP6: Okay. NS2w.
S2w: Wir, ähm, haben im [Monat vor dem Interview] darüber gesprochen ob wir vielleicht noch was an unserm Schulhof verbessern können [und so.
IP6: Mhm.] [Min. 00:59]
... [aus Platzgründen gekürzt; Dauer der Sequenz insgesamt bis Min. 01:07]

Die Redeanteile der Interviewpartner (IP6: SuS) sind in diesem Ausschnitt etwa gleich groß. Durch die Verwendung von „jetzt“ und „wird“ zeigt S1 an, dass das Thema „mit der neuen Schule“ die Kinderkonferenz bis in die Gegenwart (und Zukunft) hinein beschäftigt – daher bereitet in diesem Fall die Fokussierung auf die Kinderkonferenz „im LETZTEN Schuljahr“ keinerlei Mühe. IP6 ordnet diese erste Antwort in den Kontext „Schulreform“ ein und fragt in verständnissichernder Absicht nach (... „ah-ha?“). S1 bestätigt, dass IP6 seine Ausführung richtig ver-

standen hat mit „mmh“ und ergänzt, dass es Überlegungen für eine gemeinsame Schülerzeitung gibt. Auch hierauf erfolgt eine verständnissichernde Nachfrage von IP6 („Meinst du ...?“), die zugleich ihr ernsthaftes Interesse und Verstehen-Wollen transportiert. Auch der abschließende Satz von S1 bezieht sich auf die Gegenwart („darüber sprechen wir jetzt ganz viel“). S2 knüpft in der jüngsten Vergangenheit an und benennt mit „Schulhof verbessern“ ein weiteres aktuelles Thema der Kinderkonferenz.

Bei den Einstiegsfragen vom Typ II soll mit der Frage nach der „Kinderkonferenz“ Quantität und Qualität von Schülerbeteiligung evaluiert werden (vgl. BBS 2006, S. 30). Insgesamt gelingt es jedoch nicht, Antworten auf Fragen zu erhalten, die auf einen länger zurückliegenden Zeitraum abzielen. In zwei Fällen scheint zudem eine Differenz zwischen dem Evaluationsinteresse und kindlichem Relevanzsystem auf. Letzteres ist nach Episoden emotionaler Bedeutsamkeit strukturiert und nicht nach den abstrakte(re)n Gesichtspunkten von Beteiligung und Gremienarbeit.

Der Fall vom Typ III zeigt eine weitere Variante des Intervieweinstiegs:

Tabelle 7: Schülerantworten zur Einstiegsfrage von IP3

IP3: So. Ähm (--), als erstes möchte ich gerne von euch wissen, äh, wie kommt es dazu, dass grade jetzt IHR (.) jetzt (.) hier (.) heute zu diesem Interview gekommen seid. (---) Wer hat das bestimmt und (--) wer hat das ausgesucht und wie habt ihr überhaupt davon erfahren. (2 Sek.) NS1m.
S1m: Also, ich/ ich stand auffe Tafel (--), dass ich so, also, die ham gesch (.) geschrieben, wohin einer geht und ich stand auffe Tafel mit NSw noch, mh, dass wir zum (.) hierher komm' sollten.
IP3: Mhm. Okay, ja. NS2w, wie war das.
S2w: Wir standen auch heute Morgen auf der Tafel ((unverständlich)) weil das ist ja ein Wort auf Englisch [und
IP3: Aha]
S2w: dann ham wir noch mal alle Sachen, die auf der Tafel standen, besprochen und dann hat N. gesagt, dass ((unverständlich)) [Min. 01:03]
... [aus Platzgründen gekürzt; Dauer der Sequenz insgesamt: Min. 02:00]

Hier wird zunächst danach gefragt, wie es dazu kam, dass gerade diese Schülerinnen und Schüler nun beim Interview anwesend sind. Es folgen zwei weitere Fragen: Nach den Personen, die über die Interviewteilnahme von Schülerinnen und Schülern entscheiden, und nach den schulischen Informationswegen. Damit handelt es sich um ein Bündel von Kontrollfragen, die sich nicht auf die Schule als solche, sondern darauf beziehen, ob die Schule die Vorgaben der Schulinspektion – in diesem Fall die Vorabinformation der Schülerinnen und Schüler bzw. Eltern – eingehalten hat. Dabei lässt die erste Frage – „wie kommt es dazu“ – viel

Raum für mögliche Antworten. Vielleicht hat IP3 dies bemerkt, da sie nach einer kurzen Pause spezifizierte Fragen folgen lässt, die zugleich bestimmte Annahmen explizieren, nämlich darüber, dass jemand „das bestimmt“ bzw. „das ausgesucht“ hat. Die Antworten zeigen, dass die Schülerinnen und Schüler sich an der letzten Frage – „wie habt ihr überhaupt davon erfahren“ – orientieren. Die Redeanteile lassen sich hier aufgrund längerer unverständlicher Passagen nicht ermitteln. Es lässt sich schlussfolgern, dass eine Fragenreihung, die verschiedene Foki setzt, Schülerinnen und Schüler im Grundschulalter tendenziell überfordert.

3.2 Vergleich der Abschlussfragen

Als „Abschlussfrage“ wurde die Frage definiert, die den letzten inhaltlich zusammenhängenden Abschnitt des Interviews einleitet. In drei Fällen wurde die jeweilige Frage von den Inspektorinnen und Inspektoren als „letzte Frage“ oder „Abschlussfrage“ explizit benannt (IP3, IP5, IP6).

Tabelle 8: Abschlussfragen

IP1 GS A Index 1	Dann möchte ich euch bitten, dass ihr euch was vorstellt, und zwar, dass ihr für einen Tag Schulleiter (.) sein dürft. Und ich möchte von euch wissen, was wäre das ERste, das ihr entscheiden, verändern, verbessern würdet. Möch'st du anfangen, NS1m?
IP2 GS B Index 2	So. Wie seid ihr denn mit der/ mit dem Online-Fragebogen zurechtgekommen, war der schwierig oder ging das so? Ähm, (1,5 Sek.) NS1w?
IP3 GS C Index 3	Gut. Dann stell' ich jetzt meine letzte Frage und dann (.) könnt ihr auch gleich wieder los. Wenn ihr jemand anders jetzt erzählen solltet, de (.) der Hamburg nicht kennt und hier [Stadtteil] nicht kennt, ähm, was ist das TOLLE an dieser Schule. (1 Sek.) Was würdet ihr dem erzählen. (1,5 Sek.) Warum soll jemand, der hierher zieht, vielleicht sein Kind unbedingt hier in die Schule schicken.
IP4 GS D Index 4	DAS kommt jetzt, die nächste Frage. ALSO: Wenn jetzt ich euch frage, was erzählt ihr denn diesem fremden Menschen, was es hier noch zu verbESSERN gibt, was vielleicht NICH so gut an dieser Schule ist?
IP5 GS E Index 5	Okay. Und jetzt kommt die ALLERletzte Frage und da würde ich EUCH gerne fragen, JEden von euch, wenn ihr EINE Sache (.) ändern könntet an der Schule, was würdet ihr dann sofort ändern?

IP6 GS F Index 6	Okay. Dann hab ich jetzt meine Abschlussfrage und danach dürft ihr dann auch gern fragen. Ähm, und zwar stellt euch mal vor, es kommt ein neues Kind hier in die Stadt (.) und das kennt Hamburg nich' und muss aber jetzt irgendwie zur Schule gehen und möchte gerne wissen, ob diese Schule gut is oder nich. Und ihr sollt dem Kind jetzt GANZ ehrlich sagen, was is an dieser Schule GUT (.) und genauso ehrlich sollt ihr dem Kind aber auch sagen, was ist an der Schule vielleicht nicht ganz so gut. Und ich hab jetzt die Bitte, dass wir jetzt gleich noch mal so 'ne Reihe machen und jedes Kind sagt EINE Sache, die es richtig toll findet, aber wirklich nur eine, und EINE Sache, die es, wenn euch was einfällt, NICH so gut findet. Ja?
-------------------------------	---

Auch hier lassen sich entlang inhaltlicher Kriterien wiederum drei Fragetypen unterscheiden:

- I. Frage danach, was die Schülerinnen und Schüler verbessern bzw. ändern würden (IP1, IP5).
- II. Frage danach, was an der Schule toll oder gut bzw. nicht toll oder gut ist (IP3, IP4, IP6).
- III. Frage nach dem Online-Fragebogen (IP2).

Zunächst zu den Fällen vom Typ I:

IP1 fordert die Schülerinnen und Schüler auf, sich vorzustellen, für einen Tag Schulleiterin oder Schulleiter zu sein, also mit der Macht ausgestattet zu sein, „entscheiden, verändern, verbessern“ zu können. Die Schülerinnen und Schüler sind damit gefordert, ihre Perspektive mental umzuorganisieren – sie sollen die Schülersicht verlassen und sich in die Rolle der Schulleitung hineinversetzen. Ihre Antworten orientieren sich auch hier (wie oben bei der Fragenreihung von IP3) am zuletzt genannten Begriff: „verbessern“. Wie bereits bei der Einstiegsfrage schränkt IP1 auch hier die Anzahl der Antworten ein – es soll nur „das Erste“ genannt werden.

Tabelle 9: Schülerantworten zur Abschlussfrage von IP1

IP1: Gut. Dann möchte ich euch bitten, dass ihr euch was vorstellt, und zwar, dass ihr für einen Tag Schulleiter (.) sein dürft. Und ich möchte von euch wissen, was wäre das ERste, das ihr entscheiden, verändern, verbessern würdet. Möch'st du anfangen, NS1m? Wir gehen einmal der Reihe nach rum? Gut, dann ist so – möch'st Du anfangen? [Min. 39:48]
S1m: (1,5 Sek.) Ich hab noch nichts.
IP1: [Du hast noch nichts.
Mehrere SuS: Ich auch nicht.]
IP1: Wer hat denn schon was. Du hast was, dann fang' du an.
S2w: Ähm,
IP1: GERNE auch in Stichworten, also, nicht ganz so lang, aber erzähl. ((Klatschgeräusch))

S2w: das, das, ähm, die Kinder, wir ha'm ja auch so 'ne Smileys und so, wenn man besser ist und so, und ich würde verbessern, dass, ähm, die Kinder und so, dass man das da aufschreibt, an so 'ne Tafel, wer nicht, ähm, leise ist und so, dass machen wir auch, aber was ich verbessern würde, dass ähm, die Kinder, ähm, alle nich so LAUT sind und so und, ähm, dass wir mehr Kopfhörer kriegen als nur drei.
IP1: Okay, gut, dann hab ich das verstanden. Ich würd gerne das einmal so rundrum laufen lassen, damit euch die Arme nicht so schwer werden. Möch'st du weiter machen, NS3w?
S3w: Ähm, (--)[als Schulleiterin würde ich alle
IP1: Kannst auch den Arm runter nehmen]
S3w: Wenn ich eine Schulleiterin wäre, dann hätte ich dafür gesorgt, dass die Kinder hier glücklich sind und, ähm, dass wir viele Geräte haben,
IP1: Mhm.
S3w: damit die Kinder Spaß haben und, ähm (---)
IP1: Hast du denn, wenn ich mal kurz dazwischen fragen darf, hast du denn den Eindruck, dass hier viele Kinder nicht so glücklich sind oder wolltest du lieber, dass sie noch glücklicher sind?
S3w: Nein, also wenn die vielleicht mal traurig sind, oder so, damit man die glücklich machen kann, ähm, mit Spielgeräten oder so, oder von der Pausenhalle, ähm, die Spielsachen, damit man mit denen mal spielen kann.
IP1: Gut. Dankeschön. NS5m?
S4m: Isch hab keine ((unverständlich)).
IP1: Mhm. Dann vielleicht noch mal am Ende, wir gehen dann einmal rum, ne?
S5w: Dass die Kinder sich besser an die Regeln halten, sodass man nur sagen, ja, ich hab nicht, und dann machen sie das nicht.
IP1: Mhm. NS6m.
S6m: Äh, länger Schule, dafür drei Pausen.
IP1: Länger Schule, dafür drei Pausen. Okay (---).
S7w: Ich würd das so machen, dass sich alle Kinder hier auf der Schule richtig wohlfühlen und sie glücklich sind ((Stimme aus Lautsprecher)) (2,5 Sek.)
IP1: Mhm. (1,5 Sek) und wie würdest du das versuchen zu erreichen? Haste 'ne Idee? (1,5 Sek.) Gut. Aber (--) das wär' dein Ziel auf jeden Fall. S1m.
S1m: Mir ist es gut, dass wir Pausenfußball haben und auch die Fußballtore.
IP1: Mhm, die sind jetzt neu, ne?
S1m: Ja, das kleine.
IP1: Das kleine Feld, mhm.
S1m: Das große nicht.
IP1: Mhm. Gut.
S8w: Ich würde (--) Zensuren bei dem Zeugnis machen.
IP1: Die würdest du dir wünschen. Dir war noch was eingefallen.

S3w: Ähm, dass äh, die Pausenhalle, ähm, die ähm, auch die zweite Pause auf hat und ähm, dass, dass das ähm, kleine Feld größer ist. ((<i>Pausenlärm</i>))
IP1: Okay. NS4m, ist dir noch was eingefallen? (--) Gut. Aber dir ist noch was eingefallen.
S6m: Ähm, und ich würde [auch noch ein zweites, kleines Fußballfeld machen damit man mehrmals die Tore hat und (--) ähm,
((<i>Geräusche von draußen, Lehrerin-Stimme: Auf die Plätze, fertig, los!</i>))]
IP1: Okay. Gut.
S6m: weil ich, unsere Klasse hat sich nur auf dem kleinen Feld mit der 3D gestritten,
IP1: Mhm.
S6m: und denn ha (.) denn mussten wir das so machen, dass ähm, eine Woche, eine Woche die, [die haben und
IP1: Und eine Woche ihr.]
S6m: die andere Woche wir.
IP1: Ja, mhm. Gut. Dir ist noch was eingefallen?
S4m: Äh, isch hätte es so, isch hätte noch mehr Schaukels gemacht bei der ((<i>unverständlich</i>)), wenn die immer bei Schaukels sind,
IP1: Mhm.
S4m: und das zählen sie immer so schnell, die streiten sich manchmal sehr, und ich bin jetzt und der andere sagt dann, nee, ich bin doch dran.
IP1: Mhm. Okay. Gut. Dann bedank' ich mich ganz herzlich und mach' mal hier aus. [Min. 43:15]

Der zuerst von IP1 angesprochene Schüler – wie auch weitere Schülerinnen und Schüler – können noch keine Antwort geben. Dies könnte damit zusammenhängen, dass es ihnen schwer fällt, sich in die Rolle des Schulleiters zu versetzen, oder auch damit, dass der Prozess des Fokussierens auf mögliche Verbesserungen zu diesem Zeitpunkt noch nicht abgeschlossen ist. IP1 muss vom Vorhaben, die Schülerinnen und Schüler reihum antworten zu lassen, abweichen und formuliert neu: „Wer hat denn schon was.“ S2 setzt mit ihrer Antwort an, IP1 unterbricht jedoch mit einem nachgeschobenen, diskurssteuernden Eingriff: „GERNE auch in Stichworten.“ Mit dieser Formulierung scheint IP1 die Schülerinnen und Schüler von der Mühe, in ganzen Sätzen zu antworten, entlasten zu wollen. Die nachfolgende Präzisierung „also, nicht ganz so lang“ macht jedoch deutlich, dass es IP1 vor allem um kurze Antworten geht; die abschließende Aufforderung „aber erzähl“ nimmt die Begrenzung allerdings wieder zurück. S2 nimmt sich entsprechend Zeit, um ihr Anliegen („dass die Kinder nicht so LAUT sind“ und „dass wir mehr Kopfhörer haben“) zu erläutern. Mehrfach wechselt die Schülerin dabei zwischen „Kinder“ und „wir“. Sie sagt zweimal „ich würde verbessern“ bzw. „was ich verbessern würde“ und nutzt zwischendurch das unpersönliche man („wenn man besser ist“, „dass man das da aufschreibt“). Daran zeigt sich, dass der Wechsel in die Schulleiterrolle nicht

vollständig gelingt, sondern die Sprecherin zwischen verschiedenen Perspektiven balanciert, um ihr Anliegen zu begründen. Die von IP1 gewählte Form der Abschlussfrage erscheint somit unnötig verkomplizierend – auch wenn, wie nachfolgend gezeigt wird, einige Schülerinnen und Schüler den Wechsel der Sichtweise erfolgreich managen.

IP1 signalisiert S2 mit der Phrase „dann habe ich das verstanden“ sowohl, dass sie sich durch ihren Beitrag über den Sachverhalt ausreichend informiert fühlt, als auch, dass sie nun einen Sprecherwechsel anstrebt. Die nächste Schülerin (S3) organisiert den vorgestellten Rollenwechsel (bis auf einen Tempusfehler) sprachlich korrekt und geschlechtsangepasst über die Formulierung „wenn ich eine Schulleiterin wäre, dann hätte ich dafür gesorgt, dass die Kinder hier glücklich sind und ähm, dass wir viele Geräte haben, damit die Kinder Spaß haben“. Geräte sind nach Meinung von S3 offensichtlich nicht oder in nicht ausreichender Zahl vorhanden und dieses Faktum ist auf eine nicht näher bestimmte Art und Weise mit dem Glücklich-Sein der Kinder verbunden. Klar wird aber, dass „viele Geräte haben“ für „Spaß haben“ sorgt. IP1 versucht durch ihre Zwischenfrage Aufklärung darüber zu erlangen, ob S3 die Mitschülerinnen und Mitschüler für „nicht so glücklich“ hält oder ob sie möchte, „dass sie noch glücklicher sind“ – eine Fragestellung, für deren Beantwortung S3 eine komplexe Überlegung anstellen müsste und die am Kern ihres Verbesserungsvorschlags vorbei zielt, wie ihre Antwort zeigt: Sie verneint die Frage schlicht und erläutert („also“), dass es um ein *mögliches* Unglücklich-Sein – „wenn die vielleicht mal traurig sind“ – geht. Dann nämlich sollte der Zugang zu den vorhandenen Spielgeräten doch gegeben sein, „damit man die glücklich machen kann“ und „damit man mit denen [den Spielgeräten von der Pausenhalle] mal spielen kann.“ Der Verweis auf Gefühle des Unglücklich-Seins hat also den Zweck, endlich oder häufiger den ersehnten Zugang zu den Spielsachen zu erlangen.

Auch S7 nennt das Wohlfühlen und Glücklich-Sein aller Kinder an der Schule als Ziel, kann oder will jedoch auf Nachfrage von IP1 keine konkreten Ideen dafür angeben. S5 möchte, „dass die Kinder sich besser an die Regeln halten“. Weitere Wünsche der Schülerinnen und Schüler beziehen sich auf eine andere Rhythmisierung des Schultags („länger Schule, dafür drei Pausen“) und die Einführung von Zeugnisnoten. Viele wünschen sich Verbesserungen für die Pausen: Zugang zur Pausenhalle, Vergrößerung des kleineren Fußballfeldes bzw. Einrichten eines zweiten kleinen Feldes, mehr Tore, mehr Schaukeln. Streit wird auf Mangel (an Fußballfeldern und Schaukeln) zurückgeführt und soll durch ein „Mehr“ beseitigt werden. Lediglich S1 hat, auch als er zum zweiten Mal aufgerufen wird, keinen Verbesserungsvorschlag, sondern äußert sich positiv („mir ist es gut“) über Pausenfußball und die Fußballtore.

Weitere Interaktionen von IP1 beziehen sich insgesamt dreimal darauf, dass die Schülerinnen und Schüler reihum antworten sollen; einmal erfolgt Hinweis, dass

der Arm herunter genommen werden kann. Neben den schon bekannten Elementen der Sprecher-Hörer-Steuerung („mhm“, „gut“, *Echoing* usw.) nutzt IP1 neben der oben erwähnten eine weitere Nachfrage, um sicher zu gehen, dass sie die Sachlage richtig erfasst hat („die sind jetzt neu, ne?“). Im Anschluss an die Antwort von S8 („Zensuren bei dem Zeugnis machen“) ordnet IP1 die Antwort mit „die würdest du dir wünschen“ ein und signalisiert damit der Schülerin, dass sie ihre Äußerung wahrgenommen hat. Weitere Meldungen von Schülerinnen und Schülern werden mit „dir ist noch was eingefallen“ aufgegriffen – IP1 verzichtet somit darauf, nochmals um die Einnahme der Schulleiterperspektive zu bitten. Insgesamt hat IP1 – durch steuernde Eingriffe sowie Zwischenfragen bedingt – einen relativ hohen Redeanteil (262 Wörter, das entspricht rund 43 Prozent). Im dreieinhalbminütigen Verlauf fällt (wie schon in der Einstiegssequenz von IP1) der Gebrauch höflicher Formulierungen auf: „möchte ich euch bitten“, „wenn ich mal kurz dazwischen fragen darf“ und „Dankeschön“. Das Interview wird mit „Dann bedank’ ich mich ganz herzlich“ beendet.²⁰

Bei IP5 nimmt die Sequenz rund zwei Minuten ein, ihr Redeanteil liegt bei etwa einem Drittel, ihr Redestil ist insgesamt knapper als der von IP1. Der Fokus der Abschlussfrage liegt darauf, was die Schülerinnen und Schüler gerne verändern würden – ohne den Umweg über die vorgestellte Einnahme der Schulleiterrolle.

Tabelle 10: Schülerantworten zur Abschlussfrage von IP5

IP5: Und jetzt kommt die ALLERletzte Frage und da würde ich EUCH gerne fragen, JEden von euch, wenn ihr EINE Sache (.) ändern könntet an der Schule, was würdet ihr dann sofort ändern? NS1m, fangen wir bei dir mal an. EINE Sache (.) dürftest du ändern. [Min. 56:51]
S1m: Hm. (--) Also, dass jede auf dem Fußballfeld, dass jedes Tor groß ist.
IP5: Wir gehen einmal der Reihe nach rum? Gut, dann ist so (.) möch’st Du anfangen? (2 Sek.) Mhm.
S1m: Wie das, wie das auch auf dem [Basketball-Ballfeld.
IP5: Mhm.] Okay. NS2w, was würdest du ändern?
S2w: Ähm, vielleicht ’n andern Klingelton.
IP5: [((lacht)). Ja. Vielen Dank.
IP5_2: ((lacht))]
IP5: NS3w, was würdest du ändern.
S3w: Hm. (--) Dass die Jungs sich nicht mehr prügeln.
IP5: Mhm. NS4w.
S4w: Ähm, dass keiner (.) also dass, dass man auch auf Sch (.) auch auf Stopp also reagiert.

20 Die Wortwahl „herzlich“ scheint fast ein Standard in der Kommunikationskultur der Hamburger Schulinspektion zu sein – alle IP bis auf IP5 verwenden „herzlich“ in den Routineformeln zur Kommunikationsbeendigung.

IP5: Mhm, jaha, NS5w.
S5w: Ähm (4 Sek) Dass jedes Kind, ähm, also gleichberechtigt ist, weil manche Kinder, ähm, die, ähm, sagen einfach: ‚Nö, du, ähm, darfst das jetzt nich‘. Oder, (.) dass halt jedes Kind genau gleichberechtigt ist zum Beispiel zu schaukeln und an die Stange, an die Kletterstange zu gehen.
IP5: Mhm. Danke dir. NS6m.
S6m: Also, ich finde, es sollte noch ’n einzelnes Basketball(.)feld geben, und nich so mit, ähm, Handball-Basketballfeld, sondern so ein EINzelnes,
IP5: Mhm.
S6m: wo man nur Basketball spielen kann.
IP5_2: Nur so ’n Tor.
IP5: Mhm, mhm, mhm.
IP_2: ’n Korb. ((flüsternd:)) Entschuldigung.
IP5: ’n Korb, genau, [’n einzelnes.
S6m: ’n Korb]
IP5: Ja, NS7m, [willst du ((unverständlich))
S7m: Also, ich finde ja,] dass man, weil die meisten ja eh Fußball spielen, dass man das wieder zurück baut, dass die Gummipätze halt weg sind und da wieder der große Fußballplatz, weil da hatte man viel mehr Platzraum zum Spielen und nicht einfach diese kleinen Kinder, weil dann hat man da Liegen und da hat man ja doch viel mehr Platzraum und das Feld ist größer.
IP5: Mhm. NS8w.
S8w: Dass man (.) dass man den anderen nich schlimme (.) nich schlimme Wörter zu den anderen sagt und dann wird der vielleicht traurig oder so.
IP5: Mhm. Vielen Dank. Ich danke (.) euch allen noch mal, die ihr heute hier wart, für die viele Zeit und von daher, hier ist noch ’n bisschen was, also knabbert gerne los, bevor ihr nach Hause geht. Vielen Dank. [Min. 58:54]

Die Antworten der Schülerinnen und Schüler fallen nach Geschlecht unterschiedlich aus: Ausschließlich die *Schüler (m)* wünschen Veränderungen hinsichtlich der Bewegungsflächen im Freien („jedes Tor groß“, „noch ’n einzelnes Basketballfeld“ usw.), während sich die *Schülerinnen* (bis auf S2) sich auf die (Nicht-)Einhaltung von Regeln beziehen („dass die Jungs sich nicht mehr prügeln“, „auch auf Stopp also reagiert“ usw.).²¹ Eine Schülerin (S5) nutzt für ihre Argumenta-

21 Möglicherweise ein Hinweis auf eine Schul- und Unterrichtskultur an dieser Schule, die geschlechterstereotypes Verhalten fördert, z. B. durch die Dramatisierung von Geschlecht (vgl. Faulstich-Wieland 2005). Die Berücksichtigung der „Gender-Dimension“ ist jedoch nicht explizit im Qualitätsrahmen von 2006 (eine der Grundlagen für die Hamburger Schulinspektion) verankert, es erfolgt dort lediglich ein Verweis auf den Erziehungsauftrag gemäß Schulgesetz (BBS 2006, S. 24).

tion das abstrakte Prinzip der Gleichberechtigung: „das halt jedes Kind genau gleichberechtigt ist.“ Der Wunsch von S2 nach einem anderen Klingelton löst beim Inspektionsteam Lachen aus – der Grund der Erheiterung ist jedoch aufgrund fehlender Kontextinformationen nicht eindeutig zu klären: Ist der Wunsch von S2 ungewöhnlich oder der Klingelton unangenehm? Durch das gemeinsame Lachen wird in dieser Sequenz die zweite Inspektionsperson (IP5_2) erstmals hörbar. Beim zweiten Mal ergreift sie das Wort, als es um den Wunsch von S6 nach einem Basketballkorb geht – sie verwendet fälschlicherweise den Begriff „Tor“ für das Spielgerät und korrigiert sich kurz darauf, gefolgt von einer geflüsterten Entschuldigung. Auch hier fehlen Kontextinformationen für die Deutung der Entschuldigung: Sie könnte sich sowohl auf die Verwendung des falschen Begriffs, als auch auf die Intervention in die Interviewführung von IP5 beziehen. Die kurze Interaktion zeigt, dass solche Interaktionen tendenziell heikel sind, da sie den Interviewfluss behindern (können). Der abschließende Dank von IP5 fällt ausführlich aus und bezieht sich auf „die viele Zeit“, die – „von daher“ – mit etwas Essbarem belohnt werden soll.²² Das „hier“ in der Formulierung „hier ist noch ’n bisschen was“ folgt unmittelbar auf das „von daher“, sodass geschlossen werden kann, dass die Belohnung erst nach Abschluss des Interviews angeboten wurde.

Die Abschlussfrage vom Typ II – Frage danach, was an der Schule gut oder toll/nicht gut oder toll ist – liegt in drei Varianten vor: Die Schülerinnen und Schüler sollen sich jeweils in die Situation hinein versetzen, „jemand anders“ (IP3), „diesem fremden Menschen“ (IP4) bzw. „einem neuen Kind“ (IP6) von der Schule zu erzählen. Damit ist Aufforderung zu einem Perspektivenwechsel verbunden, also in der Vorstellung von sich selbst abzusehen – zu abstrahieren –, um dann Wissensbestände so zu strukturieren und zu versprachlichen, dass sie für eine schulfremde Person verständlich und nachvollziehbar werden. Auch dies stellt (wie der Rollenwechsel in die Schulleiterposition) eine unnötige Verkomplizierung dar, zumal die fremde Person – in Gestalt der IP – vor Ort und somit der Umweg über die Vorstellung überflüssig ist.

Die erste Antwort auf die Frage von IP3 zeigt, dass von S1 auf den Kern der Frage – „was ist das TOLLE an dieser Schule“ – Bezug genommen wird.

22 Dies ist das einzige Mal in den sechs Interviews, dass eine Belohnung Erwähnung findet.

Tabelle 11: Schülerantworten zur Abschlussfrage von IP3

IP3: Gut. Dann stell' ich jetzt meine letzte Frage und dann (.) ihr auch gleich wieder los. Wenn ihr jemand anders jetzt erzählen solltet, der Hamburg nicht kennt und hier [Stadtteil] nicht kennt, ähm, was ist das TOLLE an dieser Schule? (1 Sek.) Was würdet ihr dem erzählen? (1,5 Sek.) Warum soll jemand, der hierher zieht, vielleicht sein Kind unbedingt hier in die Schule schicken? (1 Sek.) [Min. 34:36]
S1w: Also, die Schule ist groß, man kann hier eigentlich ganz viel erleben (1 Sek.) und einfach, mh, dann is es so, man mag einfach die Schule, weil hier fast immer, äh, etwas aufregend, also au (.) etwas Aufregendes [ist.
IP3: Okay.] Mhm. Gut. Was würdet ihr andern sagen. (2 Sek.) Warum gehst Du gerne in die Schule hier, NS2w, oder gehst du überhaupt gern hier in die Schule?
S2w: Ja.
IP3: Warum?
S2w: Weil ((unverständlich))
IP3: Weil dein Bruder im Kindergarten ist, ja, das is 'n gutes Argument.
S2w: Weil wir uns dann immer sehen und zusammen ((unverständlich))
IP3: Ja, okay. NS3w.
S3w: ((unverständlich))
IP3: Kannst du 'n bisschen lauter sprechen?
S3w: Ich finde die Schule toll, weil hier (.) ich finde die Schule toll, weil hier mehr Räume sind als in der andern Schule. ((unverständlich))
IP3: Ja, okay, mhm. (1 Sek.) Ja?
S4?: Ich finde des toll, weil (.) die Schule ist viel größer als die alte Schule
IP3: Aha.
S4?: Und die Klassenräume sind bisschen breiter als früher und [anstatt
IP3: Mhm. Okay.] Wenn ihr jetzt aber mal die RÄUME weglasst, es sind ja die gleichen Lehrerinnen und Lehrer, ne, die ihr habt, an der alten Schule und an der neuen. Was findet ihr ÜBERhaupt toll hier.
S5?: Es würde besser sein, dass wir die Lehrer auch mal tauschen.
IP3: Kannst Du das erklären, was du damit meinst?
S5?: Also, nicht, dass es immer die gleichen Lehrer sind, sonst bringt es manchmal nicht so Spaß, weil sonst machen die, die ham ja (.) manchmal haben die ja auch andere Sachen so, die wir machen sind so auch, mh, tz, so, also ((unverständlich)), wenn die nicht so gut schreiben können, sind es (.) ist es auch schwer für die.
IP3: Aha? Okay. Und NS6w?
S6w: Also, ich finde diese Schule ganz toll, aber ich finde irgendwie doch, dass hier überall Baustellen sind [und auch die großen ((unverständlich))
IP3: Okay, mhm, das versteh' ich gut.] Mhm.
S7?: Und hier ist auch die, ähm, Luft schlecht. [Min. 36:50]
... [aus Platzgründen gekürzt bis Min. 37:55] ;

IP3: Mhm. Gut. So! Ich denke, ihr könnt bestimmt ganz viel erzählen, aber (.) ich weiß, dass einige ganz schnell jetzt weg müssen von euch
--

S?: Ich muss weg.

I1: und deswegen bedank ich mich recht herzlich bei euch für dieses Interview, für die vielen Informationen. [Min. 38:09]

S1 nennt als positive Aspekte die Räume der Schule und den hohen Erlebnisgrad – Hintergrund ist, dass die Schule erst vor Kurzem neue und größere Räume bezogen hat. Im Anschluss daran formuliert IP3 ihre Frage neu und ändert deren Fokus: Sie fragt S2, warum bzw. ob sie gerne zur Schule geht. S2 bejaht dies und beantwortet die Warum-Frage von IP3 mit der räumlichen Nähe zu ihrem Bruder. Die nächsten Antworten von S3 und S4 beziehen sich wieder auf das Schulgebäude. IP3 reicht dies offenbar nicht, sie möchte den Fokus erweitern auf das, was die Schülerinnen und Schüler „ÜBERhaupt toll hier“ finden, sie sollen „mal die RÄUME“ weglassen. Der Einschub „es sind ja die gleichen Lehrerinnen und Lehrer, ne, die ihr habt, an der alten Schule und an der neuen“ ist für die Öffnung des Blicks auf weitere positive Aspekte jedoch nicht hilfreich. Die Antwort von S5 macht deutlich, warum: Er oder sie folgt der Anforderung, von den Räume abzusehen und knüpft direkt an „die gleichen Lehrerinnen und Lehrer“ an mit dem Vorschlag, „dass wir die Lehrer auch mal tauschen.“ Dies erscheint logisch: Wenn man die Räume tauschen kann und dies offensichtlich zu positiven Veränderungen geführt hat, warum dann nicht auch die Lehrer? Für IP3 ist die Antwort nicht ohne Weiteres verständlich, sie bittet S5 um eine Erklärung. Diese erfolgt dahingehend, dass es mit immer den gleichen Lehrern „manchmal nicht so Spaß“ bringt. Dabei belässt es S5 jedoch nicht, sondern sie oder er versucht eine Begründung dafür zu formulieren, die jedoch diffus bleibt, weil die Bezüge unklar bleiben: Für wen ist es schwer? Für die Lehrkräfte oder diejenigen, die nicht so gut schreiben können? Oder sind es sogar die Lehrkräfte, die nicht gut schreiben können? Das fragende „Aha?“ von IP3 zeigt ihr Nichtverstehen an, aber sie belässt es dabei: „Okay.“ Die Antwort von S6 zeigt wiederum Begeisterung für die Schule – „ganz toll“ – und richtet dann den Blick auf negative Aspekte der Umgebung der Schule: „überall Baustellen“ – eine Fokuserweiterung, die von IP3 aber so nicht intendiert war. Es folgen weitere Antworten, in denen die Schülerinnen und Schüler Mängel der schulischen Umgebung benennen. IP3 äußert Verständnis und verzichtet darauf, nochmals ihre Frage einzubringen. Mit „So! Ich denke, ihr könnt ..., aber ich weiß, dass einige ...“ leitet IP3 den Abschluss des Interviews ein. Sie nutzt dafür eine höfliche Standardformel, welche zunächst eine Vermutung über ein schülerseitiges Vermögen – „ihr könnt bestimmt ganz viel erzählen“ – aufstellt; die Umsetzung dieses Vermögens in die Tat wird, eingeleitet mit „aber“, durch die dringliche Präsentation von Faktenwissen unterbunden („ich weiß, dass einige ganz schnell jetzt weg müssen von euch“). Die Eile wird nachfolgend

durch eine Schüleräußerung („ich muss weg“) bestätigt. „Deswegen“ bedankt sich IP3 abschließend und wählt dafür die Formulierung „recht herzlich“.

Die letzte Frage von IP4 setzt den Fokus zunächst auf eventuelle Verbesserungen, spitzt die Erzählaufforderung dann jedoch auf negative Aspekte („NICH so gut“) der Schule zu. Daraufhin schildert eine Schülerin eine Situation, die erkennen lässt, dass die ihr als Klassensprecherin vom Lehrer zugewiesenen Aufgaben eine Überforderung darstellen:

Tabelle 12: Schülerantworten zur Abschlussfrage von IP4

IP4: DAS kommt jetzt,* die nächste Frage. ALso: Wenn jetzt ich euch frage, was erzählt ihr denn diesem fremden Menschen, was es hier noch zu verBESSern gibt, was vielleicht NICH so gut an dieser [Schule ist. NS1w. [36:54]
S1w: Ja, das,] also
IP4: NS1w.
S1w: Viele PRÜgeln sich auch manchmal, ähm, wenn, fang (.) ich mag das auch nich', wenn, wenn der Lehrer was kopieren geht und dann soll'n die Klassensprecher dafür sorgen, dass es leise sein soll und
IP4:Mhm
S1w: und denn ha (.) sag ich auch manchen, die gehn dann einfach an die Tische [wenn die
IP4: also während] des Unterrichts gehn die Lehrer dann.
S1w: Sobald der Lehrer weg ist und ich sag, [die soll'n auf ihren Platz
IP4: Okay. Mhm.]
S1w: und das [machen die nicht.
IP4: Das ist 'n Problem.] Das notieren wir. [Min. 37:20]
... [aus Platzgründen gekürzt bis Min. 38:22]
IP4_2: Denkt noch mal nach, wenn ihr 'ne SCHWÄCHE jetzt erzählt, vielleicht ist das jetzt so in euerm Kopf, weil das gerade jetzt passiert ist. Aber uns interessiert so (.) ganz generell (.) ihr seid ja vier Jahre an dieser Schule, was (.) ihr glaubt, [was, was nicht so gut ist.
IP4: was nicht so gut ist.] SN1m.
S2m: Ähm, früher hatten wir ein Baummikado und jetzt ist da, ähm, 'ne Schaukel und das Baummikado, äh, war eigentlich viel besser [und das würd ich gern wieder haben.
IP4_2: Okay. (---) Okay.
IP4: Mhm. (---) Okay.] Mhm. [Min. 38:49]
... [aus Platzgründen gekürzt bis Min 40:41] ...
IP4: Dann ganz recht herzlichen Dank für das, was ihr uns alles erzählt habt. [Min. 40:46]

* IP4 hatte in Min. 35:33 folgende Frage gestellt: „DANN kommt die letzte Frage, ähm, wenn ihr jemand, der hier gar nicht aus der Gegend kommt, sondern ganz woanders her, aus'm, aus 'ner andern Stadt oder einem andern Land sogar, erZÄHLN sollt, WAS ist das Tollste hier an dieser Schule? Was gefällt euch hier besonders gut? Und jeder darf EINE Sache jetzt nennen. Eine Sache, was hier euch richtig gut an eurer Schule gefällt. Wer mag anfangen?“ Die Bemerkung einer Schülerin – „Ich hab auch eine Sache, was ich nich' so gern mag“ – greift IP4 mit „DAS kommt jetzt“ auf.

S1 beklagt sich über sich prügelnde Mitschülerinnen und Mitschüler und die Abwesenheit der Lehrkraft. IP4 stellt durch eine Aussagenwiederholung sicher, ob bzw. dass sie richtig verstanden hat: Lehrkräfte verlassen die Klasse während des Unterrichts. Die Schülerin bestätigt dies, indem sie ihre Beschwerde weiter ausführt, woraufhin IP4 mit „das ist ’n Problem“ die Wahrnehmung der Schülerin anerkennt. Mit „das notieren wir“ signalisiert IP4, dass sie das Anliegen der Schülerin ernst nimmt und dieses offiziell zu Protokoll genommen wird.²³ (Zugleich bleibt unklar, was genau notiert wird.) In dem darauf folgenden Interviewabschnitt (gekürzt) geht es um eine größere Schlägerei, die sich in Abwesenheit der Lehrkraft zugetragen hatte. Mit der Einordnung dieser Geschehnisse als „Schwäche“ verwendet IP4_2 einen Begriff aus dem Beurteilungsvokabular²⁴ der Schulinspektion. Sie verdeutlicht anschließend, dass die Schulinspektion nicht nur ein Interesse an aktuellen Ereignissen hat, sondern „ganz generell“ interessiert ist an dem, „was nicht so gut ist“ – also an den Schwächen. In dem Zusammenhang verweist sie darauf, dass die Schülerinnen und Schüler „ja vier Jahre an dieser Schule“ seien. Die Antwort von S1 fällt jedoch ganz konkret aus und bezieht sich auf ein früheres Spielgerät – „das Baummikado“, das als „besser“ als die jetzt dort stehende Schaukel bezeichnet wird. Auch die folgenden Antworten weiterer Schülerinnen und Schüler (gekürzt) beziehen sich auf konkrete Sachverhalte. Die Verallgemeinerungsanforderung wird somit nicht bedient, was unterschiedliche Gründe haben kann: Wie bereits anhand der Einstiegsfragen nach der Kinderkonferenz (Typ II) gezeigt, fällt es Schülerinnen und Schülern eher schwer, sich über längere Zeiträume hinweg zu erinnern. Möglicherweise gibt es aber auch nichts, das „ganz generell [...] nicht so gut ist.“ Das Interview wird auch hier mit einer Dankesformel beendet, die in diesem Fall „ganz recht herzlich“ ausfällt.

IP6 orientiert die Schülerinnen und Schüler, indem sie die „Abschlussfrage“ als solche benennt. Daran schließt sich die Aufforderung an die Schülerinnen und Schüler an, sich in eine Situation hineinzusetzen, in welcher sie einem anderen Kind je „EINE Sache“ von der Schule sagen, die sie gut bzw. nicht so gut finden. Diese im Vergleich längste Abschlussfrage (140 Wörter) enthält eine weitere Besonderheit: Zweimal ergeht an die Schülerinnen und Schüler die Aufforderung „GANZ ehrlich“ bzw. „genauso ehrlich“ zu sprechen. Auch hier soll mit der „Bitte, dass wir jetzt gleich noch mal so ’ne Reihe machen“, Übersichtlichkeit hergestellt werden. Die erste Antwort überrascht die IP:

23 An dieser Stelle wird eine strukturelle Ähnlichkeit zu polizeilichen Aufgaben und deren ent- bzw. belastender Funktion deutlich.

24 Im Schulinspektionsbericht werden definierte schulische Bereiche anhand einer vierstufigen Skala – stark/eher stark/eher schwach/schwach – beurteilt (vgl. IfBQ 2013 und den Beitrag von Diedrich zu Aufbau und Rolle der Schulinspektion Hamburg in diesem Band).

Tabelle 13: Schülerantworten zur Abschlussfrage von IP6

IP6: Okay. Dann hab ich jetzt meine Abschlussfrage und danach dürft ihr dann auch gern fragen. Ähm und zwar stellt euch mal vor, es kommt ein neues Kind hier in die Stadt (.) und das kennt Hamburg nich' und muss aber jetzt irgendwie zur Schule gehen und möchte gerne wissen, ob diese Schule gut is' oder nich'. Und ihr sollt dem Kind jetzt GANZ ehrlich sagen, was is' an dieser Schule GUT (.) und genauso ehrlich sollt ihr dem Kind aber auch sagen, was ist an der Schule vielleicht nicht GANZ so gut. Und ich hab jetzt die Bitte, dass wir jetzt gleich noch mal so 'ne Reihe machen und jedes Kind sagt EINE Sache, die es richtig TOLL findet, aber wirklich nur eine, und EINE Sache, die es, wenn euch was einfällt, NICH' so gut findet. Ja? (--) NS1w, diesmal fangen wir hier an. [Min. 48:17]
S1w: Diiie Hausaufgaben sind ganz gut.
IP6: Die HAUSaufgaben sind gut?!
S1w: [Mhmm.
IP6: Okay] und was is nicht ganz so gut?
S1w: Hm. (--) Böse Kinder.
[(<i>mehrere SuS lachen</i>)]
IP6: (<i>lacht</i>) Böse Kinder. Okay.] NS2m.
S2m: Also, wir haben auch in der Pause 'ne gute Auswahl zu den Fahrzeugen.
IP6: Mhm.
S2m: Und (--) eigentlich gibt's an der Schule alles.
IP6: Gut. [NS3w.
S3w: Ich finde] auch alles gut.
IP6: Gibt's was, was du BESONDERS gut findest?
S3w: Hm (--) Nö.
IP6: Nee. Okay. NS4w.
S4w: Ich finde (1,5 Sek.) Deutsch ganz gut (.) und (.) ich mag es nicht, wenn es Streit gibt.
IP6: Mhm, Okay.
S5?: Ich finde Sport ganz gut und Mathe.
IP6: (<i>schmunzelt</i>) Und was findest du an der Schule nicht so gut?
S5?: Weiß ich [nicht
IP6: fällt] dir nichts ein, okay. NS6m, was gefällt dir besonders gut hier?
S6m: Ich find's gut, dass man hier so 'ne große Auswahl hat, äh, dass man Hockey spielen kann [und so
IP6: Mhm.] Mhm.
S6m: Und schlecht finde ich manchmal, wenn (1,5 Sek.) irgendwie (2,5 Sek.) ach, keine Ahnung.
IP6: (<i>lacht</i>) Okay. NS7m.
S7m: Also, ich find alles an der Schule super. Vor allem, weil sie nah dran liegt. (1 Sek.) Und das schlechte is' nicht direkt in der Schule, sondern die Busse sind (.) fahrn (.) warten nich direkt auf uns, also, sie warten nich' SO lange.

IP6: Okay. Okay. NS8w.
S8w: Ich find die Lehrer gut. ((<i>unverständlich</i>)) Und (---) ja, und was ich nicht so gut finde, das weiß ich eigentlich gar nich'.
IP6: ((<i>schmunzelt</i>)) In Ordnung. NS9w.
S9w: Ich find alles sehr gut und äh (2 Sek.) und ich find eigentlich (.) nichts nich' gut.
IP6: [NS10m.
S10m: Ich find] die Bücherei ganz toll.
IP6: Die Bücherei.
S10m: Mhm. Die ich noch testen muss.
IP6: In Ordnung.
S11?: Ich finde, es gut dass wir eine friedliche Schule sind.
IP6: Mhm.
S11?: Und nich' so gut find ich das frühe Aufstehen.
IP6: ((<i>schmunzelt</i>)) Versteh ich. NS12w.
S12w: Ähm, ich find alles gut und nichts schlecht.
IP6: Wunderbar. (--) Ja, damit ist das Interview zu Ende, ganz, ganz herzlichen Dank [Min. 50:28]

Die positiven Antworten umfassen die Hausaufgaben, die gute Auswahl an Fahrzeugen, die Fächer Deutsch, Sport und Mathe, die große Auswahl (vermutlich an Aktivitäten jenseits des Unterrichts) und die Möglichkeit Hockey zu spielen, die Lehrer, die Bücherei und die Einschätzung, dass die Schule eine friedliche ist. Vier Schülerinnen und Schüler (S3, S7, S9, S12) finden alles gut, weiteren vier (S2, S5, S6 und S8) fällt nichts Negatives ein. S10 wird gar nicht mehr danach gefragt, was er nicht so gut findet. S1 nennt – in verschmitztem Tonfall – auf Nachfrage von IP6 „böse Kinder“. Die Antwort löst bei IP6 wie auch den anderen Schülerinnen und Schülern Lachen aus. Vor dem Hintergrund der insgesamt sehr positiven Rückmeldungen lässt sich schließen, dass „böse Kinder“ von der Realität dieser Schule weit entfernt sind. Insofern kann diese Äußerung, wie möglicherweise auch schon die erste Antwort von S1 („Diiiie Hausaufgaben sind ganz gut“) als ironischer Kommentar und als Unterlaufen der zweifachen Ehrlichkeitsanforderung verstanden werden.²⁵ S4 äußert, dass sie keinen Streit mag, S11 findet „das frühe Aufstehen“ nicht so gut – diese Sätze sind eher als Selbstaussagen und weniger als Kritik an der Schule zu verstehen. Die Interaktionen von IP6 beschränken sich auf aktives Zuhören und einige wenige Nachfragen. Ihr abschließender Kommentar „wunderbar“ ist vieldeutig und kann sich auf die letzte Antwort von S12, auf das Interview insgesamt oder darauf beziehen, dass

25 Zur Ironiefähigkeit von Kindern im Grundschulalter s. bspw. Kotthoff 2007.

das Interview nun zu Ende gebracht ist, wie nachfolgend expliziert wird. Der Dank wird hier „ganz, ganz herzlich“ ausgedrückt.

Die Abschlussfrage vom Typ III ist eine offene, einfach und direkt formulierte Frage, die sich auf ein anderes Befragungsinstrument der Schulinspektion bezieht. Mit fast vier Minuten fällt diese Sequenz sehr lang aus, was auf ausführliche Antworten seitens der Schülerinnen und Schüler und Nachfragen von IP2 sowie auf eine Unterbrechung, als ein Schüler mit seinem Stuhl umkippt, zurückzuführen ist. Daher wurden in diesem Fall zusammenfassende Paraphrasen eingefügt.

Tabelle 14: Schülerantworten zur Abschlussfrage von IP6

IP2: So. Wie seid ihr denn mit der/ mit dem Online-Fragebogen zurechtgekommen, war der schwierig oder ging das so? Ähm (1,5 Sek) NSw1? [Min. 51:42]
Sw1: Also manchmal, wenn ich jetzt 'ne Frage hatte, das war ein komisches Wort, was ich nicht verstanden hatte, hab ich halt 'n Lehrer geholt, aber sonst war das eigentlich immer ganz klar.
IP2: Mhm. Okay. [Min. 51:52]
<i>Zusammenfassung 1: Weitere SuS antworten, zum Teil unverständlich. Ein Sm fand es schwer mit dem Ankreuzen. IP2 will wissen, ob die SuS sich an irgendwelche Begriffe erinnern können, die schwierig waren. Das würde ihr helfen. Eine Sw nennt die Frage „Wie viel liest du am Tag?“ Fast zeitgleich in Min. 52:41 ein lautes Poltern, IP2 zieht hörbar die Luft ein, sagt, „steh erst mal in Ruhe auf“. Offensichtlich ist ein Sm mit dem Stuhl umgefallen. Es wird geklärt, ob der Sm mit dem Kopf gegen eine Schublade gefallen ist. Mehrere SuS lachen, Unruhe, weiteres Gelächter. Sw versucht, wieder den Faden aufzunehmen, aber beide IP sind noch mit Begutachtung des verunfallten Sm beschäftigt; sie stellen schließlich gemeinsam fest: „Sieht gut aus.“ Die Schülerin fährt mit ihren Ausführungen zur „Lesefrage“ fort bis Min. 53:13.</i>
IP2: Also, das fand'st du schwierig einzuschätzen.
<i>Zusammenfassung 2: Es geht weiter um die Lesefrage vom Fragebogen. Andere Schülerinnen und Schüler führen ebenfalls aus, dass die Frage schwer zu beantworten war – bis Min 54:03.</i>
IP2: Okay. (--) Seid ihr, also, haben euch die Lehrer gehOLfen, bei der (.) beim Ausfüllen des Fragebogens? NSw3?
S3w: Also bei mir (.) ich ha (.) Ja.
IP2: Und wie haben sie euch dabei geholfen?
S4w: Ich (--) hab (---) meistens, aber nicht bei allen Aufgaben, sondern ganz seltenen Aufgaben, aber ich weiß nicht mehr, welche Aufgaben, da bin ich überhaupt nicht weitergekommen. Ich hab die Wörter alle überhaupt nicht mehr verstanden.
IP2: Ja.
S4w: Und (--) auf diese Art ist (.) da war Frau NL da, also unsere Lehrerin, da hab ich die gefragt, und die hat mir die Wörter einzeln übersetzt oder so erklärt. [Min. 54:45]

Zusammenfassung 3: Es werden noch weitere SuS zu der Frage aufgerufen. Auch andere Schülerinnen und Schüler haben einzelne Wörter nicht verstanden und dann die Lehrerin gefragt, die dann jeweils durch Erklärungen geholfen hat. Zuletzt erklärt eine Schülerin detailliert, wie die Lehrerin beim Einloggen in den Online-Fragebogen geholfen hat.

IP2: Super. Technische Hilfe. Klasse. Dann dank ich euch ganz herzlich für eure Mithilfe und für das tolle Interview. Kommt gut nach Hause. [Min. 55:32]

Der Fokus der ersten Frage von IP2 liegt ganz bei den Schülerinnen und Schülern und ihren Problemen mit dem Fragebogen. Die Antworten seitens der Schülerinnen und Schüler beziehen sich auf einzelne Wörter, das Ankreuzverfahren und auf die Frage, wie viel Zeit am Tag sie mit Lesen verbringen. Die *Zusammenfassung 1* zeigt, dass das Inspektionsteam sich sorgsam und verantwortungsvoll um den umgefallenen Schüler kümmert. Währenddessen behält die Schülerin ihr Thema – die Schwierigkeiten mit der Lesefrage – fest im Blick und sorgt so dafür, dass das Gespräch nahtlos fortgeführt werden kann. Mit den sich anschließenden Kontrollfragen von IP2 wird der Blick auf die Lehrkräfte gerichtet: Wie sah deren Hilfe aus? Die zugrunde liegende Absicht zielt darauf, zu klären, ob hier nicht etwa eine Manipulation des Antwortverhaltens der Schülerinnen und Schüler durch die Lehrkräfte stattgefunden hat. Es wird also, wie mit der Einstiegsfrage von IP3, überprüft, ob Verfahrensvorschriften seitens der Lehrkräfte eingehalten wurden. S4 antwortet darauf nicht direkt, sondern zunächst mit einer Erläuterung ihrer Schwierigkeiten: „Ich hab die Wörter alle überhaupt nicht mehr verstanden.“ Solchermaßen begründet erscheint die Hilfe der Lehrerin nicht nur verständlich, sondern unumgänglich. Es kann vermutet werden, dass S4 genau verstanden hat, dass es um eine Kontrolle ihrer Lehrerin geht; ihre Antwort bietet eine Vorlage für ihre Mitschülerinnen und Mitschüler, die ähnlich antworten, sodass der Lehrkraft kein Fehlverhalten anzulasten ist. Mit den Antworten scheint IP2 zufrieden, der Inhalt der letzten Antwort wird von ihr als „technische Hilfe“ eingeschätzt und durch zwei positive Kommentare gerahmt. Auch hier wird zur Beendigung des Interviews „ganz herzlich“ gedankt. Darüber hinaus werden die Schülerinnen und Schüler mit der Formel „Kommt gut nach Hause“ aus der Situation entlassen.

4. Schülerinterviews in der Schulinspektion: Sprachliches Handeln und evaluative Zwecke

In diesem Beitrag wurden die Eingangs- und Abschlussfragen von Interviews, die im Rahmen der Hamburger Schulinspektion an sechs verschiedenen Grundschulen geführt wurden, unter der Frage untersucht, inwiefern sie für den Interview-

zweck – die Evaluation von Schulqualität – geeignet sind. Dazu wurden zunächst gesetzliche, theoretische und empirische sowie praxisbezogene Begründungszusammenhänge für den Einbezug von Schülerinnen und Schülern in das Evaluationsverfahren aufgezeigt. Anschließend wurde erläutert, wie das Schülerinterview im Verfahren der Hamburger Schulinspektion eingebettet ist.

Da die einführende Begrüßung der Schülerinnen und Schüler in den Audio-Aufnahmen nicht erfasst wird, wurde das Dokument einer Eröffnungssequenz („Setting, Warm-up“) vorgestellt. Als für die Zwecke und den weiteren Verlauf des Interviews wesentliche Inhalte und Merkmale der Kommunikation wurden herausgearbeitet: Höflichkeit, kurze und klare Sätze, Konkretheit, anschauliche Explikation der Aufgabe von Schulinspektion, Zusicherung von Vertraulichkeit und Meinungsfreiheit, Transparenz über die Ziele (Doppelfokussierung auf Gutes und nicht so Gutes), Themen und Dauer des Interviews sowie über die Funktionen aller Interviewteilnehmenden. Als dem Evaluationszweck zuträgliche Teilzwecke des Warm-up sind das Werben um Kooperation und Auskunftsbereitschaft der Schülerinnen und Schüler bzw. die Etablierung eines guten Gesprächsklimas zu sehen.

Vor der Präsentation der Fallbeispiele wurden Bedingungen und mögliche Schwierigkeiten der Konstellation „Schülerinterviews“ benannt, um eine Angemessenheit der Analyse zu gewährleisten. Die Fragestellung dieses Beitrags wurde im Kontext der funktionalen Pragmatik reformuliert. Schließlich wurden Gütekriterien für Inspektionsinterviews und Richtlinien für die Analyse vorgestellt.

Ergebnisse

Die Ergebnisse aus den Analysen der Einstiegs- und Abschlussfragen im Zusammenhang mit den Schülerantworten werden im Folgenden tabellarisch dargestellt. Anhand dieser Tabelle wird die eingangs gestellte Frage beantwortet: *Welche inspektorensseitigen sprachlichen Äußerungen im Kontext der Grundschule sind förderlich bzw. weniger förderlich sind, um die Ziele und Zwecke des Interviews – die Evaluation von Schulqualität – unter der Bedingung eines relativ engen zeitlichen Rahmens zu erreichen?*

Die empirischen Befunde – die sprachlichen Handlungen – werden kategorisiert und nach *eher förderlich* bzw. *weniger förderlich im Sinne des Evaluationszwecks* (Tabelle 15) unterschieden. Bei den Kategorisierungen geht es weniger um die Trennschärfe, sondern mehr um eine Vielfalt an Hinweisen. Die Differenzierung *eher förderlich/weniger förderlich* wurde aus der Analyse der IP-Fragen unter Einbezug der Schülerantworten abgeleitet. Um einen Nachvollzug zu gewährleisten, werden jeweils (exemplarische) Belegstellen (Indizien) genannt (Angabe der IP, der Frage, des Stichwortes oder Zitats). Wo es sinnvoll und möglich ist, werden förderliche und weniger förderliche Aspekte einander gegenübergestellt.

Tabelle 15: Sprachliche Handlungen – eher förderlich und weniger förderlich für den Evaluationszweck der Schulinspektion im Kontext Grundschule

Eher förderliche sprachliche Handlungen	Belegstelle/Indiz	Weniger förderliche sprachliche Handlungen	Belegstelle/Indiz
<i>Aktives Zuhören (Sprecher-Hörer-Steuerung)</i>	Alle IP: Verwendung von „mhm“, okay“, „ja“	<i>Vagheit von Formulierungen</i>	IP2, Einstiegsphase: „vielleicht“
<i>Eindeutige Fokussierungen</i>	IP1, 5, Einstiegsfrage	<i>Häufung von Fokussierungen oder Wechsel von Foki</i>	IP3, Einstiegsfrage: Fragenreihungen; IP2, Einstiegsphase
<i>Gegenwartsbezogenheit</i>	IP1, 5: Einstiegsfrage	<i>Vergangenheitsbezogenheit</i>	IP2, 4, 6, Einstiegsfrage: Frage nach Kinderkonferenz
<i>Einfachheit, Direktheit</i>	IP2: Abschlussfrage	<i>Aufforderung zum Perspektiven- oder Rollenwechsel (Verkomplizierung)</i>	IP1, 3, 4, 6, Abschlussfragen: „Stellt euch vor“
<i>Konkretheit</i>	IP1, Einstiegsfrage; IP5: Bitte um Beispiel	<i>Abstraktheit</i>	IP2, Einstiegsphase: Frage nach generellen Abläufen
<i>Nachfragen zur Verständnissicherung</i>	IP6, Einstiegsphase	<i>Nachbohren</i>	IP4, Einstiegsphase: Mehrfaches Fragen, Festhalten am Leitfaden
<i>Höfliche Formulierungen</i>	Alle IP: Namentliche Ansprache v. SuS; Äußerung von Dank; IP1: Konjunktiv; IP1, 5: Aufforderung als Bitte	<i>Forderung nach Ehrlichkeit</i>	IP6, Abschlussfrage
<i>Wie-Fragen als Erzählimpuls</i>	IP2, Abschlussfrage	<i>Nachfragen zur Kontrolle der Lehrkraft</i>	IP2, Abschlussphase
<i>Begrenzende Formulierungen</i>	IP1, Einstiegsfrage; IP5, Abschlussfrage: „Eine Sache“	<i>Wechsel der interviewenden Person</i>	IP5_2, Abschlussphase; IP4_2, Abschlussphase
<i>Aufforderung zur Nennung von Verbesserungen an der Schule</i>	IP1, Abschlussfrage	<i>Aufforderung zur Kritik an der Schule</i>	IP4, 6, Abschlussfrage

Für den Evaluationszweck sind der klare Fokus der IP-Frage sowie der Verzicht auf „Verkomplizierungen“ von herausragender Bedeutung. Als Einstiegsfrage er-

scheint die Frage danach, was die Schülerinnen und Schüler an ihrer Schule toll finden, besonders gut geeignet. Hervorzuheben ist außerdem, dass die Frage nach Verbesserungen der Frage danach, was nicht gut ist, vorzuziehen ist. Letztere, wie auch Kontrollfragen, die sich auf das Verhalten der Lehrkraft beziehen, bringen die Schülerinnen und Schüler in Loyalitätskonflikte, verengen den Blick auf Negatives und führen mental tendenziell zur Schließung, weil nicht nach Beteiligung gefragt wird. Demgegenüber öffnet die Frage nach Verbesserungen, sie lädt zur Beteiligung ein und weckt die Kreativität. Vor diesem Hintergrund wäre zu empfehlen, dass die Doppelfokussierung gut/nicht gut aufgegeben und verschoben wird in Richtung auf: Was ist gut, was könnte (noch) verbessert werden.

Zieht man rückblickend die von Tosana et al. (2011) genannten Gütekriterien heran, kann festgestellt werden, dass 1. „Nicht-Beeinflussung“ eingehalten wurde – es konnten keine Suggestivfragen festgestellt werden – und dass 2. „Spezifität“ gegeben war, d. h., individuelle Sichtweisen von Schülerinnen und Schülern erhielten Raum und Zeit, niemand wurde übergangen, zum Teil wurden Schülerinnen und Schüler wiederholt angesprochen. Die Erfüllung des 3. Kriteriums „Tiefe“ – die Einschätzung der unterschiedlichen Dimensionen einer Antwort – war nur dort erforderlich, wo Schülerantworten nicht eindeutig waren: Hier erfolgten Nachfragen (siehe z. B. Abschlussfrage von IP1, Antwort von S3; Abschlussfrage IP3, Antwort von S5). Das 4. Kriterium „Reichweite“ – die Erhebung aller relevanten Themen einschließlich neu auftauchender Gesichtspunkte – spielte hier aufgrund der Beschränkung auf die Einstiegs- und die Abschlussfrage keine Rolle. Auch das Kriterium 5. „personaler Kontext“ – die Berücksichtigung des individuellen Hintergrunds – fand in den analysierten Interviews keine Beachtung, da es sich um Gruppeninterviews handelt.

Betrachtet man die Antworten der Schülerinnen und Schüler mit Blick auf den jeweiligen Sozialindex der Schule, so können kaum Unterschiede in der sprachlichen Performanz festgestellt werden – es antworteten also Schülerinnen und Schüler der Schule mit KESS-Index 1 nicht weniger kompetent als die Schülerinnen und Schüler der Schule mit KESS-Index 6. Dies könnte damit zusammenhängen, dass von den Schulleitungen Schülerinnen und Schüler für das Interview ausgewählt wurden, die im Zusammenhang mit ihrer Funktion (z. B. Klassensprecher/-in) als sprach- und sprechfähig eingeschätzt wurden.

Diskussion und Ausblick

Unsere Ergebnisse bieten einen Einblick in die vielfältige Praxis der Interviewführung an Grundschulen in der Mitte des 1. Zyklus. Es konnten einfach umzusetzende Hinweise gewonnen werden, die für die Interviewpraxis des 2. Zyklus genutzt werden können. Darüber hinaus könnten die Ergebnisse als Grundlage für

die Entwicklung eines Kriterienkatalogs genutzt werden, anhand dessen eine Stichprobe von Interviews von Zeit zu Zeit systematisch überprüft werden könnte,²⁶ um weitere praxisorientierte Hinweise zu gewinnen. Da die in diesem Beitrag praktizierte Vorgehensweise recht aufwendig ist und im laufenden Geschäft der Schulinspektion nicht zu leisten sein dürfte, wäre hier z.B. an eine Vergabe von Qualifikationsarbeiten zu denken.

Vor dem Hintergrund der anfangs vorgetragenen Begründungszusammenhänge steht die Beteiligung von (Grund-)Schülerinnen und Schülern im Rahmen von *Schulentwicklung* außer Frage – fragwürdig ist aber, ob Aufwand und Ertrag von Schülerinterviews (und damit auch von deren stichprobenhafter Überprüfung) im Rahmen von *Schulinspektion* in einem vertretbaren Verhältnis stehen. „Vertretbar“ ist der Aufwand, wenn er nachweisbar und entscheidend zur Qualität der Urteile von Schulinspektion beiträgt. Dass Interviews Eingang in die Bewertung und den Bericht der Schulinspektion finden, ist zwar punktuell nachweisbar,²⁷ bislang aber nicht systematisch erfasst. Insgesamt wird die Relevanz von (Schüler-)Interviews im Kontext von Schulinspektion als gering eingeschätzt: So urteilt Pant, dass „Interviews [die] schwächste Methode der Datenerhebung“ (Pant 2012, Folie 21) seien; er schlägt daher vor, die standardisierte Befragung der Schüler/-innen im Hinblick auf deren Wahrnehmung von Unterricht zu erweitern. Für das Argument „schwächste Methode“ spricht auch die Untersuchung von Tosana et al. (2011), die anhand von Fallbeispielen zeigen konnten, dass die Interviewprotokolle regelhaft vor allem Auslassungen, daneben aber auch Verzerrungen (z. B. in Form von unzulässigen Verallgemeinerungen) aufweisen. Pant schlägt eine „Weiterentwicklung“ der Interviews dahingehend vor, dass sie zwar eine „Abschwächung der Bedeutung“ erfahren, aber weiterhin „als offene Methode“ genutzt werden sollten, um „Rückmeldungen von Stärken/Schwächen der Schule aus Sicht von Lk, SuS und Eltern“ zu erhalten (2012, Folie 29). Dem schließt sich die Autorin unter dem Vorbehalt, dass der Fokus auf Stärken/Verbesserungen verschoben wird, an. Die hier vorgestellten Hinweise könnten der von Pant angeregten Weiterentwicklung dienlich sein.

26 Die Durchführung von Interviews gilt als einer der „Standardprozesse“ von Schulinspektion, für die „periodische Evaluationszyklen festgelegt“ (Bernt & Diedrich 2014, S. 55) sind, um deren Qualität zu sichern und weiterzuentwickeln.

27 So heißt es bspw. in einem Inspektionsbericht aus dem Jahr 2011: „Zwar schätzen die Schülerinnen und Schüler der Befragung zufolge mehrheitlich ihren Leistungsstand auch selbst ein (SFB, Frage 18), doch zeigt das Interview mit ihnen, dass sich dies hauptsächlich auf die Lernentwicklungsgespräche bezieht. Nach der vorliegenden Online-Erhebung bestätigt eine knappe Mehrheit der Schülerinnen und Schüler, dass die Arbeit mit Lerntagebüchern, Kompetenzrastern oder Portfolios vorkommt (SFB, Frage 17). Im Schülerinterview ergibt sich jedoch ein anderes Bild.“

Literatur

- BBS (2006). *Orientierungsrahmen Qualitätsentwicklung an Hamburger Schulen*. Hamburg: Behörde für Bildung und Sport.
- Bernt, T. & Diedrich, M. (2014). „Ist das QM oder kann das weg?“ Qualitätssicherung in der Schulinspektion. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine daten- und theoriegestützte Schulentwicklung* (S. 51–65). Münster u. a.: Waxmann.
- Bourdieu, P. (1993). *Sozialer Sinn. Kritik der theoretischen Vernunft*. Frankfurt a. M.: Suhrkamp.
- Brünner, G. (2009). Analyse mündlicher Kommunikation. In M. Becker-Mrotzek (Hrsg.), *Mündliche Kommunikation und Gesprächsdidaktik* (S. 52–65). Baltmannsweiler: Schneider-Verlag Hohengehren.
- BSB (Hrsg.) (2008). *Jahresbericht der Schulinspektion Hamburg 2008*. Hamburg: Behörde für Schule und Berufsbildung.
- BSB (Hrsg.) (2014). *Hamburgisches Schulgesetz*. Hamburg: Behörde für Schule und Berufsbildung. Verfügbar unter: <http://www.hamburg.de/contentblob/1995414/data/schulgesetzdownload.pdf> [Zugriff 30.03.2015]
- Bührig, K. & Redder, A. (2013). Praxeogramm und Handlungsmuster als Methoden der Mehrsprachigkeitsanalyse. In A. Redder, J. Pauli, R. Kießling, B. Brehmer, I. Breckner & Y. Androutsopoulos, *Mehrsprachige Kommunikation in der Stadt. Das Beispiel Hamburg* (S. 81–104). Münster: Waxmann.
- Christen, F. (2004). *Einstellungsausprägungen bei Grundschulern zu Schule und Sachunterricht und der Zusammenhang mit ihrer Interessiertheit*. Kassel: University Press.
- Ditton, H. & Arnoldt, B. (2004). Schülerbefragungen zum Fachunterricht – Feedback an Lehrkräfte. *Empirische Pädagogik 2004*, 18 (1), 115–139.
- Ehlich, K. & Rehbein, J. (1986). *Muster und Institution. Untersuchungen zur schulischen Kommunikation*. Tübingen: Narr.
- Ehren, M. C. M. & Visscher, A. J. (2006). Towards a Theory on the Impact of School Inspections. *British Journal of Educational Studies*, 1, 51–72.
- Galinski, A. (2004). *Zweierlei Perspektiven auf Gespräche. Ethnomethodologische Konversationsanalyse und Diskursanalyse im kontrastiven Vergleich*. Verfügbar unter: http://www.linse.uni-due.de/linse/esel/pdf/konversation_diskurs.pdf [Zugriff am 30.03.2015]
- Faulstich-Wieland, H. (2005). *Spielt das Geschlecht (k)eine Rolle im Schulalltag? Plädoyer für eine Entdramatisierung von Geschlecht. Vortrag in der Reihe Gender Lectures an der Humboldt-Universität Berlin am 11.7.05*. Verfügbar unter: http://www.genderkompetenz.info/veranstaltungen_publications_und_news_archiv/genderlectures/faulstich-wieland_manuskript_genderlecture.pdf [Zugriff am 30.03.2015]
- Heintzel, F. (2010). Zugänge zur kindlichen Perspektive – Methoden der Kindheitsforschung. In B. Frieberthäuser, A. Lange & A. Prengel (Hrsg.), *Handbuch qualitative Methoden in der Erziehungswissenschaft* (3. Aufl., S. 707–721). Weinheim, München: Juventa.

- Helfferrich, C. (2011). *Die Qualität qualitativer Daten. Manual für die Durchführung qualitativer Interviews* (4. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hopf, C. (1978). Die Pseudo-Exploration – Überlegungen zur Technik qualitativer Interviews in der Sozialforschung. *Zeitschrift für Soziologie*, 7 (2), 97–115.
- IfBQ (Hrsg.) (2013): *Musterbericht der Schulinspektion*. Verfügbar unter: <http://www.hamburg.de/contentblob/4018006/data/pdf-bewertung-musterbericht.pdf> [Zugriff am 30.03.2015]
- Kameyama, S. (2004). *Verständnissicherndes Handeln*. Münster u. a.: Waxmann.
- Kotthoff, H. (2007). *Ironieentwicklung unter interaktionslinguistischer Perspektive*. Verfügbar unter: <http://www.inlist.uni-bayreuth.de/issues/45/> [Zugriff am 30.03.2015]
- Lamnek, S. (2010). *Qualitative Sozialforschung*. Weinheim: Beltz.
- Merkens, H. (2007). Auswahlverfahren, Sampling, Fallkonstruktion. In U. Flick, E. v. Kardorff & I. Steinke (Hrsg.), *Qualitative Forschung: ein Handbuch* (5. Aufl., S. 286–299). Reinbek: Rowohlt.
- Merton, R. K. & Kendall, P. L. (1993). Das fokussierte Interview. In C. Hopf & E. Weingarten (Hrsg.), *Qualitative Sozialforschung* (S. 171–204). Stuttgart: Klett-Cotta.
- Nietzschmann, R. & Vieluf, U. (2006). Evaluation. Schulentwicklungsdaten zur Sprache bringen. *Pädagogik* 2006, (3), 30–34.
- Pant, H.A. (2012). Was macht die Professionalität von Schulinspektoren aus? Beitrag auf der Tagung der Deutschen Gesellschaft für Bildungsverwaltung (DGBV) vom 2.–3. März. Verfügbar unter: http://www.dgbv.de/mediapool/118/1188921/data/Tagung_2012-03-02_Pant.pdf [Zugriff am 30.03.2015]
- Pietsch, M., Bensen, M. & Bos, W. (2006). Ein Index sozialer Belastung als Grundlage für Rückmeldungen von Leistungsergebnissen an Schulen und Klassen und für ‚faire Vergleiche‘ von Grundschulen in Hamburg. In W. Bos & M. Pietsch (Hrsg.), *KESS 4 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen* (S. 225–245). Münster u. a.: Waxmann.
- Pruiken, C. (2005). *Interessen und Hobbys hochbegabter Grundschulkinder*. Münster u. a.: Waxmann.
- Teaching and Learning Research Briefing* (2003). University of Cambridge. Verfügbar unter: http://www.tlrp.org/pub/documents/no5_ruddock.pdf [Zugriff am 30.03.2015]
- Tosana, S., Lambrecht, M., Perels, F. & Bardowiecks, S. (2011). Die Protokollierung von Interviews im Rahmen von Schulinspektionen. In S. Müller, M. Pietsch & W. Bos (Hrsg.), *Schulinspektion in Deutschland* (S. 57–78). Münster u. a.: Waxmann.
- Weber, P. & Becker-Mrotzek, M. (2012). *Funktional-pragmatische Diskursanalyse als Forschungs- und Interpretationsmethode*. Verfügbar unter: http://www.fallarchiv.uni-kassel.de/wp-content/uploads/2012/06/weber_mrotzek_diskurs_ofas.pdf [Zugriff am 30.03.2015]

Der Blick aus der Praxis – die Schulformexpertinnen und -experten

Peter Schulze & Christiane von Schachtmeyer

Zusammenfassung

Die Autorin, Leiterin eines Gymnasiums und Schulformexpertin, und der Autor, hauptamtlicher Schulinspektor, stellen das Hamburger Modell der Schulformexpertinnen und Schulformexperten (SFE) vor. SFE sind Mitglieder einer Schulleitung, die bei der Inspektion einer anderen Schule der gleichen Schulform als drittes Mitglied neben zwei Hauptamtlichen das Team der Schulinspektion vervollständigen. Sie bringen dabei vor allem den Blick aus dem Leitungsallday in das Inspektionsteam ein. Eingangs wird das Modell vorgestellt, aus Sicht der Schulinspektion werden Stärken und Herausforderungen aufgezeigt. Anschließend werden aus Sicht der Schulformexpertin die Mitwirkung im Inspektionsteam und der Nutzen für die eigene Arbeit dargestellt.

„Guten Morgen, wir sind das Team der Schulinspektion und freuen uns schon, Ihre Schule kennenzulernen!“ So begrüßen die Inspektorinnen und Inspektoren oft alle Interessierten, die sich am ersten Inspektionstag vor Unterrichtsbeginn im Lehrerzimmer versammelt haben. Dabei geben die Gäste von der Inspektion nicht nur einen kurzen Überblick über die bevorstehenden Unterrichtsbesuche, die Inspektionstage und das Selbstverständnis der Inspektion, sondern sie stellen sich auch persönlich vor. Die Lehrkräfte interessiert dabei oft vor allem die Qualifikation der einzelnen Inspektorinnen und Inspektoren. Auch wenn die folgenden Fragen selten so direkt gestellt werden, oft meint man, sie aus skeptischen Blicken zu errahnen: „Verfügen die Inspektorinnen und Inspektoren über genügend Praxiserfahrungen? Können sie überhaupt die alltäglichen Anforderungen unserer Schule einschätzen?“ Aufmerksam wird das Kollegium gerade dann, wenn sich die Kollegin oder der Kollege aus der Schulleitung einer anderen Schule als „Schulformexpertin“ oder „Schulformexperte“ vorstellt.

Auch bei Gästen der Inspektion, vor allem solchen aus den Inspektoraten anderer Bundesländer, erregt das Modell der „Schulformexpertinnen/Schulformexperten (SFE)“ häufig großes Interesse. Sie ist eine „Hamburgensie“, eine Besonderheit, die es nur in Hamburg gibt. Die Erfahrungen mit dem SFE-Modell sind überwiegend positiv.

In diesem Beitrag berichten wir – eine Schulformexpertin und ein hauptamtlicher Inspektor – über dieses Hamburger Modell der Schulformexpertinnen und

-experten. Eingangs stellen wir die Genese und die Ziele vor. Wir bilanzieren Vorteile und Kritikpunkte. Anschließend berichten wir aus der Praxis.

1 Einführung in das Modell der Schulformexpertinnen und -experten – aus der Perspektive eines Inspektors

1.1 Wer sind Schulformexpertinnen und Schulformexperten?

In Hamburg besteht das Team, das eine Schule inspiziert, mindestens aus drei Personen: Die beiden hauptamtlichen Inspektorinnen und Inspektoren werden von einer Schulformexpertin bzw. einem Schulformexperten begleitet. SFE besitzen dabei immer Leitungserfahrung in der inspizierten Schulform. An der Inspektion eines Gymnasiums nimmt beispielsweise immer ein Mitglied des Schulleitungsteams (Schulleiterin/Schulleiter, Stellvertretung oder Abteilungsleitung) eines anderen Hamburger Gymnasiums teil. In der Einsatzplanung der SFE wird darauf geachtet, dass diese nicht eine Schule aus dem gleichen Schulbezirk inspizieren. So bauen wir Befangenheiten durch Vorwissen vor und verhindern, dass überhaupt die Befürchtung aufkommt, die Konkurrenz zwischen Nachbarschulen könnte die SFE in ihrer Bewertung beeinflussen.

Die SFE nehmen fast an allen Inspektionsschritten teil: Sie können, müssen aber nicht, das Vorgespräch in der Schule mit durchführen, sie lesen die eingereichten Dokumente der Schule, interpretieren die Fragebogenergebnisse, führen beim Schulbesuch Unterrichtsbeobachtungen durch und nehmen an den Interviews teil. Nach einer individuellen Vorbewertung werten sie im Inspektionsteam schließlich die Ergebnisse mit aus, legen die Bewertung fest, lesen den Inspektionsbericht gegen und können an der schulöffentlichen Rückmeldung teilnehmen. In dem Verfahren und in der Bewertung sind die SFE gleichberechtigte Teammitglieder. Einzig an der Rückmeldung der Ergebnisse an die Schulleitung nehmen sie nicht teil. Auch führen sie nicht verantwortlich die Interviews, können diese aber durch eigene Fragen ergänzen. Ähnliches gilt für das Schreiben des Berichts und die Durchführung der schulöffentlichen Präsentation, die jeweils Aufgaben der Teamleitung sind.

In der Anfangszeit wurden Schulleitungen auf Empfehlung der Schulaufsichten angesprochen, ob sie Interesse hätten, SFE zu werden. Heute wird man SFE, indem man gegenüber der Inspektion sein Interesse anmeldet. Sie werden etwa drei Tage für ihre Aufgabe fortgebildet. Themen sind dabei u. a. der Auftrag, die Funktion und das Selbstverständnis der Inspektion, das Qualitätsverständnis und die Bewertungsgrundlage und natürlich die Arbeitsweise und die verschiedenen Instrumente der Inspektion, wozu auch eine Schulung zur Unterrichtsbeobachtung gehört. Bei ihrem ersten und zweiten Einsatz werden neue Schulformexpertinnen und -experten durch die Teamleitung besonders unterstützt und begleitet.

Ein paar Zahlen: Seit dem Start der Schulinspektion Hamburg im Jahr 2007 bis zum Sommer 2014 haben etwa 140 SFE Schulen inspiziert. Etwa 45 SFE haben dabei mindestens 4 Inspektionen durchgeführt, in einem Fall sogar 12. Gleichzeitig haben ebenfalls 45 SFE nur an einer einzigen Inspektion teilgenommen haben. Die Gründe dafür sind vielfältig und reichen von zu starker zeitlicher Belastung in der eigenen Schulleitung über den Stellenwechsel oder die Pensionierung und das Warten neuer SFE auf den zweiten Einsatz bis hin zu – in Einzelfällen – abweichenden Vorstellungen von der Inspektionstätigkeit. Von den 400 Hamburger Schulen wird damit rechnerisch jede dritte von einer Leitung (mit)geführt, die Inspektionserfahrung hat, und in jeder zehnten Schule gibt es jemanden mit Inspektionsroutine. Tatsächlich sind jedoch in vielen Schulen mehrere Leitungsmitglieder SFE, in einem Hamburger Gymnasium sogar 5 der 6 Leitungsmitglieder. Die Zahl der bereits ausgebildeten SFE, die (noch) nicht zum Einsatz gekommen sind, liegt im unteren zweistelligen Bereich.

1.2 Exkurs: Wirtschaftsvertreterinnen und -vertreter, Unterrichtsbeobachterinnen und -beobachter

Bei den Inspektionen der etwa 40 berufsbildenden Schulen in Hamburg ergänzt neben den beiden hauptamtlichen Inspektorinnen und Inspektoren, der Schulformexpertin bzw. dem Schulformexperten auch noch ein viertes Teammitglied das Inspektionsteam: die Wirtschaftsvertreterin bzw. der Wirtschaftsvertreter (WV). Dieses sind Personen, die die Perspektive der Ausbildungsbetriebe mit in das Inspektionsteam einbringen können, weil sie in der Wirtschaft tätig sind, oftmals eben in der Ausbildung. Die insgesamt 12 Wirtschaftsvertreterinnen und -vertreter werden ansonsten ähnlich wie die SFE qualifiziert und eingesetzt.

An der Inspektion allgemeinbildender weiterführender Schulen nehmen als vierte Personen auch Unterrichtsbeobachterinnen und -beobachter teil. Ihr Einsatz ist notwendig, damit an zwei Tagen ausreichend viele Unterrichtsbeobachtungen durchgeführt werden können. Sie ergänzen das Team, nehmen aber ausschließlich an den Unterrichtsbeobachtungen teil. Auch sie werden hierfür durch Mitglieder der Inspektion speziell geschult.

1.3 Wie entstand die Idee, SFE zu integrieren?

Die Konzeption der Hamburger Schulinspektion entwickelte eine Projektgruppe, in der 2004 und 2005 Vertreterinnen und Vertretern der Schulbehörde, des Landesinstituts und drei Schulleitungen gemeinsam arbeiteten. Die Inspektionsteams sollten dabei multiprofessionell besetzt sein. Aufgrund ihrer unterschiedlichen

beruflichen und fachlichen Hintergründe würden so die Teammitglieder sehr verschiedene Sichtweisen auf die inspizierten Schulen einbringen und den Blickwinkel des Inspektionsteams insgesamt deutlich erweitern. Die dann eingestellten hauptamtlichen Inspektorinnen und Inspektoren brachten ein Bündel an Qualifikationen aus unterschiedlichen Professionen in das Inspektionsteam ein: als Schulleitungen, Lehrkräfte oder Schulaufsichten, als wissenschaftliche Mitarbeiterinnen und Mitarbeiter in der Schulforschung, als Beraterinnen und Berater zu Qualitätsmanagement und Schulentwicklung oder als Referent/-innen in der Fort- und Weiterbildung. Ziel dieser Multiprofessionalität war und ist es, neben dem Wissen von Schulpraktikerinnen und -praktikern die Erfahrung schulnahen Fachpersonals in der Inspektionsarbeit zu nutzen. Eine Spezialisierung auf einzelne Schulformen, wie sie in anderen Flächenländern teilweise üblich ist, schied aber auch schon aus organisatorischen Gründen aus, da die Anzahl der Schulen einiger Schulform im Stadtstaat Hamburg zu gering ist, um ein Inspektionsteam nur diese Schulform inspizieren zu lassen.

Auf die Frage „Wie können wir bei der Multiprofessionalität des Teams auch eine ausreichende schulformspezifische Kompetenz im jedem Inspektionsteam gewährleisten?“ fand die Projektgruppe mit der Integration von SFE die passende Antwort. Ein ähnlicher Ansatz wurde mit den Wirtschaftsvertreterinnen und -vertretern als Expertinnen und Experten für die berufsbildenden Schulen geschaffen.

Bereits in den Überlegungen der Projektgruppe, die von Behörde und Politik übernommen worden sind, spielten viele der heutigen Funktionen und Vorzüge des SFE-Modells eine Rolle.

1.4 Vorzüge des Modells

1. Praxisexpertise für die Schulform

Die Schulformexpertinnen und -experten kommen aus der Leitung einer Schule der gleichen Schulform. Sie bringen sehr detaillierte Erfahrungen und Kenntnisse aus der alltäglichen Praxis mit in das Inspektionsteam. Damit ergänzen und bereichern sie den Blick des Inspektionsteams. Die Hauptamtlichen haben dabei neben ihrer eigenen Fachlichkeit vor allem den breiten Überblick über die Praxis in vergleichbaren Schulen und sind Experten für die inhaltlichen Standards und Verfahren der Inspektion.

2. Höhere Akzeptanz

In Hamburg war die Einführung der Inspektion anfangs in den Schulen umstritten. Nach Wahrnehmung der Inspektion hat sich das jedoch überraschend schnell geändert. Von den Schulformexpertinnen und -experten erhoffte man sich eine

größere Akzeptanz in den Schulen allgemein sowie bei den Personalräten und den Gewerkschaften im Besonderen. Die SFE erhöhten vor allem bei einigen Kritikerinnen und Kritikern die Akzeptanz des Inspektionsteams. Heute genießen die multiprofessionellen Teams allgemein eine hohe Akzeptanz. Dazu trägt auch bei, dass alle hauptamtlichen Inspektorinnen und Inspektoren zuvor lange im Bildungswesen gearbeitet und sich durch langjährige Arbeit in der Inspektion ein profundes Wissen über die Praxis in Hamburgs Schulen erarbeitet haben.

3. Vermittler und Multiplikatoren in das System

Viele Schulformexpertinnen und -experten sind gut mit anderen Schulen, der Schulbehörde und dem Landesinstitut vernetzt und somit wichtige Multiplikatorinnen und Multiplikatoren. Durch ihr vertieftes Verständnis des Qualitätsbegriffs und der Arbeitsweise der Inspektion wirken sie formell und informell als Vermittlerinnen und Vermittler in ihrem Umfeld. Sie werden so auch zu wichtigen Partnerinnen und Partnern der Inspektion in den Schulen und der Behörde.

4. Feedback an die Inspektion

Die SFE geben durch ihre Mitwirkung der Inspektion immer wieder Rückmeldung zu deren Arbeit und setzen so Entwicklungsimpulse. Sie sind zu kritischen Freundinnen und Freunden, zu langjährigen Begleiterinnen und Begleitern geworden, die so auch die Qualitätsentwicklung der Hamburger Schulinspektion mit fördern. Wenn es beispielsweise Neuerungen im Inspektionsverfahren, wie etwa die Ratingkonferenz als neue Befragungs- und Interviewform im zweiten Inspektionszyklus, oder eine Veränderung im Format der Inspektionsberichte gibt, dann ist das Feedback der SFE oft direkt, manchmal unbequem und gerade deshalb besonders hilfreich.

5. Routinen aufbrechen

In den Inspektionsteams, in denen die Hauptamtlichen oftmals schon sehr häufig zusammen inspiziert haben, gewährleisten SFE immer wieder einen „frischen“ Blick auf Routinen. So diskutieren sie beispielsweise Items der Unterrichtsbeobachtung oder hinterfragen manchmal, ob die Maßstäbe, die die Inspektion auf Grundlage der Anforderungen des Orientierungsrahmens Schulqualität an die Personalentwicklung anlegt, angesichts der großen Arbeitsbelastung von Leitungen in der Praxis realistisch sind. Die SFE stoßen so immer wieder Reflexionen in den Inspektionsteams über die eigene Arbeit, die Beurteilungsmaßstäbe und die Ansprüche an eine gute Schule an. Diese Reflexionen eröffnen neue Perspektiven und helfen, die Arbeit der inspizierten Schule besser einordnen zu können. Die SFE melden den Hauptamtlichen jedoch ebenfalls immer wieder zurück, wie beeindruckend und bereichernd sie das stete gemeinsame Ringen um das tiefere Verständnis einer Schule finden. Auch persönlich ist die Zusammenarbeit mit den

SFE für die Hauptamtlichen oft sehr bereichernd. Die Teams pflegen nicht selten eine sehr herzliche Atmosphäre.

6. Inspektion als Personalentwicklung

Eine wahre Erfolgsgeschichte des SFE-Modells liegt vor allem in einem vermeintlichen Nebenaspekt begründet: Die Inspektionstätigkeit stellt für die SFE auch eine Form der Personalentwicklung dar. Viele SFE berichten, dass sie durch die Inspektionserfahrung nun ihre Schule anders leiten und ein vertieftes Verständnis für ihre Arbeit entwickelt haben. Ob bei der Organisation der Förderung oder dem Aufbau der Konferenzstruktur, fast in jeder Inspektion kommt der Moment, wo die oder der SFE nach einem Interview schnell zum Notizblock greift, um eine gerade gewonnene Best-Practice-Idee für die eigene Schulleitungsarbeit zu sichern. So ist neben der mehrtägigen Qualifizierung von Schulleitungsmitgliedern zu den Verfahren und Standards der Inspektion vor allem die praktische Tätigkeit während der Inspektion eine Weiterbildung besonderer Art. Wenn man eine andere Schule anhand der standardisierten Kriterien der Inspektion evaluiert, die auf dem Orientierungsrahmen Schulqualität, der Schulforschung und den politischen Schwerpunktsetzungen basieren, dann gibt das auch viele Impulse für die eigene Leitungspraxis. Die Gesamtwirkung der Inspektion durch diese Qualifizierung der SFE, die vermeintlich ein Nebenaspekt der eigentlichen Inspektionsarbeit ist, dürfte nicht gering sein und verdient ggf. eine eingehendere Evaluation.

1.5 Kritik und Herausforderungen

Wo Licht ist, da ist auch Schatten, und natürlich bringt das SFE-Modell auch Herausforderungen mit sich.

1. Peer Review: Akzeptanzprobleme

Vor allem das Team des Peer Reviews¹ 2012 äußerte sich teilweise kritisch über das SFE-Modell. Hierzu gehört die unterschiedliche Akzeptanz der SFE durch die

1 Ein Peer Review ist eine Form externer Evaluation. 2012 haben vier Bildungsexpertinnen und -experten, davon zwei Leitungen von Inspektionen anderer Bundesländer, eine Büroleiterin eines anderen Kultusministeriums und eine Professorin, auf Einladung des IfBQ die Schulinspektion Hamburg evaluiert. Dabei hat das IfBQ Fragen an das Peer-Review-Team formuliert und ihm zahlreiche Dokumente zur Verfügung gestellt. Das Peer-Review-Team hat nicht nur mit den Mitarbeiterinnen und Mitarbeitern der Inspektion Interviews geführt, sondern auch mit wichtigen Partnern wie Schulleitungen, Schulformexperten, Schulaufsichten und der Leitung sowie der Aufsicht des IfBQ. Anders als im Inspektionsverfahren ist es bei einem Peer Review üblich, dass Empfehlungen ausgesprochen werden (vgl. Düring et al. 2012).

Leitungen der inspizierten Schulen: „Die Qualität dieser Personengruppe in der Ausgestaltung ihrer spezifischen Rolle wird nicht durchgängig positiv bewertet (vor allem von den Schulleiterinnen und Schulleitern nicht).“ (Düring et al. 2012, S. 16). Diese Rückmeldung der Peers nach dem Interview mit Schulleitungen ist bemerkenswert, weil sie bei der Evaluation der Inspektionen nahezu nie an das Inspektionsteams oder das Inspektorat zurückgemeldet wird (vgl. Pietsch 2011).

2. *Begrenzter Beitrag*

Die Peers äußern sich weiterhin kritisch zu der beschränkten Perspektiverweiterung durch die SFE: „Ihre fachlich-pädagogische Beratung und damit das Einbringen ihrer Expertise beschränken sich auf die punktuelle Mitarbeit in einzelnen Inspektionsteams während der Schulbesuche.“ (Düring et al. 2012, S. 16). In der Praxis ist der Beitrag der SFE durchaus unterschiedlich. Vor allem unter dem Gesichtspunkt der Funktion „Inspektion als Form der Personalentwicklung“ (siehe oben) nehmen neue SFE bei ihrer ersten Inspektion häufig zunächst eine eher beobachtende Position ein.

3. *Organisation und Ressourcen*

Die Planung des Einsatzes der Inspektionsteams ist mit den SFE, den WV und den Fachleitungen wesentlich aufwendiger, als würden nur hauptamtliche Inspektorinnen und Inspektoren eingesetzt.

Eine Besonderheit ist manchmal die Inspektion einer Schule, die von einem/einer SFE geleitet wird. Denn die Schulleitung kennt sehr genau die Arbeit des Inspektionsteams. Sie ist mit den Standards und Verfahren selbst gut vertraut. Die Schulleitungen mit eigener Inspektionserfahrung begleiten und hinterfragen die Arbeit des Inspektionsteams teilweise verständnisvoller, manchmal aber auch kritischer, auf jeden Fall fundierter.

2 Praxisbericht – aus der Perspektive einer Schulformexpertin

2.1 Was motiviert SFE zu dieser Arbeit?

Schulleiterinnen und Schulleiter in Hamburg haben, v.a. seit Einführung der Selbstverantworteten Schule (2006), ein vielfältiges Aufgabenfeld und können über mangelnde Arbeit nicht klagen. Die Teilnahme an einer Schulinspektion führt dazu, dass man fast drei Tage nicht in der Schule ist. Diese Tage sind sehr anregend, aber auch sehr arbeitsintensiv, gefüllt mit 5 Interviewrunden und ca. 20 Unterrichtsbesuchen à 20 Minuten. Zur Entlastung erhält man sogenannte WAZ

(Wochenarbeitszeit-Stunden), die der Stundenbilanz der eigenen Schule zugutekommen. Diese Entlastung ist allerdings nicht langfristig einkalkulierbar und wird als Motivation eher eine untergeordnete Rolle spielen. Daher kann man davon ausgehen, dass Schulleitungen diese zeitintensive Arbeit nur tun, wenn sie sich auch einen persönlichen Ertrag erhoffen. Im Prinzip gilt das, was unter dem Aspekt „Personalentwicklung“ schon beschrieben wurde.

Wann hat man schon, außer bei der Tätigkeit als SFE, die Gelegenheit, derartig intensiv eine Schule kennenzulernen und einen Blick hinter die Kulissen zu bekommen? Wann kann man sonst so viel Unterricht beobachten und einen so gründlichen Einblick in didaktische Standards Hamburgs erhalten? In jeder Schule bekommt man darüber hinaus einen Strauß von Anregungen unterschiedlichster Art geboten, von pfliffigen organisatorischen Lösungen, pädagogischen Haltungen bis hin zu interessanten Raumgestaltungen, die auf direkte oder indirekte Art Einfluss auf die Leitungstätigkeit haben. So war für meine Leitungstätigkeit die Inspektion eines Hamburger Gymnasiums mit sehr großem Migrationsanteil besonders prägend. Hier hat mich besonders der Stolz der gesamten Schulgemeinschaft auf die Leistungen ihrer Schülerinnen und Schüler beeindruckt. Diese Grundhaltung war und ist für mich ein großes Vorbild. Ein weiterer positiver Aspekt ist, dass einem immer wieder der Orientierungsrahmen für Schulqualität, d. i. der Bewertungsmaßstab der Schulinspektion, ins Gedächtnis gerufen wird. So ist eine Schulinspektion auch immer wieder ein Spiegel, der der eigenen Arbeit vorgehalten wird. Allerdings ist es im Alltag dann schwierig, die frisch gewonnenen Erkenntnisse umzusetzen; das ist ein langwieriger Prozess, für den man dann die notwendige Energie aufbringen muss. Groß ist das Risiko, dass einen der Alltag zu schnell im Griff hat. Das bedeutet, dass man für sich als Schulformexperte ein Konzept entwickeln muss, wie man die Eindrücke aus den Inspektionen unter Wahrung der Vertraulichkeit weiterverarbeiten möchte. Nützlich ist für mich ein kleines Notizbuch zum Notieren von Eindrücken und Formulieren von Fragestellungen, die ich dann auch in meine Leitungsrunde bringe, sodass man durch die Inspektionen auch wieder einen Blick auf die eigenen wunden Punkte wirft: Wie ist es bei uns selbst um die Kooperationsstrukturen bestellt? Wie sieht es bei uns mit der Implementierung von Feedback aus?

Ingeheim hoffen sicherlich die SFE auch, durch ihre Arbeit auf die Inspektion ihrer eigenen Schule besser vorbereitet zu sein. Das könnte ein fragwürdiges Motiv sein, ist es aber nicht, wenn man wie oben beschrieben die Inspektion nutzt, um praktisch von außen einen Blick auch auf die eigene Schule zu werfen. Meine persönliche Erfahrung zeigt jedoch, dass diese Hoffnung nur z. T. begründet ist. Zwar kennt man die Instrumente der Schulinspektion sehr gut, doch als zu Inspizierende hat man eine ganz andere Rolle und ist gefangen im eigenen System. Deswegen sind die gewonnenen Erfahrungen und Kenntnisse nur begrenzt übertragbar. Das habe ich deutlich gespürt, als die Schulinspektion in unserem

Hause zu Gast war und ich nun selbst interviewt wurde. Obwohl ich alle Fragen schon häufig gehört hatte und auch schon viele Antworten auf diese Fragen kannte, fielen mir die Antworten auf einige Fragen dennoch nicht leicht. Dasselbe gilt auch für den Umgang mit den eigenen Inspektionsergebnissen. Sicherlich weiß man viel über die Vorgehensweise und das Entstehen des Inspektionsergebnisses. Das kann, muss aber nicht zwangsläufig die Akzeptanz des eigenen Ergebnisses erhöhen.

2.2 Die Arbeit als SFE im Inspektionsteam

Die Arbeit im Inspektionsteam ist geprägt von klaren Ritualen, hoher Professionalität (vgl. den Beitrag von Sowada in diesem Band) und gegenseitiger Wertschätzung. Die Arbeit der Schulinspektion ist bedeutsam und seriös, sie ist interessant, anregend und fordernd. Sie fordert von einem selbst hohe Professionalität, Kriterienklarheit und Kondition. Ein Inspektionstag kann intensive zwölf Stunden dauern mit hohen Anforderungen an Aufmerksamkeit und Professionalität. Man vergisst dabei nicht, dass so lange in der eigenen Schule die Arbeit liegen bleibt. Die Inspektionsarbeit ist ausgesprochen gewinnbringend, wegen der Vielzahl von Impressionen, Erkenntnissen und Kontakten. Es wäre schade, wenn man auf sie verzichten müsste. Es gibt eigentlich nur wenige Schwierigkeiten und diese liegen in der Konstruktion des Schulformexperten per se begründet. Man ist im Inspektionsteam für die schulformspezifische Kompetenz zuständig und muss darauf achten, dass man nicht nur die eigene Schule als Maßstab nimmt. Außerdem weiß man natürlich, dass die eigene Schule und damit auch man selbst früher oder später inspiziert wird. Man sollte also sehr zurückhaltend in der Kommunikation über alles sein, was bewertungsrelevant für die eigene Schule ist. Das gilt auch für das persönliche Verhältnis zu den hauptberuflichen Inspektoren, das bei aller Kollegialität auch von einer gewissen professionellen Distanz geprägt ist. So ist z. B. das Duzen ein Tabu. Insgesamt erfordert die Tätigkeit als SFE eine große Professionalität von jeder Schulformexpertin und jedem Schulformexperten. Die ist dann noch einmal besonders gefordert, wenn die Schulinspektion im eigenen System vor der Tür steht.

2.3 Wie gelingt der Rollenwechsel Schulleitung – Inspektorin/Inspektor – Kollegin/Kollege im Kreis anderer Schulleitungen?

Fordernder und anspruchsvoller ist es meiner Erfahrung nach, den Kontakt mit den zu inspizierenden Schulleitungskolleginnen und -kollegen angemessen zu gestalten. Dabei ist es eine kleine Erleichterung, dass man keine Nachbarschulen

inspiziert. Gelöst wird das Problem dadurch aber nicht, denn in einem kleinen Bundesland mit rund 60 Gymnasien und insgesamt 400 Schulen wie Hamburg hat man zu vielen Schulen und Schulleitungen gute Kontakte.

Nicht jede Schulleitung betrachtet die Schulinspektion als begrüßenswertes Instrument der externen Evaluation und selbst wenn dies doch so sein sollte, heißt man die Schulinspektion vielleicht mit einem klammen Bauchgefühl willkommen. Man weiß als SFE also nie, wie man wahrgenommen wird – als Unterstützer oder gar als Verräter? Bei besonders schlechten Hospitationsergebnissen, schwierigen Kollegien oder auffälligen Fehlentscheidungen der Leitung kann es schwierig werden, der Leitung gegenüber neutral aufzutreten. Besonders belastet kann das Verhältnis werden, wenn man an einem aus Sicht der Schule schlechten Inspektionsergebnis beteiligt war und die Schule die Ursache dafür nicht in der eigenen Arbeit, sondern dem Team der Inspektion wöhnt. Es kann daher sinnvoll sein, auf die Anwesenheit des SFE bei der schulöffentlichen Präsentation des Inspektionsergebnisses zu verzichten. Wie tritt man dann dem Kollegen beim nächsten Aufeinandertreffen gegenüber?

Die hier formulierten Schwierigkeiten treten in der Praxis aber nur selten auf. Dabei helfen die Leitungserfahrung und der klar geregelte Ablauf der Inspektion.

3 Gemeinsames Fazit

Die Hamburgensie „Schulformexpertin/Schulformexperte“ hat sich als Modell bewährt. Es steigert die Qualität der Inspektionen, weil durch diesen Ansatz zusätzliches Know-how in das Inspektionsteam kommt und der Blickwinkel des Teams erweitert wird. Die Resonanz in den Schulen ist überwiegend positiv, wenngleich es auch vereinzelte Kritik gibt. Die Schulformexpertinnen und -experten erleben diese Tätigkeit als persönliche Bereicherung und gewinnen aus den Einblicken oftmals wichtige Impulse für die eigene Leitungsarbeit. So wirkt Schulinspektion auch auf ganz anderen Wegen als in anderen Bundesländern auf die Leitungspraxis in vielen Hamburger Schulen. Gleichwohl müssen die Schulformexpertinnen und -experten auch lernen, mit dem Rollenkonflikt umzugehen, mal Mitglied eines Inspektionsteams zu sein und bei der nächsten Schulleiterdienstbesprechung oder Schulleitungsfortbildung wieder als Leitungskollege einer anderen Hamburger Schule aufzutreten.

Literatur

- Düring, K., Friede, A., Perels, F., Stenke, D. (2012). *Bericht zum Peer Review der Schulinspektion Hamburg 27. bis 29. Februar 2012*. Verfügbar unter: <http://www.hamburg.de/contentblob/4022662/data/pdf-peer-review-bericht-zur-qualitaet-der-schulinspektion.pdf> [Zugriff am 27.1.2015]
- Pietsch, M. (2011). *Nutzen und Nützlichkeit der Schulinspektion Hamburg. Ergebnisse der Hamburger Schulleiterbefragung*. Verfügbar unter: <http://www.hamburg.de/contentblob/4022654/data/pdf-zufriedenheitsstudie-2011.pdf> [Zugriff am 27.1.2015]

Teil 3

Die Befunde der Schulinspektion

Evidenzbasierte Referenzrahmen zur Schulqualität als Grundlage von Schulinspektion

Melanie Ehren & Jaap Scheerens¹

Zusammenfassung

Dieser Beitrag beschreibt, wie Schulinspektorate unterschiedliche Ziele (Kontrolle, Entwicklung und Verbesserung, Verbindung zwischen verschiedenen Akteuren im Schulsystem) im konzeptionellen Rahmen der Inspektionen und ihrer Indikatoren operationalisieren. Es bietet einen Überblick über die Indikatoren, die in einer Reihe von Ländern eingesetzt werden, und zeigt beispielhaft, wie sie in den konzeptionellen Rahmen von Inspektionen einfließen und die Schulen mit dem Ziel der Kontrolle, Entwicklung und Verbindung bewerten und beurteilen. Wir beschreiben die Inspektion und Bewertung der Prozesse und Ergebnisse schulischer Bildung (wozu auch gehört, Expertenmeinungen zu formulieren) und vergleichen sie mit Inspektionskonzepten, deren Hauptaugenmerk auf der Kontrolle von Input-Erfordernissen und der Einhaltung von gesetzlichen Vorschriften liegt. Außerdem erörtern wir, wie sinnvoll und angemessen unterschiedliche Konzepte im Hinblick auf aktuelle Ergebnisse der Schulwirksamkeitsforschung sind.

Schulinspektionen können unterschiedliche Ziele verfolgen, z. B. Kontrolle, Unterstützung und die Verbindung zwischen Lehrerinnen und Lehrern, Schulen und den zuständigen Stellen im Bildungssystem. *Kontrolle* knüpft an die ursprüngliche Bedeutung des Begriffs „Inspektion“ an und steht im Zentrum einer Überwachung der Einhaltung gesetzlicher Vorgaben. In vielen Ländern gilt die Kontrolle als die wesentliche Aufgabe von Schulinspektorinnen und -inspektoren. Im Mittelpunkt steht die Überprüfung der gesetzlich vorgeschriebenen Erfordernisse, Regeln und Pflichten. *Unterstützung* geht mit dem Ziel der Entwicklung einher, sie konzentriert sich auf eine Bewertung der Schulqualität, gibt Rückmeldungen an die Schule, was ihre Stärken und Schwächen betrifft, und macht Vorschläge, wie sich die Schwächen beseitigen lassen. Kontrolle und Unterstützung und regelmäßige Schulbesuche haben oftmals Auswirkungen auf die dritte Funktion der *Verbindung* von Schulinspektionen, bei der sie als Bindeglied zwischen der Spitze des Bildungssystems (wo Normen und Regeln festgelegt werden) und den Schulen selbst fungieren (wo Bildung stattfindet und konkretisiert wird). Inspektoren und Inspektorinnen neh-

1 Aus dem Englischen übersetzt von Rita Kloosterziel.

men die Rolle der Mittler ein: Sie haben einerseits die Aufgabe, die Schulen über Beschlüsse der zentralen Stellen in Kenntnis zu setzen, und müssen andererseits die Entscheidungsträger über die Realität auf Schulebene informieren.

Laut De Grauwe (2007) spiegeln diese drei Ziele unterschiedliche Systeme im Qualitätskonzept wider, das der Inspektion zugrunde liegt, und machen deutlich, worauf jeweils das größte Gewicht gelegt wird. Die Kontrolle von Schulen beinhaltet meist eine Evaluation der Input-Indikatoren und eine Überprüfung, inwieweit die Schulen gesetzliche Vorgaben einhalten. Dagegen bewerten Schulinspektorate, die darauf abzielen, eine Entwicklung von Schulen zu fördern und anzuregen, häufig die Bildungsprozesse und den Output der Schulen. Liegt der Schwerpunkt auf dem Aspekt der Verbindung, spielen Schulinspektorate vor allem die Rolle des Vermittlers, indem sie Wissen und Informationen an die entsprechenden Stakeholder innerhalb des Schulsystems weitergeben. Evaluationen mit einem bestimmten thematischen Schwerpunkt, bei denen Schulinspektorate Informationen zu spezifischen, für die Richtlinien relevanten Themen sammeln (z. B. wie neue Richtlinien umgesetzt werden oder wie das Bildungssystem in einem bestimmten Bereich abschneidet), sind Beispiele für diese Mittlerrolle eines Schulinspektorats.

1. Kontrolle von Input, Regeln und Bestimmungen

De Grauwe (2007) und Eddy Spicer et al. (2014) stellen dar, wie Inspektions- und Monitoringsysteme den Input einer Schule in den Fokus stellen können, also etwa die Anzahl der Lehrbücher pro Schüler oder Schülerin, die Ausbildung der Lehrerinnen und Lehrer, die Klassenstärke usw. Bei solchen Systemen geht es in erster Linie darum, die Einhaltung von Regeln zu kontrollieren. Sie wollen sicherstellen, dass sich Schulen an gesetzlich festgelegte Normen halten und auf der Verwaltungsebene in Übereinstimmung mit den gültigen Verordnungen handeln. Dazu gehört, dass Leitlinien und Vorschriften etwa in Bezug auf Zugangs- oder Sicherheitsbestimmungen und immer öfter auch auf das zufriedenstellende Ausfüllen der Unterlagen zur Selbsteinschätzung durch die Schulen vorliegen und angewandt werden. Als Beispiele seien das schwedische Schulinspektorat, das kontrolliert, in welchem Maße Schulen allen Schülerinnen und Schülern den gleichen Zugang zu Bildung ermöglichen, und das niederländische und das irische Schulinspektorat genannt, die überprüfen, ob Schulen ein Minimum an Unterrichtsstunden einplanen und anbieten (vgl. Ehren et al. 2013). Nach De Grauwe (2007) entspricht diese Art der Kontrolle der ältesten bürokratischen Form des Monitorings: überprüfen, ob die Regeln und Bestimmungen respektiert werden. Das klassische Inspektoratsystem in Verbindung mit unterschiedlichen Formen administrativer Selbstberichte durch Schulen (Formulare ausfüllen) ist das wesentliche Instrument, auf das sich diese Art des Monitorings stützt.

2. Evaluation und Förderung von Bildungsprozessen

Zu Bildungsprozessen zählen die Qualität des Unterrichts an einer Schule, die Interaktion zwischen Lehrerinnen und Lehrern, Schülerinnen und Schülern und Curriculum während des Unterrichts und die administrativen Organisationsprozesse innerhalb der Schule. Als Teil des konzeptionellen Rahmens von Inspektionen werden diese Prozesse immer bedeutsamer, da man übereinstimmend davon ausgeht, dass bei dem Versuch, Unterschiede in der Qualität von Schulen zu erklären, Prozessvariablen wichtiger sind als Input-Variablen. Und Informationen über die Qualität einer Schule sind notwendig, um die Qualität von Schulen insgesamt zu verbessern. Das gilt vor allem in Ländern mit hohem Einkommen, wo sich der Input der Schulen nur geringfügig voneinander unterscheidet (siehe Hanushek 1997).

Eine Übersicht von Van Bruggen (2010) zeigt z. B., dass achtzehn europäische Schulinspektorate Indikatoren und Kriterien zu „Organisation und Management in der Schule“ und „Lehren und Lernen“ in ihren konzeptionellen Rahmen aufgenommen haben, um eine nationale Perspektive auf gute Bildung zu gewährleisten und um Schulen vor dem Hintergrund eines gemeinsamen Kriterienbündels zu beurteilen, das eine solche Perspektive auf gute Bildung repräsentiert. Laut Ehren et al. (2013) orientieren sich viele dieser Konzepte an Schulwirksamkeitsforschungen. Ihre vergleichende Studie zu Inspektionskonzepten aus sechs europäischen Ländern belegt eine starke Konzentration auf Bildungsprozesse wie die Lernmöglichkeiten und Lernzeit, Leistungsorientierung, klare und gut strukturierte Unterrichtsführung, anspruchsvolle didaktische Ansätze und eine geordnete Lernumgebung. Diese Indikatoren sind in gewissem Maße Teil der Inspektionskonzepte aller sechs Länder, die Ehren et al. in ihrer Studie untersuchen (England, die Niederlande, das österreichische Bundesland Steiermark, die Tschechische Republik, Irland und Schweden).

3. Evaluation des Outputs von Schulen

Die zunehmende Verfügbarkeit und Verwendung von Daten zu Schülerleistungen haben dazu geführt, dass das Hauptaugenmerk bei Schulinspektionen immer mehr auf dem Output von Schulen liegt. Einige der Schulinspektorate (z. B. in den Niederlanden) wählen diesen Fokus auch, um die Autonomie der Schulen in Bezug auf die Gestaltung ihrer schulischen Prozesse zu respektieren. Ein verstärktes Interesse am Output von Schulen geht oft Hand in Hand mit dem allgemeinen Wunsch auf Regierungsseite, Leistungsmessungen vorzunehmen und Rankinglisten anzulegen, um auf diese Weise den Wettbewerb der Schulen untereinander zu fördern (vgl. Tolofari 2005).

Das Ziel dieser Art des Monitorings ist es, sicherzustellen, dass Schülerinnen und Schüler das lernen, was sie lernen sollen, und den Output der Schule im Zusammenhang mit Lernfortschritten der Schülerinnen und Schüler zu evaluieren. Dazu gehören z. B. die Schülerleistungen in einer Reihe von Unterrichtsfächern und die Anzahl der Schulabschlüsse bzw. Schulabbrüche. Die Evaluation von Leistungen kann auch eine Zielverschiebung verhindern, wenn sich Schulen auf die Vorlage von Vorgaben und Prozeduren konzentrieren, um die Schulinspektorate über ihre schulischen Prozesse zu unterrichten. Die Ergebnisse von Schülerleistungen in landesweit standardisierten Lernstandserhebungen und Vergleichsarbeiten werden aggregiert, um die Leistung von Schulen zu evaluieren. In manchen Ländern werden auch Rankinglisten erstellt.

In jüngster Zeit haben einige Schulinspektorate (z. B. in Norwegen, Schottland und den Niederlanden) damit begonnen, auch soziale Outcomes von Schulen zu evaluieren. Soziale Outcomes werden von Ehren und Dijkstra (2014) definiert als „der individuelle und kollektive Nutzen von Bildung bei interpersoneller Interaktion in den nicht-ökonomischen Lebensbereichen“. Auf den schulischen Bereich bezogen bestehen die sozialen Outcomes von Bildung nach Angaben der Autoren in der Kompetenz, mit anderen zusammenzuleben, und aus den sozialen Kompetenzen, die Menschen benötigen, um ihre Ziele zu erreichen und in einer Vielzahl von Situationen mit anderen zu agieren, sowohl im Beruf als auch in anderen Zusammenhängen. Zu den sozialen Outcomes von Bildung gehören auch die staatsbürgerlichen Kompetenzen, die es braucht, um einen Beitrag zu Gesellschaft, Demokratie und den gesellschaftlichen Netzwerken zu leisten, in denen Menschen leben. Soziale Outcomes werden in den konzeptionellen Rahmen von Inspektionen aufgenommen, um ein breiter gefächertes Bild vom Output einer Schule zu erlangen und zu verhindern, dass sich Schulen allein darauf beschränken, Rechnen, Lesen und Schreiben zu lehren. Man geht davon aus, dass Testergebnisse aus kognitiven Leistungstests ein unvollständiges Bild von den Kompetenzen junger Menschen zeichnen. In vielen Ländern herrscht die Ansicht vor, dass der Kanon der Kompetenzen und Fertigkeiten weiter gefasst sein sollte, wenn Schülerinnen und Schüler all das lernen sollen, was sie brauchen, um auf dem Arbeitsmarkt zu bestehen und aktive Mitglieder der Gesellschaft sein (vgl. Dijkstra et al. 2014).

Die am weitesten verbreiteten Verfahren, die bei der Evaluation des Outputs von Schulen zum Einsatz kommen, sind die regelmäßigen Messungen von Schülerleistungen in Form von standardisierten Tests und Prüfungen, in Verbindung mit der Veröffentlichung von Rankinglisten und einem systematischen (externen) Audit von Schulen. Bei den Tests geht es vor allem um kognitive Domänen wie Mathematik und Literacy, während soziale Kompetenzen häufiger mithilfe von Beobachtungen im Schulalltag gemessen werden (vgl. ebd.). Erst seit Kurzem haben einige Länder wie die Niederlande damit begonnen, landesweit standardisierte Tests zur Messung sozialer Kompetenzen einzuführen. Bisher hat nur eine

begrenzte Anzahl von Schulen diese Tests eingesetzt, sodass auch die Benchmarks und Ziele, mit denen die Schulen im Rahmen dieser Maßnahmen verglichen und evaluiert werden könnten, noch begrenzt sind.

In einigen Ländern werden neuerdings Indikatoren für den Output von Schulen verwendet, um die Inspektionsbesuche in Schulen im Rahmen von „risikobasierten Inspektionen“ gezielter zu planen. Solche risikobasierten Inspektionsmodelle fußen auf der Annahme, dass der Output einer Schule (wie er anhand von Schülerleistungen gemessen wird) ein adäquater Prädiktor für die Unterrichtsqualität dieser Schule ist und in Risikoanalysen und Früherkennung von Problemen eingesetzt werden kann, um potenziell versagende Schulen zu ermitteln. Schülerergebnisse (um die sozioökonomischen Faktoren bereinigt) in standardisierten Tests und Prüfungen werden in den Niederlanden z.B. verwendet, um Schulen in eine von drei Kategorien (grün, orange, rot) einzuordnen. Von Schulen in der grünen Kategorie nimmt man an, dass bei ihnen kein Risiko zu scheitern vorliegt. Schulen in der Kategorie Orange weisen ein potenzielles Versagensrisiko auf. Rote Schulen haben dagegen ein hohes Risiko zu versagen. Schulen in der grünen Inspektionskategorie werden lediglich einer einfachen Inspektion unterzogen, was bedeutet, dass es in dem Jahr, in dem die Kategorisierung vorgenommen wurde, keine weiteren Aktivitäten im Zusammenhang mit einer Inspektion gibt. Zusätzlich wird eine Sichtung und Analyse von Selbsteinschätzungen und anderen Dokumenten der Schule (*Desk Research*) für die Schulen im orangefarbenen Bereich angesetzt, während für Schulen in der roten Kategorie sofort Inspektionsbesuche geplant werden.

Der verstärkte Fokus auf den Output von Schulen rechtfertigt eine detailliertere und differenziertere Analyse und Verwendung von Daten zu Schülerleistungen, um die Genauigkeit der Bewertung der Inspektionsergebnisse und die Verlässlichkeit der Frühwarnanalysen zu verbessern. Aus diesem Grund haben einige Schulinspektorate unlängst damit begonnen, Messungen des *Value Added* vorzunehmen, um die Wirkung zu erfassen, den die Einzelschule und ihre Lehrerinnen und Lehrer auf die Qualität der Bildungsarbeit haben. *Value-Added*-Messungen bedienen sich mathematischer Algorithmen bei dem Versuch, den Beitrag der Schule zur Lernleistung der Schülerinnen und Schüler getrennt von allen anderen Faktoren zu betrachten, die schulische Leistungen und Fortschritte beeinflussen können, z.B. Fähigkeiten des einzelnen Schülers, Höhe des Familieneinkommens, Bildungshintergrund der Eltern oder der Einfluss von Peergroups. Das Ziel dieses *Value-Added*-Ansatzes ist es, „die Auswirkungen einzelner Lehrerinnen und Lehrer oder Schulen auf Schülerleistungen zu erfassen und dabei die Unterschiede im sozialen und familiären Hintergrund der Schülerinnen und Schüler in Betracht zu ziehen“ (ASA 2014, S. 1; Übersetzung: R.K.).

Hamilton und Koretz (2002, S. 23) unterscheiden zwischen zwei Möglichkeiten, Testergebnisse so darzustellen, dass die Leistung der Schule nachvollziehbar wird (wenn sie nach Schultyp aggregiert wird). In normorientierten Berichten wird

die Leistung einer Schule in Bezug auf ihre Stellung innerhalb einer Verteilung von Punktwerten anderer Schulen beschrieben. Solche Berichte können folgende Basis haben:

1. *Prozentrang*, der den prozentualen Anteil einer Referenzgruppe anzeigt (wobei es sich oft um die nationale Schülerpopulation in Schulen handelt), die eine niedrigere Punktzahl erzielt hat als eine bestimmte Einzelschule. Eine Schule, die im nationalen Vergleich einen durchschnittlichen Prozentrang (*National Percentile Rank* – NPR) von 75 erzielt, erreicht also eine höhere Punktzahl als 75 Prozent der Schulen, die in der nationalen Stichprobe vertreten sind.
2. *Standardwert*: Stellt die Leistung einer Schule so dar, dass man sieht, wie weit das Testergebnis vom Mittelwert/Durchschnittswert entfernt ist. Die Punktzahlen werden transformiert, sodass sie einen bestimmten Mittelwert und eine Standardabweichung (SD) aufweisen. Beispiele sind z-Werte (Mittel = 0, SD = 1), T-Werte (Mittel = 50, SD = 10) und *Normal Curve Equivalents* (NCE) (Mittel = 50, SD = 21,06). Eine Schule mit einem T-Wert von 60 liegt also um eine Standardabweichung über dem Mittel, was ungefähr einem Prozentrang von 84 entspricht.
3. *Klassenstufenäquivalente* stellen dar, wie die Leistung eines Schülers oder einer Schülergruppe im Hinblick auf die Klassenstufe einzuordnen ist, für die eine solche Leistung typisch wäre. Die Ergebnisse werden meist als Dezimalzahl dargestellt, z. B. 5,7. Dabei steht die Zahl vor dem Komma für die Klassenstufe und die Zahl hinter dem Komma für den Monat (wenn man von einem zehnmonatigen Schuljahr ausgeht, steht die Null für die Leistung der Schülerinnen und Schüler, die ganz am Anfang dieser Klassenstufe stehen). Ein Schüler oder eine Schülerin, der oder die einen Wert von 5,7 in einem Test für Viertklässler erreicht, weist dasselbe Leistungsniveau auf wie ein durchschnittlicher Schüler und oder eine Schülerin im siebten Monat der fünften Klasse, der denselben Test absolviert. Klassenstufenäquivalente stellen eine *Entwicklungsskala* dar, mit deren Hilfe Wachstum untersucht werden soll. Der mittlere Leistungszuwachs im Laufe eines Wachstumsjahres beträgt in Fach und Stufe grundsätzlich 1,0 Klassenstufenäquivalent.

Die Alternative zum normorientierten Bericht (aggregierter) Testwerte ist die kriteriums- oder lehrzielorientierte Darstellung. Bei dieser Art des Berichts geht es nicht um einen Vergleich mit anderen Schulen oder Gruppen von Schulen. Vielmehr wird die Schulleistung mit einem oder mehreren zuvor definierten Leistungsniveaus verglichen. Solche festgelegten Leistungsniveaus schließen normalerweise Zielvorstellungen im Hinblick auf minimale Testwerte und Stoffe ein, die die Schülerinnen und Schüler in spezifischen Fächern und Bereichen bewältigen können sollen.

Bei der Zielsetzung unterscheiden Hamilton und Koretz (2002) zwischen zwei breiten Ansätzen: Statusmessungen und Veränderungsmessungen. Die Statusmessung vergleicht die Leistung einer Einheit zu einem bestimmten Zeitpunkt mit einem einzigen Standard, wobei es sich um ein vom Schulinspektorat vorgegebenes Leistungskriterium wie die durchschnittliche Leistung vergleichbarer Schulen (z. B. mit ähnlichen Schülerpopulationen) oder einen historischen Durchschnitt (z. B. den Durchschnitt einer Gruppe von Schulen über einen Zeitraum von fünf Jahren hinweg) handeln kann.

Dagegen vergleichen Veränderungsmessungen die Leistung einer Einheit zu einem bestimmten Zeitpunkt mit der Messung einer früheren Leistung. Veränderung lässt sich durch einen Querschnittsansatz messen, bei dem die aktuellen Viertklässler mit den Viertklässlern aus dem Jahr zuvor verglichen werden. Man kann auch einen quasi-längsschnittlichen Ansatz wählen, bei dem die aktuellen Viertklässler mit den Drittklässlern aus dem Jahr zuvor verglichen werden. Daneben gibt es noch den Längsschnittansatz, bei dem die Leistung einzelner Schülerinnen und Schüler dazu genutzt wird, Schülerinnen und Schüler über einen bestimmten Zeitraum hinweg mit sich selbst zu vergleichen. Die Ziele quantifizieren das Ausmaß der Veränderung, das man von den Schulen erwartet. Nach Hamilton und Koretz (2002) lassen sich Testwerte auf der Schul-, Klassen- und Fächerebene sowie auf der Ebene spezifischer Schülergruppen darstellen. Die Entscheidung, ob die Werte der Schulebene, der Klassen-/Fächerebene oder der Schülerebene dargestellt und ob für spezifische Gruppen disaggregiert werden soll, sollte nach Angaben dieser Autoren von der Überlegung abhängen, für welche Zwecke die Testergebnisse gedacht sind und welche Informationen die Stakeholder wünschen. Bei jedem dieser Modelle lassen sich die Testwerte an eine Reihe von Merkmalen von Schulen und Schülerinnen/Schülern anpassen, wobei die enge Beziehung zwischen Schülerleistungen und soziökonomischem Status sowie anderen Aspekten des Hintergrunds der Schülerinnen und Schüler in Betracht gezogen wird.

Beispiele für Schulinspektorate, die *Value-Added*-Messungen verwenden, lassen sich in den Niederlanden finden, wo Schülerleistungsdaten in separaten Leistungsgruppen auf der Grundlage des Grads der Benachteiligung klassifiziert werden (der sich meist am Bildungsstand der Eltern bemisst). Diese Klassifizierung wird benutzt, um den Output von Schulen zahlenmäßig zu bestimmen und einzustufen. Sie fließt auch in die Frühwarnanalysen ein, um potenziell versagende Schulen („*Failing Schools*“) für Inspektionsbesuche zu markieren. *Value-Added*-Messungen sind auch im konzeptionellen Rahmen von Schulinspektionen in England, Kanada und Irland vorgesehen. In Irland hat z. B. das *Department of Education and Skills* eine Strategie zur Verbesserung der Leistungen im Lesen, Schreiben und Rechnen entworfen. Der Entwurf empfiehlt die Verwendung eines Analysewerkzeugs für Benchmark-Daten, das „*Schools Like Ours*“ (etwa: „Schu-

len wie unsere“) genannt wird. Es soll einer Einzelschule nicht nur den Zugang zu ihren eigenen Daten ermöglichen, sondern auch zu denen vergleichbarer Schulen. In Kanada hat das *Literacy and Numeracy Secretariat* des Bildungsministeriums von Ontario ein Benchmark-Modul entwickelt, das ebenfalls mit „*Schools Like Ours*“ betitelt ist und bei dem mit beliebigen Kombinationen der verfügbaren Indikatoren gearbeitet wird, z. B. ähnliche demographische Voraussetzungen, aber bessere Leistungen.

Dass *Value-Added*-Indikatoren im konzeptionellen Rahmen von Schulinspektionen neuerdings mehr Beachtung finden, hängt mit der Notwendigkeit zusammen, die Beurteilung der Schulpraxis auf eine korrekte Verwendung der Ergebnisse der Schülerleistungen zu stützen. Scheerens, Glas und Thomas (2003) stellen z. B. fest, dass mehr Informationen über einzelne Schülerinnen und Schüler, Untergruppen von Schülerinnen/Schülern und allen Schülerinnen/Schülern einer Einzelschule sowie vergleichbare Daten für eine ganze Population (bzw. eine repräsentative Stichprobe) von Schulen eine zuverlässigere und aussagekräftigere Analyse von Schülerleistungsergebnissen ermöglicht (vgl. Scheerens et al. 2003).

Donaldson und Johnson (2010) geben jedoch an, dass noch Unsicherheit herrscht, wie groß der Beitrag tatsächlich ist, der Schulen beim Lernzuwachs ihrer Schülerinnen und Schüler zugeschrieben werden kann. Sie gehen davon aus, dass sich solche Modelle noch im Entwicklungsstadium befinden und daher fehleranfällig sind. Trotzdem sehen viele Schulinspektorate den Nutzen einer Entwicklung von *Value-Added*-Messungen, um die Zuverlässigkeit und Stichhaltigkeit ihrer Bewertungen zu verbessern, vor allem, wenn man sie mit ihren derzeit gebräuchlichen, eher groben Methoden vergleicht. Dabei werden die Schulleistungen mit dem durchschnittlichen Rohwert einer Population verglichen oder Angaben zu kostenlosem Schulmittagessen oder andere Daten zum sozioökonomischen Hintergrund der Schülerinnen und Schüler verwendet, um Schulen in separate ähnliche Leistungsgruppen zu klassifizieren und zu vergleichen.

4. Einsatztauglich: Die Modellierung von Schuleffektivität und drei Funktionen der Schulinspektionen

Viele Inspektionssysteme haben ihre Zielsetzung in den letzten Jahren dahingehend geändert, dass sie auf eine Verbesserung von Lehren und Lernen hinarbeiten. Dieses Ziel der Schulentwicklung hat in den letzten Jahren im Zuge der verstärkt verfolgten Strategie, Schulen autonomer und selbstbestimmter zu machen, an Bedeutung gewonnen. Einem hohen Maß an Schulautonomie steht in manchen Ländern die systematische Bewertung von Schulen gegenüber, um die Qualität und

Effektivität von Entscheidungen auf Schulniveau sicherzustellen. Nachlassende Schülerleistungsergebnisse, wie sie bei internationalen Vergleichen wie PISA und TIMSS zutage getreten sind, haben auch oft zu einem Anstieg an Beurteilungen und Kontrollen von Schulen geführt, selbst in vermeintlich dezentralisierten Bildungssystemen.

Wenn Schulinspektionen ihren Zweck erfüllen sollen, müssen die Art der Inspektion und vor allem die Standards im konzeptionellen Rahmen der Inspektion an die beabsichtigten Ziele des verbesserten Lehrens und Lernens und schlussendlich der verbesserten Schülerleistung anpasst sein. Die Fachliteratur zu Bildung und Schuleffektivität bietet wichtige Hilfestellungen bei der Definition einer „guten Schule“ und bei kritischen Überlegungen, inwieweit Inspektionsstandards in unterschiedlichen Ländern durch Forschungsergebnisse belegt sind. Ganz allgemein formuliert kann man „Bildungseffektivität“ als den Grad der Zielerreichung eines Bildungssystems verstehen. Ein Bildungssystem mag ein landesweites Bildungssystem sein, eine Einzelschule, eine Gruppe von Schülerinnen und Schülern oder sogar ein einzelner Schüler oder eine einzelne Schülerin. Im Zusammenhang mit Schulinspektionen konzentrieren wir uns hier auf Schulen und Schuleffektivitätsforschung als Kernthema.

Schuleffektivitätsforschung versucht, sich mithilfe wissenschaftlicher Methoden mit den kausalen Aspekten zu beschäftigen, die dem Konzept der Effektivität innewohnen. Anhand des Outcomes wird eine Einschätzung der Wirksamkeit von Schule vorgenommen, vor allem aber wird die Tatsache beleuchtet, dass Unterschiede in der Wirksamkeit von Schule beeinflussbaren Voraussetzungen zugeschrieben werden, zu denen sowohl Input als auch Prozesse gehören. Normalerweise ist die Einschätzung von Schulwirkung das Ergebnis von Vergleichen, wenn z. B. die durchschnittlichen Leistungswerte der Schulen miteinander verglichen werden. Leistungswerte in Kernfächern, die am Ende eines festgelegten Programms erfasst werden, stellen die wahrscheinlichsten „Schulwirkungen“ dar, obwohl man auch alternative Kriterien heranziehen kann, etwa die Zufriedenheit der Lehrerinnen und Lehrer oder die Art, wie die Schule mit ihrer Umgebung interagiert. Um die „Nettowirkung“ beeinflussbarer Bedingungen wie den Einsatz verschiedener Lehrmethoden oder eine bestimmte Art des Schulmanagements zu erfassen, bedarf es einer Anpassung der Leistungsmessungen, sodass sie die unterschiedlichen Aufnahmepraktiken von Schulen berücksichtigen. Zu diesem Zweck werden Merkmale des Hintergrunds der Schülerinnen und Schüler wie der sozioökonomische Status, die allgemeine Lernfähigkeit oder die Ausgangsleistungen in einem Fach als Kontrollvariablen eingesetzt. Diese Art der statistischen Anpassung, wie sie in Forschungsarbeiten vorgenommen wird, findet eine praktische Entsprechung in dem Bemühen, „faire Vergleiche“ zwischen Schulen anzustellen und den *Value Added* zu bemessen (siehe oben, vgl. Scheerens 2013, S. 4).

Was Schulwirksamkeitsforschung und die Wissensgrundlage, die sich aus dieser Forschungsrichtung für die Schulinspektion an sich ergeben hat, miteinander verbindet, ist in erster Linie ein Interesse am Outcome von Schule als dem ultimativen Qualitätsstandard. Für den Bereich der Schulinspektion ist die Schulwirksamkeitsforschung jedoch vor allem aus einem Grund besonders wichtig: Sie liefert empirisch gestützte Informationen zu der Frage, welche beeinflussbaren schulischen Bedingungen den größten Einfluss auf den Outcome von Schule haben, und bietet so eine wissenschaftlich fundierte Begründung für die Wahl der Input- und Prozessindikatoren.

Wenn wir von den drei grundlegenden Funktionen von Schulinspektionen ausgehen, die wir zu Beginn dieses Kapitels angeführt haben, so ist es vor allem der Bereich der Entwicklung und Verbesserung, der von einer Verknüpfung von Schulwirksamkeit und Schulinspektionen profitiert. Wenn Schulinspektionen nicht nur über die Wirkungen von Schule berichten, sondern auch über die Input- und Prozessindikatoren, könnte man erwarten, dass ein Feedback zu diesen Indikatoren unmittelbare Ansatzpunkte für Handlungen bietet, die zu einer Schulentwicklung führen. Wenn z. B. ein Schulprozessindikator wie die Verbindung zwischen Schulcurriculum und Aufgaben oder Prüfungen (oft als „Lerngelegenheiten“ angegeben) einen niedrigen Wert aufweist, wäre eine Optimierung der Übereinstimmung von Unterrichtsinhalt und Aufgabeninhalt eine plausible Vorgehensweise, um die Schulleistung zu steigern.

Beim Monitoring als einer der Funktionen von Schulinspektionen ist die Verbindung zur Schulwirksamkeit komplexer. Erstens gibt es keine Verbindung zwischen der Überwachung der Einhaltung von Regeln und der Wirksamkeit von Schule, soweit es die Erfüllung von standardmäßigen Vorschriften und Bestimmungen betrifft. Einhaltungskontrollen passen jedoch zum Verständnis von Wirksamkeit, wenn grundlegende Inputs in Schulen evaluiert werden, die sich unmittelbar auf den Outcome von Bildung auswirken. Solche Inputs sind leicht zu benennen, z. B. Lehrerqualifikationen, zahlenmäßiges Schüler-Lehrer-Verhältnis und festgeschriebene Unterrichtszeiten. Inwieweit solche Inputmessungen einen Unterschied ergeben, hängt unter anderem vom Kontext ab. Meist zeigen solche Inputs in Entwicklungsländern deutlichere Auswirkungen als in industrialisierten Ländern, da in industrialisierten Ländern normalerweise alle Schulen über diese grundlegenden Inputs verfügen und die Schulen sich in Bezug auf Lehrerqualifikationen oder reguläre Unterrichtszeiten kaum voneinander unterscheiden (vgl. Hanushek 1997). So stellt Hanushek (1986, S. 1161) fest, dass die Variable „Berufserfahrung der Lehrerinnen und Lehrer“ als einzige unabhängig vom Kontext einen zuverlässigen Einfluss auf den Outcome von Schule hat: 30 Prozent der geschätzten Koeffizienten sind statistisch signifikant. Hanushek kommt zu dem Schluss, dass Ausgaben im Bildungsbereich anscheinend nicht durchgehend im Zusammenhang mit Leistung stehen

und dass wichtige Effekte erst bei größeren Unterschieden in den Inputs zu erwarten seien.

Was die Verbindungsfunktion von Schulinspektionen angeht, so besteht lediglich ein eher theoretischer Zusammenhang zu Schulwirksamkeit und Schulinspektionen, und zwar insofern, als sich Mehrebenen-Modelle von Schulwirksamkeit vorrangig mit einer Angleichung von Facetten und Elementen befassen, die auf verschiedenen Ebenen operieren (siehe z.B. Scheerens 2007; Creemers und Kyriakides 2008). Wenn man Bildungssysteme als hierarchisch geordnete Systeme betrachtet, kann man Schulwirksamkeit von Unterrichtswirksamkeit und von Systemeffektivität unterscheiden. Dieser letztgenannte Begriff ist nicht besonders weit verbreitet, er gehört zu einer neueren Forschungsrichtung, die ihre Impulse vor allem aus dem starken Anstieg internationaler Schulleistungsuntersuchungen bezieht. Bei solchen Untersuchungen können von politischen Entscheidungen beeinflussbare Bedingungen im Rahmen des nationalen Systems mit Schülerleistungen in Verbindung gebracht werden. Beispiele für solche Entscheidungen sind Regelungen zur Förderung der Autonomie, *Accountability* und Wahlmöglichkeiten. Unterrichtswirksamkeit konzentriert sich auf effektives Lehrerhandeln im Rahmen des Unterrichts. Diese Strömung der Bildungsforschung zeichnet sich dadurch aus, dass Prozessmerkmale von Bildung auf Lehrer- oder Unterrichtsebene untersucht werden. Wenn wir uns also mit Variablen dieser Ebene beschäftigen, bei denen sich eine Verbindung zu Leistung gezeigt hat, tauchen wir in die primären Prozesse schulischer Ausbildung ein. Schulwirksamkeit richtet ihr Augenmerk dagegen auf die Voraussetzungen, die eine effektive Schule mitbringt, und befasst sich mit der Frage, welche Bedingungen für einen Zuwachs bei Schülerleistungen (*Value Added*) sorgen. Das Ziel besteht im Allgemeinen darin, Schulmerkmale aufzuspüren, die sich ganz sicher mit dem Output einer Schule in Verbindung bringen lassen. Dieser Output wird meist anhand von Schülerleistungen gemessen.

Schulwirksamkeit meint die Verschmelzung der Wirksamkeitsforschungen auf diesen drei Ebenen. Konzeptionelle Beiträge zu diesem Forschungszweig sprechen von Schulen als Ansammlungen von „*Nested Layers*“ (Purkey und Smith 1983), ineinander geschachtelten Schichten. Ihnen liegt die zentrale Annahme zugrunde, dass höhere organisatorische Levels effektivitätsfördernde Bedingungen auf niedrigeren Levels begünstigen (vgl. Scheerens und Creemers 1989). Die Mehrebenenanalyse hat einen wesentlichen Beitrag zur Entwicklung solcher integrierten Schulwirksamkeitsmodelle geleistet (vgl. Scheerens 2013).

Diese Vorstellung von Schulwirksamkeit und ineinander geschachtelten Klasebenen innerhalb der Schulebene, die wiederum Teil eines nationalen Bildungssystems ist, knüpft an die Verbindungsfunktion von Schulinspektionen an. Sie greift die besonderen Verbindungen zwischen den verschiedenen Levels auf, die Inspektionen bei ihrer Bewertung von Unterricht, Schulen und Bildungssystem

und mit ihren Informationen für die Akteure auf diesen Ebenen schaffen können. Schulinspektorate könnten die Abstimmung und Verknüpfung der ineinander geschachtelten Ebenen fördern, indem sie z. B. eine Kohärenz nationaler curricularer Konzepte mit Evaluation, Inspektion und Beurteilungskonzepten anstoßen. Theoretische Modelle zu Bildungseffektivität erkennen jedoch auch an, dass es bei vielen Bildungssystemen lediglich eine lose Kopplung zwischen den Schichten des Bildungssystems gibt. Außerdem herrschen in den meisten Ländern verschiedene Erwartungen vor, wie weit nationale Inspektorate als Instrumente der Zentralverwaltung fungieren und eine enge Kopplung fördern sollen. In manchen Fällen geht man davon aus, dass Inspektorate unabhängig oder nahezu unabhängig funktionieren und bei der Evaluation von Schulen ihrer eigenen Agenda folgen. Soweit diese Art der systemischen Angleichung bis jetzt erforscht wurde (z. B. Mourshed et al. 2010; Scheerens et al. 2015), lässt sich feststellen, dass die Funktionsweise von Inspektoraten bisher nicht thematisiert wurde.

5. Identifizierung von effektivitätsfördernden Schulvoraussetzungen; Übereinstimmung in Reviews

Der wesentliche Teil der Schuleffektivitätsforschung besteht in der Identifizierung von effektivitäts- und verbesserungsorientierten Bedingungen. Im folgenden Abschnitt werden neuere und ältere Forschungsreviews genannt und auf Übereinstimmungen in Bezug auf die wichtigsten effektivitätsfördernden Voraussetzungen untersucht. Eine solche Zusammenfassung ermöglicht den Vergleich der grundlegenden Forschungen zur Schuleffektivität mit Inspektionskonzepten in einem späteren Abschnitt.

Scheerens (2014) fasst die Ergebnisse von Review-Studien aus den 1990er Jahren und Review-Studien jüngerer Datums zusammen (z. B. Reynolds et al. 2014; Muijs et al. 2014 und Hopkins et al. 2014). Bei den älteren Reviews handelt es sich um die von Purkey und Smith (1983), Scheerens (1992), Levine und Lezotte (1990), Sammons et al. (1995) und Cotton (1995). Diese früheren Review-Studien nennen die folgenden Bedingungen als zuträglich für hohe Schülerleistungen:

- Leistungsorientierung und hohe Erwartungen: ein produktives Schulklima, eine Schulphilosophie, die auf Leistung abzielt, gemeinsame Vision und Ziele, hohe Erwartungen in die Leistungsfähigkeit aller Schülerinnen und Schüler
- Kooperative Atmosphäre und ein geordnetes Klima: kooperative Planung, Übereinstimmung in Bezug auf lernorientierte Atmosphäre, geordnetes Klima
- Klare Zielsetzung in Bezug auf Grundfertigkeiten: Fokus auf Schülerlernen, Konzentration auf Unterricht

- Häufige Evaluation: angemessenes Monitoring, evaluatives Potenzial der Schule, Beurteilung
- Berufliche Weiterentwicklung: Personalentwicklung, berufliche Fortbildungen, eine lernende Organisation
- Einbeziehung der Eltern: Unterstützung der Eltern, Zusammenarbeit von Schule und Familie
- Schulleitung: Schulleitungshandeln, Schulmanagement und -organisation, entwicklungsorientierte Schulleitung
- Effektive Unterrichtsplanung: Klassenführung, *Time-on-Task*, strukturierter Unterricht, Lerngelegenheiten, Koordination von Lehrplan und Unterricht

Die folgenden Faktoren werden von den Autoren der früheren Review-Studien übereinstimmend als die wesentlichen Unterrichtsbedingungen eingestuft: Leistungsorientierung (die eng mit „hohen Erwartungen“ verbunden ist), Kooperation, Schulleitungshandeln, häufige Evaluation, Zeit, Lerngelegenheit und „Struktur“.

Neuere Reviews von Reynolds et al. (2014), Muijs et al. (2014) und Hopkins et al. (2014) bieten einen Überblick über die wichtigsten Voraussetzungen in der Schulwirksamkeitsforschung (*Educational Effectiveness Research* – EER) und der Unterrichtswirksamkeitsforschung (*Teaching Effectiveness Research* – TE). Das Review von Hopkins et al. (2014) fügt dieser Forschung eine Komponente der Verbesserung und Entwicklung hinzu (*School System Improvement* – SSI) und zielt darauf ab, unser Verständnis effektiver Interventions- oder Entwicklungsprogramme und der Voraussetzungen in Schulen zu fördern, die zu einer wirksamen Schulentwicklung beitragen. Die Zusammenfassung dieser Studien in Tabelle 1 (zitiert nach Scheerens 2014) verdeutlicht, dass es über die Jahre einen Konsens hinsichtlich der wesentlichen Bedingungen von Schule und Unterricht gibt. Die fünf Faktoren, zu denen in den früheren Reviews die größte Übereinstimmung auszumachen war, tauchen auch in diesen neueren Arbeiten auf. Die wichtigste Entwicklung zeigt sich im Review zu Unterrichtswirksamkeit in der Form von neu hinzugekommenen, vom Konstruktivismus inspirierten Unterrichtsstrategien. Sie sind in Tabelle 1 in der zweiten Spalte kursiv gesetzt.

Tabelle 1: Effektivitätsfördernde Bedingungen, Zusammenstellung aus den Review-Studien von Reynolds et al. (2014); Muijs et al. (2014) und Hopkins et al. (2014)

EER	TE	SSI
<ul style="list-style-type: none">• Effektive Führung• Akademischer Fokus• Positives, geordnetes Klima• Hohe Erwartungen• Beobachtung der Lernfortschritte• Einbeziehung der Eltern• Effektiver Unterricht (Zeit)• Mitarbeiterfortbildung• Schülerbeteiligung	<ul style="list-style-type: none">• Innerschulische Lerngelegenheiten (<i>Opportunity to learn</i>)• Zeit• Klassenmanagement• Struktur und unterstützende Lerngerüste, u. a. Feedback• Produktives Klassenklima• Klarheit der Präsentation• <i>Selbstgesteuertes Lernen fördern</i>• <i>Vermittlung meta-kognitiver Techniken</i>• <i>Vermittlung von Modellbildung</i>• <i>Differenziertere Diagnose</i>• <i>Bedeutung von Vorwissen</i>	<ul style="list-style-type: none">• Aspekte organisatorischer Gesundheit• Schulinterne Reflexion und Verbesserungsstrategien• Schulentwicklungsplan• Umfassende Schulreform• Facetten des Schulleitungshandelns (transformativ, pädagogisch, dezentralisiert)• Effektive systemische Reform, u. a. Schülerleistungen und Betonung auf Qualität des Unterrichts

EER – Schulwirksamkeitsforschung (*Educational Effectiveness Research*); TE – Unterrichtswirksamkeitsforschung (*Teaching Effectiveness Research*); SSI – Verbesserung und Entwicklung von Schule (*School System Improvement*). Quelle: Scheerens 2014

6. Weniger Konstanz der Effektgrößen in Synthesen quantitativer Forschung

Der vorangegangene Abschnitt zeigte eine deutliche Überschneidung bei wirksamkeitsfördernden Bedingungen sowohl in qualitativen als auch in quantitativen Reviews. Die quantitativen Reviews und Metaanalysen deuten jedoch stets auf bedeutsame Unterschiede in den spezifischen Effektgrößen aller Schlüsselvariablen hin. Diese Unterschiede sind in Tabelle 2 dargestellt (zitiert nach Scheerens 2013). Die Unterschiede zwischen den von Hattie (2009) berichteten Effektgrößen einerseits und den anderen, stärker auf Europa bezogenen Metaanalysen andererseits sind recht beeindruckend. Es sollte beachtet werden, dass Hattie Effektgrößen mithilfe des d-Koeffizienten ausdrückt und die anderen Autoren Korrelationen

darstellen. Da die Korrelationskoeffizienten halb so groß sind wie d-Koeffizienten, lassen sich die Werte nährungsweise ineinander überführen.

Tabelle 2: Ergebnisse aktueller Meta-Analysen

Schulbezogene Prozessmerkmale			
	Scheerens et al. 2007	Hattie 2009	Creemers & Kyriakides 2008
Konsens/Kohärenz Kollegium	.02	–	.16
Geordnetes Klima	.13	.34	.12
Monitoring und Evaluation	.06	.64	.18
Curriculum/Lerngelegenheiten	.15	–	.15
Hausaufgaben	.07	.30	–
Effektive Lernzeit	.15	.34	–
Einbeziehung der Eltern	.09	.50	–
Leistungsorientierung	.14	–	–
Schulleitungshandeln	.05	.36	.07
Differenzierung	.02	.18	–
Unterrichtsbezogene Prozessmerkmale			
	Scheerens et al. 2007	Hattie 2009	Seidel & Shavelson 2007
Zeit und Lerngelegenheiten	.08	.34	.03
Klassenführung	.10	.52	.00
Strukturierter Unterricht	.09	.60	.02
Vermittlung von Lernstrategien	.22	.70	.22
Feedback & Monitoring	.07	.66	.01

Koeffizienten basieren auf der Fisher-z-Transformation von Korrelationen; da Hattie Effektgröße d verwendet, erscheinen diese Angaben in Fettschrift. Quelle: Scheerens 2013, S.14

Da wir auf einen Vergleich zwischen Inspektionskonzepten und Schulwirksamkeitsforschung abzielen, fügen wir die nachfolgende Tabelle mit einer durchschnittlichen Effektgröße von wesentlichen Wirksamkeitsbedingungen an, wie sie in Metaanalysen von Marzano (2003), Scheerens et al. (2007) und Hattie (2009) beschrieben werden. Eine Mittlung der Effektgrößen stellt zwar eine etwas grobe Methode dar, vermittelt aber einen Eindruck von der relativen Bedeutung dieser entschiedenen wirksamkeitsfördernden Bedingungen. Die durchschnittlichen Effektgrößen zeigen an, dass die wichtigste Bedingung des schulischen Beitrags zu guten Schülerleistungen darin besteht, die Schülerinnen und Schüler bildungsbezogenen Inhalten (Lerngelegenheiten und Unterrichtszeit) „auszusetzen“. Organisationale

Faktoren wie Schulleitungshandeln und Kooperation weisen dagegen relativ kleine Effektgrößen auf.

Tabelle 3: Rangfolge von Schulwirksamkeitsvariablen nach durchschnittlicher Effektgröße (d-Koeffizient), Zusammenstellung aus drei Meta-Analysen

	Marzano 2003	Scheerens et al. 2007	Hattie 2009	Durch- schnittliche Effektgröße
Lerngelegenheiten	.88	.30	.39*	.523
Unterrichtszeit	.39	.30	.38	.357
Monitoring	.30	.12	.64	.353
Leistungserwartungen	.27	.28	.43**	.327
Einbeziehung der Eltern	.26	.18	.50	.313
Schulklima	.22	.26	.34	.273
Schulleitungshandeln	.10	.10	.36	.187
Kooperation	.06	.04	.18***	.093

*) operationaisiert als „Zusatzprogramm für begabte Kinder“; **) operationalisiert als „Lehrererwartungen“; ***) operationalisiert als „Teamteaching“

Quelle: Scheerens 2013, S. 24

Ein weiteres wichtiges Thema, das es zu berücksichtigen gilt, ist die Konstanz dieser Effekte durch alle Einzelschulen. Konstanz in der Schätzung von Schuleffekten durch Klassenstufen, Lehrerinnen und Lehrer und Unterrichtsfächer sowie Stabilität von Schuleffekten durch Jahre stellen bei Schulinspektionen eine wichtige Grundannahme dar. Die Beurteilung der Schulqualität durch Schulinspektionen findet normalerweise einmal in drei bis fünf Jahren statt und man geht davon aus, dass die Schul- und Unterrichtsqualität bis zum nächsten Inspektionsbesuch relativ unverändert bleibt.

Mehrere Studien zur Schulwirksamkeit stellen die potenzielle Beständigkeit der Schulwirksamkeit (oder den Mangel derselben) mithilfe einer Analyse einer Korrelationsmatrix zu fächer- und kohortenbezogenen (oder klassenbezogenen) Effekten dar und berechnen die Stärke eines allgemeinen Schulfaktors (vgl. Scheerens 2013). Dazu seien Vermeer und Van der Steeg (2011) zitiert:

„Typischerweise ist die Rangordnung der durchschnittlichen Leistung (*Value Added*) von Schulen durch die Jahre korreliert. In einer Studie zu weiterführenden Schulen in den Niederlanden fanden Bosker et al. (1989) Korrelationen, die abhängig von der Zeitspanne von ein bis vier Jahren abnahmen: .74 (ein Jahr), .62 (zwei Jahre), .49 (drei Jahre) und .49 (vier Jahre). Gray et al. (1995) betrachteten bei weiterführenden Schulen in England Zeitabstände von einem, zwei und drei Jahren und

fanden Korrelationen von .94, .96 und .81. Thomas et al. (2010) analysierten in der Grafschaft Lancashire Schuldaten über einen Zeitraum von elf Jahren hinweg. Sie kamen zu dem Schluss, dass die Schuleffekte vergleichsweise stabil waren. Wenn Schulen jedoch als durchschnittlich, über- oder unterdurchschnittlich kategorisiert wurden, ließen sich in den einzelnen Kategorien viele Veränderungen feststellen. Im Verlauf von elf Jahren waren 50 Prozent der Schulen in eine andere Kategorie gewechselt. Außerdem waren kaum beständige Fortschritte zu beobachten:

Bei den meisten Schulen scheint der Aufwärtstrend typischerweise nicht länger als drei Jahre anzudauern. Kurz gesagt erbringt die nichtlineare Modellierung den Nachweis, dass es zwar zweifelsohne Veränderungen gab, dass diese jedoch nicht ‚fortdauernd‘ waren und in vielen Fällen auch zufällig aufgetreten sein könnten. Diese Ergebnisse stehen in krassem Widerspruch zu Wunschkonzeptionen der Regierung von einer stetig fortschreitenden Schulentwicklung (Thomas et al. 2010, S. 280). Auch eine neuere niederländische Studie weist eine geringere Stabilität nach: Von den weiterführenden Schulen mit den höchsten Punktwerten waren nach drei Jahren nur noch 15 Prozent in der obersten Kategorie.“ (Vermeer & Van der Steeg 2011, S. 9 f.; Übersetzung: R. K.)

Um der mangelnden Stabilität Rechnung zu tragen, wäre es sinnvoll, die Position von Schulen über eine bestimmte Zeitspanne (z. B. drei Jahre lang) im Rahmen von Rechenschafts- und Belohnungsprogrammen zu begutachten und Schulen anhand ihres Leistungsdurchschnitts aus einer Reihe von Jahren zu vergleichen. Die Forschungsergebnisse aus verschiedenen Ländern zeigen diesbezüglich wiederum, dass die Anzahl der Jahre der Durchschnittsberechnung in jedem Land unterschiedlich sein kann, da die Stabilität von Schuleffekten von Land zu Land verschieden zu sein scheint und eventuell die Homogenität von Bildungssystemen widerspiegelt.

7. Was mit den Schlüsselfaktoren schulischer Wirksamkeit gemeint ist

Um ein wenig genauer hinter die „Etiketten“ der Faktoren zu sehen, die in den Tabellen der vorangegangenen Abschnitten genannt wurden, folgt nun eine Beschreibung der wesentlichen Merkmale und der Subkomponenten dieser Faktoren.

Leistungsorientierung

Dieser Faktor bezeichnet ergebnisorientiertes Streben und eine positive, optimistische Einstellung zur Leistungsfähigkeit aller Schülerinnen und Schüler. Zu den Datenquellen gehören Planungsunterlagen wie Schulentwicklungspläne oder Leit-

bilder, Antworten von Schulleitung und Lehrerinnen und Lehrern zu Fragebögen und Nachweise auf Verwaltungsebene, z. B. Aufzeichnungen zu Schülerleistungen. Leistungsorientierung wird oft in Subkomponenten unterteilt, z. B. „klarer Fokus auf Beherrschung der Kernfächer“, „hohe Erwartungen auf Schul- und Lehrerebene“ und „Aufzeichnungen zu Schülerleistungen“.

Schulleitungshandeln

Bei vielen operationalen Definitionen und Instrumenten im Zusammenhang mit Schulleitungshandeln stellt man einen deutlichen Fokus auf Führungsrollen im Bereich des Hauptprozesses des Lernens und Unterrichtens und im Bereich organisatorischer Voraussetzungen, die als Unterstützung für diesen Hauptprozesses angesehen werden, fest. Dazu gehören das Coaching von Lehrerinnen und Lehrern und Hilfestellung bei curricularen Entscheidungen. Oft wird auch ein Zusammenhang zum Monitoring von Einschätzungen und Fortschritten von Schülerinnen und Schülern hergestellt. Neben dem Schulleitungshandeln, das sich auf den primären Prozess des Unterrichtens und Lernens konzentriert, ist „transformationales Schulleitungshandeln“ eher auf Aktivitäten gerichtet, die die Schule auf organisatorischer Ebene verbessern. „Dezentralisiertes Schulleitungshandeln“ oder gar „*Teacher Leadership*“ bedeutet, dass Teile der Schulleitung an Lehrerinnen und Lehrer abgegeben werden. Relevante Subkomponenten von Schulleitungshandeln sind: allgemeiner Führungsstil, Führungsrollen im Zusammenhang mit Koordination (partizipative Entscheidungsprozesse und Informationsvermittlung steuern), Meta-Kontrolle von Unterrichtsprozessen, Förderung von Personalfortbildung.

Kooperation, Konsens und Kohärenz im Kollegium

Früher gab es eine leichte Betonung auf messbaren Facetten von Kooperation (wie Häufigkeit von Konferenzen) und persönlicher Zufriedenheit. Neuerdings setzt sich eher der Gedanke von Schulen als professionellen Lerngemeinschaften und *Peer Learning* durch, sodass Kooperation stärker im Zusammenhang mit schulbezogenen Entwicklungsinitiativen einerseits und Diskussionen zu Unterricht und Lernen andererseits gesehen wird. „*Teamteaching*“ gilt auch oft als Indikator für Kooperation unter Kollegen.

Kooperation, Konsens und Kohärenz im Kollegium wird normalerweise durch die Art und Häufigkeit von Zusammenkünften und Besprechungen, Zufriedenheit mit Kooperation, aufgabenbezogene Facetten von Kooperation, Konstanz in Bezug auf Unterrichtsziele und -methoden gemessen.

Curriculare Qualität und Lerngelegenheiten

Curriculare Qualität wird meist im Sinne systematischer Prozessplanung und erlebter Zufriedenheit mit dem Curriculum gemessen. Das Konzept der Lerngelegenheiten steht im Zusammenhang mit der Anpassung von schulischen Zielen, Unterricht und Schülerbeurteilung. Grundsätzlich geht es um die Übereinstimmung der unterrichteten Inhalte mit den getesteten Inhalten. In jüngeren Studien wird „Testvorbereitung“ als neue Sichtweise auf Lerngelegenheiten betrachtet. Bezogen auf den Unterricht ist „unterrichtliche Angleichung“ eine weitere neue Interpretation von Lerngelegenheiten. Im Laufe der Zeit wurden curriculare Qualität und Lerngelegenheiten in folgende Subkomponenten operationalisiert: systematisches Formulieren und Überwachen curricularer Prioritäten, Auswahl und Anwendung von Methoden und Lehrbüchern und Lerngelegenheiten im Sinne von behandelten Inhalten.

Schulklima

Ein geordnetes und sicheres Schulklima ist eine der organisatorischen Voraussetzungen von Schulen, die in internationalen Schulleistungsuntersuchungen wie PISA relativ positive Unterstützung erfahren haben. Außerdem hat die Schulwirksamkeitsforschung diesen Punkt von Anfang an im Blick. Offensichtlich ist die leistungsbezogene Facette des Schulklimas eng verbunden mit „Leistungsorientierung“ und „leistungsorientierter Schulphilosophie“, wie sie oben beschrieben wurden. Interne Beziehungen als Teil des „beziehungsorientierten Schulklimas“ betreffen die Beziehungen zwischen Lehrerinnen/Lehrern und Schülerinnen/Schülern, Lehrerinnen/Lehrern und Schulleitung sowie Lehrerinnen/Lehrern und Schülerinnen/Schülern jeweils untereinander. Relevante Subkomponenten des Schulklimas sind Disziplin, Leistungsorientierung und gute interne Beziehungen der Mitarbeiter einer Schule untereinander sowie zwischen Mitarbeitern und Schülerinnen und Schülern.

Evaluation und Monitoring

Evaluation als eine wirksamkeitsfördernde Bedingung hat mit der Evaluationsorientierung (oder dem Mangel einer solchen Orientierung) an einer Schule zu tun, sowohl auf den Unterricht als auch auf die Schülerinnen und Schüler bezogen. Die Häufigkeit der Anwendung wird ebenso gemessen wie die Zufriedenheit der Mitarbeiter mit Evaluationen und die Art und Weise, wie Evaluationsergebnisse im Sinne einer Schulentwicklung eingesetzt werden. Zur Messung von Eva-

uation und Monitoring wird zunächst kontrolliert, ob eine Schule regelmäßig systematische Selbstevaluationen und Schülerbewertungen vornimmt und/oder andere Arten von Tests und regelmäßige Überprüfungen des Leistungsstandes der Schülerinnen und Schüler einsetzt. Intensivere Messungen von Evaluation und Monitoring beinhalten eine Analyse einer aufgabenbezogenen Zusammenarbeit der Lehrerinnen und Lehrer und gehen der Frage nach, ob (formell standardisierte und informelle Formen von) Lehrer- und Unterrichtsevaluationen einen festen Platz im *Peer Learning* haben. Subkomponenten von Evaluation und Monitoring sind Schulevaluation, Klassenevaluation und Überprüfung von Schülerleistungen.

Einbeziehung der Eltern

Wesentliche Bestandteile der Einbeziehung der Eltern sind das Mitspracherecht bei der Formulierung von Schulprinzipien, die aktive Beteiligung der Eltern an schulischen Angelegenheiten, die Unterstützung durch die Eltern entweder in Form praktischer Hilfestellungen oder im Zusammenhang mit Unterricht und Lernen. Die ambitionierteste Form der Zusammenarbeit mit den Eltern ist sicherlich der Versuch der Schule, das pädagogische Klima in der Familie zu beeinflussen. Die Einbeziehung der Eltern wird oft dadurch gemessen, dass Schulen befragt werden, welchen Stellenwert sie der Beteiligung von Eltern in ihrer Schulphilosophie einräumen, wie intensiv der Kontakt zu Eltern ist und wie zufrieden die entsprechenden Akteure (Lehrerinnen und Lehrer, Eltern und Schulleitung) mit der Einbeziehung der Eltern sind. Das Konzept der Einbeziehung der Eltern lässt sich zu einer „Einbeziehung des schulischen Umfelds“ erweitern.

Klassenklima

Das Klassenklima deckt sich in weiten Teilen mit der Beschreibung des Schulklimas, vor allem in Bezug auf die Subkomponenten. Beide Definitionen beinhalten einen disziplinarischen Teil und heben gute Beziehungen hervor, das Klassenklima konzentriert sich jedoch auf die Beziehung zwischen Lehrerinnen und Lehrern und Schülerinnen und Schülern, während das Schulklima auch die Beziehungen des Schulpersonals untereinander einbezieht. Zum Klassenklima gehört darüber hinaus die Idee von kognitiver und emotionaler Unterstützung von Schülerinnen und Schülern, in der Form von klaren und verständlichen Erklärungen, Hilfe bei Aufgaben sowie einer anregenden Einbeziehung und einem Gefühl der Selbstwirksamkeit der Schülerinnen und Schüler. Bisweilen wird beim Klassenklima auch ein „Spaßfaktor“ berücksichtigt: Die Schülerinnen und Schüler werden gefragt,

ob sie ihren Lehrer oder ihre Lehrerin mögen, ob er oder sie mit ihnen auch über außerschulische Aktivitäten spricht und ob sie auch Witze machen und lachen.

Effektive Lernzeit

Lernzeit kann ganzheitlich oder detaillierter gemessen werden. Eine differenzierte Messung unterscheidet zwischen zugewiesener Lernzeit (offiziell anberaumte Unterrichtsstunden), reiner Unterrichtszeit (der Teil der Stunde, den der Lehrer tatsächlich auf das Unterrichten verwendet, wobei die Zeit für Organisatorisches und für Ablenkungen abgezogen wird) und *Time-on-Task* (die Zeit, in der ein Schüler oder eine Schülerin aktiv beschäftigt ist). Klassenführung wird oft im Sinne einer Maximierung der reinen Unterrichtszeit und der *Time-on-Task* definiert. Eine weitere wichtige Unterscheidung gibt es zwischen der Unterrichtszeit in der Schule und der Zeit, die Schülerinnen und Schüler mit Hausaufgaben zubringen. Studien zur Unterrichtszeit beinhalten oft Informationen zu Abwesenheitszeiten von Schülern und Unterrichtsausfällen.

Strukturierter Unterricht

Strukturierter Unterricht steht im Zusammenhang mit dem Element der kognitiven Unterstützung im Rahmen des Klassenklimas. Dazu gehört die Frage nach verständlichen Erklärungen des Lehrers oder der Lehrerin und der Hilfestellung, die Schülerinnen und Schüler bei Aufgaben bekommen. Bei strukturiertem Unterricht geht es grundsätzlich um häufige Interventionen zur Förderung des Lernprozesses, z. B. klare Benennung der Lernziele, Aufteilung des gesamten Lernstoffes in relativ kleine Einheiten, Präsentation dieser Einheiten in einer gut geplanten Abfolge, zahlreiche Übungsmöglichkeiten für die Schülerinnen und Schüler, Hilfestellung durch Hinweise, häufige Fragen und Überprüfungen, um den Lernfortschritt zu kontrollieren, und Feedback an die Schülerinnen und Schüler. Relevante Subkomponenten sind klare Zielsetzungen, Vorbereitung strukturierter Sequenzen von Unterrichts- und Lerneinheiten, verständliche Erklärungen, Einsatz von Fragen und Feedback als Teil der Wissensvermittlung und die Überprüfung der Schülerfortschritte.

Konstruktivistischer Unterricht und selbstbestimmtes Lernen

Konstruktivistischer Unterricht und selbstbestimmtes Lernen scheinen im Gegensatz zu den eher behavioristischen und gelenkten praktischen Ansätzen des strukturierten Unterrichts zu stehen. Konstruktivistischer Unterricht und selbstbestimmtes

Lernen unterstreichen den Lernprozess von Schülern und die Vermittlung allgemeiner und/oder fächergebundener Lernstrategien. „Kognitive Aktivierung“ ist ein wichtiges Element des konstruktivistischen Unterrichts. Dabei geht es um eine ausreichend tief gehende Präsentation des Inhalts, die auf ein Verstehen auf einer höheren Ebene abzielt, auf der Schülerinnen und Schüler verstehen, authentische Einsatzmöglichkeiten des Gelernten und Konzepte in unterschiedlichen Kontexten anwenden können. Die Idee vom „*Scaffolding*“, von unterstützenden Lerngerüsten, geht davon aus, dass das Maß an Selbstregulierung des Schülerlernens allmählich vergrößert wird, je besser die Schülerinnen und Schüler den Lernstoff beherrschen. Sie kombiniert die Konzepte von strukturiertem und konstruktivistischem Unterricht und stellt beide Ansätze in ein Kontinuum.

Differenzierung

Differenzierung erkennt die individuellen Unterschiede der Schülerinnen und Schüler an und versucht, Raum für Variationen im Unterricht zu schaffen, die diesen Unterschieden gerecht werden. Schulen und Lehrerinnen und Lehrer können den Unterricht dadurch differenziert gestalten, dass sie Schülerinnen und Schüler in Klassen kanalisieren, die auf unterschiedlichen Fähigkeitsniveaus arbeiten, Schülerinnen und Schüler innerhalb einer Klasse in Gruppen mit unterschiedlichen Fähigkeitsniveaus zusammenfassen, das Unterrichtstempo anpassen (und Schülern so die Möglichkeit geben, den Lernstoff nach unterschiedlichen Zeitplänen zu bearbeiten) und den Unterricht innerhalb relativ heterogener Klassenbesetzungen individualisieren. Zur Differenzierung gehören auch spezielle Programme und/oder zusätzlicher Unterricht und weitergehende Hilfen für schwache Schülerinnen und Schüler sowie Zusatzaufgaben für leistungsstarke Schülerinnen und Schüler.

8. Der Facettenreichtum schulwirksamkeitsfördernder Variablen

Es ist wichtig anzumerken, dass diese Konzepte sich nicht gegenseitig ausschließen und dass es mehrere Beispiele konzeptioneller Überschneidungen gibt. Leistungsorientierung wird nicht nur im Sinne direkter Schulphilosophien beschrieben, sondern auch als eine relevante Ausrichtung im Klima einer Schule. Außerdem ist es wichtig zu sehen, dass Evaluationen und die Überprüfung von Leistungen wesentliche Elemente der Leistungsorientierung einer Schule sind. Leistungsorientierung wirkt sich ebenso wie Schulleitungshandeln auf die Ergebnisse von Schülerleistungen aus. Abgesehen von der konzeptionellen Überschneidung dieser

Hauptindikatoren passiert es auch, dass Schulen eine Reihe dieser Indikatoren zu einer Gesamtstrategie in Bezug auf (eine Verbesserung von) Lehren und Lernen so kombinieren. Lerngelegenheiten und *Time-on-Task* sind z. B. wichtige Bestandteile einer umfassenden Strategie zur gesteigerten Konfrontation der Schülerinnen und Schüler mit Inhalten. „*Focused Teaching*“ (ein von Louis et al. 2010 geprägter Begriff) verbindet dagegen direkte Wissensvermittlung und konstruktivistische Unterrichtsstrategien. Ein drittes Beispiel ist die Art, wie die inhärente Verwandtschaft von Evaluation und Monitoring Lerngelegenheiten garantiert, was in Pophams Credo „*Test Better; Teach Better*“ (etwa: bessere Tests, besserer Unterricht; Popham 2003) seinen Niederschlag findet. Und schließlich kann man Evaluation und die Überprüfung von Leistungen als wirksame Hebel bei strukturierteren Unterrichtsansätzen betrachten, um Lernziele zu formulieren und den Schülerinnen und Schülern strukturierte Unterstützung und Anleitung zu geben. (Dieser Ansatz wird gerade in den Niederlanden unter der Überschrift „ergebnisorientierte Arbeit“ probeweise angewandt; vgl. Visscher & Ehren 2011.)

Viele der Faktoren wie Klima, Leistungsorientierung, Evaluation und Monitoring und Lerngelegenheiten, die oben erörtert wurden, lassen sich sowohl klassen- als auch schulbezogen messen und können auf beiden Ebenen unterschiedlich interpretiert werden. So mag ein Lehrer oder eine Lehrerin nur eingeschränkt von der Möglichkeit Gebrauch machen, die Überprüfung der Schülerleistungen in seinen Unterricht einfließen zu lassen, während der Schulleiter in der ganzen Schule ausgefeilte Systeme zur Kontrolle der Lernfortschritte der Schülerinnen und Schüler einsetzt, um schulbezogene Entwicklungsstrategien danach auszurichten.

Der Facettenreichtum der Schulwirksamkeit, die konzeptionellen Überschneidungen der Bedingungen und die zahlreichen Möglichkeiten, jede Subkomponente zu beschreiben und zu messen, machen deutlich, wie komplex die Gestaltung effektiver Inspektionskonzepte ist. Die Frage, welche Indikatoren Inspektionskonzepte enthalten sollen und wie sie zu messen sind, bedarf gründlicher Überlegungen. Im folgenden Abschnitt werden solche Überlegungen in Bezug auf sechs Inspektionskonzepte angestellt. Sie stammen aus den Niederlanden, England, Irland, Schweden, dem österreichischen Bundesland Steiermark und der Tschechischen Republik. Sie wurden von Ehren et al. (2013) im Rahmen eines von der EU finanzierten Projekts untersucht. Mithilfe der Dissertation von De Volder (2012) und der Länderprofile auf der Website des SICI, der *Standing International Conference of national and regional Inspectorates of education*, wurde die Tabelle vervollständigt. Die Beschreibung der Inspektionskonzepte bezieht sich auf das Jahr 2010.

9. Sechs Inspektionskonzepte: eine Analyse

Die folgende Übersicht bietet einen Überblick über die Inspektionskonzepte aus sechs europäischen Ländern. Die Länder repräsentieren eine große Bandbreite an Inspektoratstypen. Einige arbeiten nach einem *Low-Stakes*-Inspektionsansatz mit Betonung auf der Entwicklungskapazität einer Schule (z. B. Irland), andere überprüfen Schulen anhand testbasierter Frühwarn-Analysen (z. B. die Niederlande). Einige Länder haben stark zentralisierte nationale Inspektorate (z. B. England), andere Inspektorate arbeiten auf Regionalebene (z. B. in Österreich). In der Übersicht fassen wir die Inspektionsstandards und Schwellenwerte zusammen, nach denen die Inspektorate Schulen evaluieren. Darüber hinaus erläutern wir kurz den hierarchischen Aufbau dieser konzeptionellen Rahmen und die Art der Datensammlung, auf die sich die Beurteilung durch die Inspektorate stützt. Es ist wichtig anzumerken, dass sich hinter einer sichtbaren Überschneidung der Standards von Inspektionskonzepten sehr wahrscheinlich eine breitgefächerte Gamma-Verteilung operationaler Definitionen und spezifischer Messinstrumente verbirgt. Die Zusammenfassung und der nachfolgende Vergleich mit der Schulwirksamkeitsforschung muss sich daher auf allgemein gefasste Kommentare und Vorschläge beschränken, wie sich Inspektionskonzepte möglicherweise verbessern ließen.

Übersicht: Inspektionsstandards bei regelmäßig stattfindenden Gesamtinspektionen in sechs europäischen Ländern (Stand 2010)

Niederlande

Input

N/A

Prozess

- Die Schule verfügt über ein System, mit dem die Qualität ihres Unterrichts sichergestellt wird.
- Der den Schülerinnen und Schülern angebotene Lernstoff stellt eine ausreichende Vorbereitung auf die berufliche/akademische Bildung nach der weiterführenden Schule dar.
- Den Schülerinnen und Schülern steht ausreichend Unterrichtszeit zur Verfügung, um den Lernstoff zu bewältigen.
- Die Schule nimmt eine systematische Bewertung der Fortschritte der Schülerinnen und Schüler vor.
- Das Schulklima ist sicher und anregend.

- Das pädagogische Vorgehen der Lehrerinnen und Lehrer entspricht den grundlegenden Erfordernissen.
- Das didaktische Vorgehen der Lehrerinnen und Lehrer entspricht den grundlegenden Erfordernissen.
- Kinder mit sonderpädagogischem Förderbedarf bekommen die nötige Unterstützung.

Output

- Die Ergebnisse der Schülerinnen und Schüler entsprechen dem erwarteten Niveau (unter Berücksichtigung der Merkmale der Schülerpopulation).

Schwellen für „Failing Schools“

- Gesamtbeurteilung als sehr schwach, schwach, ausreichend oder gut.
- Eine Schule ist schwach, wenn die Leistungsergebnisse der Schülerinnen und Schüler am Ende der Grundschulzeit in mehr als drei aufeinanderfolgenden Jahren unzureichend ist.
- Eine Schule wird als sehr schwach beurteilt, wenn die Ergebnisse nach Abschluss der Grundschule in mehr als drei aufeinanderfolgenden Jahren unzureichend ist und das Lehren und Lernen oder die Betreuung der Schülerinnen und Schüler unzureichend ist.

Datensammlung und -auswertung

- Das Konzept beinhaltet die folgenden hierarchischen Ebenen: 5 Qualitätsbereiche, 10 Qualitätsaspekte, 46 Indikatoren (von denen 10 zum Bereich der Schwellen gehören).
- Es gibt Entscheidungsregeln für die Bewertung jedes einzelnen Indikators als ausreichend oder nicht ausreichend, außerdem Entscheidungsregeln für die Aggregation der Bewertungen der Indikatoren, um zu einer Beurteilung der Qualitätsaspekte, der Qualitätsbereiche und der Gesamtbewertung der Effektivität einer Schule zu gelangen.
- Datensammlung bei Inspektionsbesuchen im Rahmen von Unterrichtsbeobachtungen (mithilfe eines detaillierten Bewertungsbogens), leitfadengestützten Interviews mit Mitarbeitern der Schule, Eltern und Schülerinnen und Schülern und Dokumentenanalysen, u. a. Analyse der Selbstevaluation der Schule.
- Der Output der Schule wird mithilfe der Analyse landesweit standardisierter Tests und Abschlussprüfungen sowie anderer Indikatoren ermittelt (z. B. Abbrecher- und Abschlussquoten).

England

Input

Bericht zum Gebäudezustand, falls er Auswirkungen auf die Unterrichtsqualität oder Gesundheit und Sicherheit hat.

Prozess

- Qualität des schulischen Angebots, u. a. Unterricht, Lehrplan und Betreuung, Anleitung und Hilfestellung für Schülerinnen und Schüler
- Führung und Management
- Allgemeine Effektivität der Schule
- Fähigkeit der Schule zur Verbesserung

Output

- Leistung und Standards (Lernfortschritte der Schülerinnen und Schüler und in nationalen Tests erzielte Standards unter Berücksichtigung des Kontexts)
- Persönlichkeitsentwicklung und Wohlbefinden der Schülerinnen und Schüler

Schwellen für „Failing Schools“

- Schulen werden als *outstanding* (hervorragend), *good* (gut), *satisfactory* (zufriedenstellend) oder *inadequate* (ungenügend) eingestuft, wobei es bei den ungenügenden Schulen eine zusätzliche Unterteilung in „*Notice to Improve*“ (Aufforderung zur Verbesserung) und „*Special Measures*“ (besondere Maßnahmen) gibt. Bei etwa 40 Prozent der zufriedenstellenden Schulen wird der Bedarf beobachtender Inspektionen festgestellt.

Datensammlung und -auswertung

- Konzept beinhaltet 6 Hauptaspekte, unterteilt in 7 Kriterien zu Schülerergebnissen, 3 Kriterien zur Effektivität des schulischen Angebots und 8 Kriterien zur Effektivität von Führung und Management der Schulen.
- Datensammlung u. a. durch Unterrichtsbeobachtung (standardisierter Evaluationsbogen), leitfadengestützte Interviews mit Mitarbeitern der Schule, Eltern und Schülerinnen/Schülern und Dokumentenanalyse. Dazu gehört die Analyse der Selbstevaluation der Schule und die Auswertung von Fragebögen, die Eltern im Vorfeld des Inspektionsbesuchs ausfüllen.
- Messung des Outputs der Schule durch Analyse von Leistungsdaten: Lernfortschritte der Schülerinnen und Schüler, erreichte Standards in landesweiten Tests und Prüfungen unter Berücksichtigung von Kontextfaktoren (Anteil sozial benachteiligter Schülerinnen/Schüler, Anteil der Schülerinnen/Schüler, deren Muttersprache nicht Englisch ist), Analyse schriftlicher Schülerarbeiten und durch die Schule zur Verfügung gestellter Leistungsdaten.

Schweden

Input

N/A

Prozess

- Der Fokus liegt auf der Einhaltung gesetzlicher Vorschriften, mit Inspektionen soll sichergestellt werden, dass die Rechte der Schülerinnen und Schüler in Bezug auf Bildung berücksichtigt werden.
- Ergebnisse, Leistungsstandards, Lernen und Lehren: wie Schulen auf eine Verbesserung der Leistungsstandards hinarbeiten, wie Lehrerinnen und Lehrer ihren Unterricht auf die individuellen Bedürfnisse der Schülerinnen und Schüler ausrichten, ob und wie sie Erkenntnisse aus den Schülerergebnissen in ihrem Unterricht umsetzen, wie und ob sie Kinder mit speziellem Förderbedarf unterstützen.
- Den Schülerinnen und Schülern die Regeln und Werte einer demokratischen Gesellschaft nahebringen.
- Management und internes Audit-Verfahren: Schulverwaltung und -management.

Output

- Ergebnisse, Leistungsstandards, Lernen und Lehren: was Schülerinnen und Schüler mit Blick auf die im nationalen Curriculum formulierten Ziele lernen.

Schwellen für „Failing Schools“

- Leitfaden für Beurteilungen wird erstellt, zur Zeit ganzheitlicher Ansatz, bei dem man sich bei der Einschätzung, ob sich eine Schule an die gesetzlichen Vorgaben hält, auf Erfahrung und Kompetenz des Inspektors verlässt.

Datensammlung und -auswertung

- Konzept beinhaltet 3 Hauptbereiche und 16 Indikatoren.
- Datensammlung u. a. durch Unterrichtsbeobachtung (Anmerkungen zu festgelegten Punkten auf einem Beurteilungsbogen, keine detaillierte Punktwertung), Befragung des Direktoriums, Dokumentenanalyse (Statistiken, Qualitätsberichte, vorangegangene Kontrollbeschlüsse und Inspektionsberichte, Selbstevaluationen), Interviews mit Schulpersonal, u. a. Schulleitung, Lehrerinnen/Lehrern, Schulkrankenschwestern und Schülerinnen/Schülern, Eltern und zuständigen Politikern.

Irland

Input

- Informationen zur Ressourcenverteilung für Schülerinnen und Schüler mit besonderem Förderbedarf, außerdem Informationen zu Bauprogrammen und Plänen zur Verbesserung oder Erweiterung der Einrichtungen (keine separate Überprüfung).

Prozess

Auf Grundschulniveau:

- Qualität des Schulmanagements
- Qualität der Schulplanung
- Qualität von Unterricht und Lernen
- Qualität der Unterstützung der Schülerinnen und Schüler

Output

- Keine Output-Standards. Testergebnisse werden während des Inspektionsbesuchs analysiert, jedoch nicht in den Bericht aufgenommen. Das Analyseergebnis von Testnoten kann mit der Schulleitung und/oder Lehrerinnen und Lehrern besprochen werden und trägt zum Urteil der Inspektoren über die Unterrichts- und Lernqualität bei. Statistische Daten fließen jedoch nicht in den Inspektionsbericht ein.

Schwellen für „Failing Schools“

- Keine Schwellen. Gegenwärtig ist die Evaluation nicht Grundlage für eine Gesamtwertung. Empfehlungen für künftige Maßnahmen des Schulmanagements oder Fachlehrerinnen und Fachlehrer werden formuliert.

Datensammlung und –auswertung

- Konzept umfasst 5 Evaluationsbereiche, unterteilt in 143 „Themen der Selbstevaluation“. Theoretisch müssen Schulen Nachweismaterial sammeln und sich dann selbst mit ihrer Leistung zu jedem Thema auf einer vierteiligen Skala einordnen. Dieser Selbstevaluationsprozess bildet die Grundlage eines Inspektionsteams, das Gesamtevaluationen von Schulen (*Whole School Evaluations*) durchführt.
- Datensammlung mittels Dokumentenanalyse (Formular mit Schulinformationen, von Schulleitung ausgefüllt, Dokumentation der Schule zu Schüleranmeldungen, Vorgaben des Lehrplans, Planungen, Schulrichtlinien usw.), Fragebögen an Schulleitung und Fachbereichskoordinatoren, Interviews mit Mitarbeitern der Schulen, Unterrichtsbeobachtungen.

- Output-Daten werden gesammelt, auf Grundschulniveau vergleichen die Inspektorinnen und Inspektoren die Schülerleistungen in standardisierten Tests in Mathematik und Englisch mit nationalen Normen.
- Auf Grundschulniveau analysieren die Inspektorinnen und Inspektoren laufende Tests, die Lehrerinnen und Lehrer in allen Fächern einsetzen, sowohl für formative als auch für summative Zwecke, außerdem standardisierte, von einer externen Forschungsagentur entwickelte Tests in Englisch und Mathematik, die die Lehrerinnen und Lehrer der Schule kaufen und stellen.

Österreich (Bundesland Steiermark)

Input

- N/A

Prozess

- Lehren und Lernen
- Klassenzimmer und Schulen als Lebensraum
- Schulpartnerschaft und Außenbeziehungen
- Schulmanagement
- Professionalität und Personalentwicklung

Output

- N/A
- Es gibt standardbasierte Leitungstests, doch Inspektorate dürfen sie nicht einsetzen oder in die Output-Messungen einbeziehen.

Schwellen für „Failing Schools“

- Keine Schwellen, ein Überblick über Stärken und Schwächen wird formuliert. Über Konsequenzen und notwendige Maßnahmen werden die Schule und die zuständigen Behörden informiert.

Datensammlung und -auswertung

- Konzept umfasst 5 Bereiche der Schulqualität, unterteilt in 16 Kriterien, mit einem Fokus auf einer kritischen Analyse der Selbstevaluationsprozedur und -ergebnisse der Schulen, wie sie im Schulprogramm formuliert ist. Nur wenn diese Dokumentenanalyse („methodisch-substantive Analyse der Selbstevaluation“) Mängel aufweist (in Bezug auf die Prozeduren der Selbstevaluation) oder aufzeigt, dass „wesentliche Anforderungen und Standards“ (bezieht sich auf die Ergebnisse der Selbstevaluation) nicht eingehalten werden, muss eine

„spezifische Untersuchung“ durchgeführt werden. Einer solchen Untersuchung gehen Beratungen und Hilfsmaßnahmen voraus.

- Datensammlung: Dokumentenanalyse des „Inspektionsordners“ (von der Schule zusammengestellt), in dem alle Schulentwicklungspläne der letzten 5 Jahre enthalten sind, außerdem das Schulprogramm für die individuelle Förderung aller Schülerinnen und Schüler, eine Liste aller Lehrer (mit Altersangabe), Stundenpläne, Liste der Schülerinnen und Schüler, Unterrichtsbeobachtungen (strukturiertes Protokoll, das auch den Lehrerinnen und Lehrern zur Verfügung steht), Interviews mit Lehrerinnen/Lehrern, Schülerinnen/Schülern, Eltern, Schulleitung und Bürgermeister (mithilfe eines strukturierten Leitfadens).

Tschechische Republik

Input

- Bei einer „Zustandsüberprüfung“ kontrollieren die Prüfer, ob gesetzliche Bestimmungen zur Bereitstellung schulischer Angebote erfüllt sind (zusätzlich zur Prüfung der Gebäude).
- Arbeitsbedingungen Personal
- Materielle Voraussetzungen
- Finanzielle Voraussetzungen

Prozess

- Gleiche Bildungschancen
- Bildungsprogramme der Schulen
- Schulmanagement
- Effektive Schulorganisation
- Effektive Unterstützung der Persönlichkeitsentwicklung der Kinder
- Partnerschaft
- Effektive Unterstützung der Entwicklung von Schlüsselkompetenzen von Kindern, Schülern und Studenten

Output

- Systematische Evaluation einzelner Schülerinnen/Schüler- und Gruppenergebnisse
- Evaluation von Gesamtergebnissen einer Schule

Schwellen für „Failing Schools“

- Das Inspektorat evaluiert die Schulen mithilfe einer vierstufigen Skala: A, B, C, D. Das Evaluationsergebnis A kann zu einer Streichung der Schule aus dem Schulregister führen. Das Evaluationsergebnis D wird als vorbildlich angesehen.

Datensammlung und -auswertung

- Konzept umfasst 7 Hauptbereiche, aufgeteilt in 22 Indikatoren, die in 60 Subindikatoren unterteilt sind.
- Datensammlung: Unterrichtsbeobachtung (mit Beobachtungsprotokoll), Analyse der Ergebnisse der Selbstevaluation der Schulen, Inspektionsanalyse der Schuldokumente, Inspektionen vor Ort, Beobachtungen der Schülerinnen/Schüler und Lehrerinnen/Lehrer im Unterricht und standardisierte Interviews mit pädagogischen Mitarbeitern sowie Kontakt zu Schulgründern.
- Output wird anhand der Abschlussprüfungen der Oberstufen weiterführender Schulen gemessen, die wesentlich in der Verantwortung der Einzelschulen liegen. Keine nationalen Prüfungen, keine Tests in Grundschulen.
- Standardisierte Tests seit 2011

Die Zusammenfassung in der Übersicht zeigt, dass sich alle sechs Inspektorate auf beeinflussbare Bedingungen und Prozesse konzentrieren, einige jedoch zusätzlich noch den Output der Schule evaluieren. Und nur eines der Inspektorate befasst sich ausdrücklich mit Aspekten des Inputs von Schulen (in der Tschechischen Republik). Interessanterweise sammeln zwei Länder (Irland und England) Informationen zum Input (z. B. die Qualität der Schulgebäude), verwenden diese Informationen jedoch nur, wenn die Inspektorinnen und Inspektoren der Ansicht sind, dass dieser Input sich besonders auf die Qualität der schulischen Prozesse auswirkt. Diese Inspektorate benoten den Input einer Schule nicht separat. Die Wahl dieses Ansatzes scheint die komplexe Verknüpfung der Input- und Prozessbedingungen mit Schulqualität und guten Schülerleistungen anzuerkennen.

Eine solche Sichtweise fehlt jedoch in der generellen Bewertung von Schulprozessen und Schul-Output durch die sechs Inspektorate. Die Übersicht zeigt auf, dass sich die Gesamtbewertung der Schulqualität auf Standards, Substandards und weitere Unterteilungen in detaillierte Kriterien stützt, wobei bei der Benotung von Schulen auf einer Skala strenge Regeln in ähnlicher Weise auf alle Schulen angewandt werden. Bei den Beurteilungsprotokollen und Leitfäden muss eine Bewertung der Bedingungen mit „vorhanden“ oder „nicht vorhanden“ oder eine Bewertung auf einer 3- oder 4-stufigen Skala angegeben werden. Dabei wird die mögliche Verbindung der Bedingungen untereinander nicht berücksichtigt.

Die Art und Weise, wie Informationen auf Klassenebene aggregiert werden, um die Wirksamkeit auf Schulebene zu evaluieren, lässt ebenfalls die Tatsache unberücksichtigt, dass diese Bedingungen auf unterschiedlichen Ebenen in der Schulhierarchie unterschiedliche Bedeutungen haben. Feststellungen zur „Leistungsorientierung“ bei Unterrichtsbeobachtungen (auf Klassenebene) werden z. B. oft einfach gemittelt, um zu einer Bewertung der Qualität der Schule in diesem Bereich zu kommen. Auf Schulebene kann Leistungsorientierung jedoch auch eine Bewertung

der Schulphilosophie in diesem Bereich beinhalten und Kontrollsysteme einschließen, die die Schule aufgebaut hat, um die Ausrichtung der Lehrerinnen und Lehrer auf gute Schülerleistungen zu unterstützen. Es sind nur einige der Schulinspektorate (z. B. Schweden und Österreich), die einen Bogen um diesen Ansatz machen, der Schul- und Klassenbedingungen über einen Kamm schert. Diese Schulinspektorate haben eine ganzheitlichere Herangehensweise an Schulevaluationen und geben Schulen einen Überblick über ihre Stärken und Schwächen, statt mit rudimentären Schwellen zu arbeiten, um „*Failing Schools*“ herauszugreifen. Der Nachteil bei diesem Ansatz besteht jedoch in einem möglichen Mangel an Genauigkeit und Transparenz in den Inspektionsbewertungen.

Zum Thema des pauschalen oder spezifischer operationalisierten Inspektionsstandards bleibt noch anzumerken, dass die Beispiele aus quantitativen Indikatorensystemen, wie sie in diesem Kapitel angeführt werden, in den meisten Fällen spezifischer sind als die Inspektionsstandards, die in Tabelle 1 gelistet sind. Dabei sollte man jedoch die Tatsache nicht aus den Augen verlieren, dass sich systematische quantitative Indikatorensysteme von Schulinspektionen insofern unterscheiden, als Letztere den großen Vorteil des professionellen Urteilsvermögens von Experten mit sich bringen, während Erstere von Messtechniken abhängen. Trotzdem könnte man die extensiveren Indikatorensets als Ressource bei der Entwicklung von Inspektionskonzepten verwenden und möglicherweise ließen sie sich teilweise als Grundlage für strukturierte operative Abläufe bei Klassenbeobachtungen und als „Stützgerüste“ für die professionellen Einschätzungen von Schulinspektoren und -inspektoren übernehmen. Was Ausrichtung und Gestaltung von Inspektionskonzepten angeht, so gibt es viele Optionen und eine große Auswahl an Instrumenten. Das von Scheerens et al. (2011) entwickelte Indikatorenset, das im Anhang vorliegt, soll als Beispiel für eine sachbezogene Zusammenstellung von Inspektionsstandards dienen.

Einige Schulinspektorate (z. B. OFSTED in England) haben solche detaillierten Checklisten jedoch aufgegeben, nachdem standardisierte und schriftlich festgelegte Rezepte zur Schulentwicklung aus ihnen geworden waren. Schulen und andere Stakeholder (z. B. Partner bei der Schulverbesserung und Entwickler von Selbstevaluationen für Schulen) entwickelten sogenannte „OFSTED-genehmigte“ Modelle für Schulorganisation und Unterricht, die immer öfter von Schulen kopiert und eingeführt wurden, ohne dass geprüft wurde, ob diese Praktiken für die spezielle Klasse oder den Schulkontext geeignet waren, auf den sie angewendet werden sollten.

Die Übersicht über die Inspektionskonzepte fasst auch zusammen, wie eine Reihe von Schulinspektoren den Output von Schulen in ihre Evaluation der Schulqualität einbeziehen. Die Beschreibung, wie Testdaten analysiert werden, um den Schul-Output zu bewerten, macht deutlich, dass diese Evaluationen immer noch recht rudimentär sind und nur in eingeschränktem Maße auf die zu Be-

ginn des Kapitels dargestellten ausgefeilteren *Value-Added*-Modelle für Analyse und Bericht von Schul-Output zurückgreifen. Die Schulinspektorate, die Zugang zu Schülerleistungsdaten haben (in den Niederlanden und England) berücksichtigen eine mögliche Instabilität in den Daten, indem sie bei der Bewertung des Schul-Outputs Durchschnittswerte berechnen (normalerweise aus drei Jahren). Die meisten Schulinspektorate scheinen jedoch keine detaillierten und qualitativ hochwertigen Leistungsdaten zur Verfügung zu haben, um solche Analysen vornehmen zu können. Entweder gibt es keine nationalen standardisierten Tests oder sie dürfen nicht auf solche Daten zugreifen. Stattdessen analysieren einige Schulinspektorate Bewertungen durch Lehrerinnen und Lehrer oder Schülerarbeiten, um sich eine Vorstellung vom Output einer Schule zu machen. Interessanterweise werden solche Analysen in manchen Fällen (z.B. in Irland) auch benutzt, um die Qualität von Prozessen zu beurteilen (z.B. die Unterrichtsqualität). Dadurch dienen sie als stellvertretende Messgröße für die Qualität bildungsbezogener Prozesse an der Schule.

An der Übersicht kann man die bemerkenswerte Tatsache ablesen, dass die Unterrichtsqualität zwar als Hauptbedingung für Schulqualität angesehen und entsprechend betont wird, dass aber keines der Schulinspektorate die Merkmale von Lehrerinnen und Lehrern in sein Konzept der Inspektionsstandards aufgenommen hat. Obwohl das niederländische Inspektionskonzept Kriterien auf Klassenebene wie das pädagogische und didaktische Verhalten von Lehrerinnen und Lehrern einbezieht, werden sie lediglich auf Schulebene bewertet, wenn die Ergebnisse einer Auswahl von Unterrichtsbeobachtungen gemittelt werden. Wie wir bereits ausgeführt haben, sind jedoch Bedingungen auf Unterrichtsebene wie hohe Erwartungen, ein anspruchsvoller Unterrichtsansatz, eine geordnete Lernumgebung und klarer und strukturierter Unterricht wichtiger für die Verbesserung der Schülerleistungen als Bedingungen auf Schulebene. Die meisten Inspektorate evaluieren Lehrerinnen und Lehrer oder Unterricht jedoch nicht ausdrücklich auf der Ebene von Klasse/Fach oder Klassenstufe, sondern evaluieren lieber die Bedingungen auf Schulebene und allgemeine Merkmale des Unterrichts oder Unterrichtsmuster wie Lernzeit, Schulleitungshandeln und Schulklima.

Eine abschließende „Empfehlung“ für zukünftige und weiterreichende Entwicklungen nationaler Schulinspektionskonzepte ist daher, die Betonung von Standards schulischer Organisation in Relation zu Lehre und Lernen oder didaktischen Standards zu setzen. Eine vorsichtige Warnung an Gestalter von Schulinspektionskonzepten könnte sein, die primären Prozesse des Lehrens und Lernens nicht aus den Augen zu verlieren.

Literatuur

- American Statistical Association (2014). *ASA statement on using value-added models for educational assessment*. Alexandria, VA: Author. Verfügbar unter: https://www.amstat.org/policy/pdfs/ASA_VAM_Statement.pdf. [Zugriff am 17.03.2015]
- Bosker, R. J., Guldemon, H. G., Hofman, R. H. & Hofman, W. H. A. (1989). De stabiliteit van schoolkwaliteit. In J. Scheerens & J. C. Verhoeven (Hg.), *Schoolorganisatie, beleid en onderwijskwaliteit* (S. 39–53). Lisse: Swets & Zeitlinger.
- Cotton, K. (1995). *Effective schooling practices: A research synthesis. 1995 Update* (School Improvement Research Series). Northwest Regional Educational Laboratory.
- Creemers, B. P. M. & Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness*. London, New York: Routledge.
- De Grauwe, A. (2007). *Module 7. Alternative models in reforming school supervision*. Verfügbar unter: http://www.unesco.org/iiep/PDF/TR_Mods/SUP_Mod7.pdf [Zugriff am 17.03.2015]
- De Volder, I. (2012). *Externe schoolevaluaties in Europa. Een vergelijkend onderzoek*. Antwerpen: Grant uitgevers.
- Dijkstra, A. B., De la Motte, P. I., Ehren, M. C. M. & Eilard, A. (2014). Discussion. School inspections and school improvement in the social domain. The assessment of social outcomes of education. In: A. B. Dijkstra & P. I. de la Motte (Hg.). *Social Outcomes of Education; The assessment of social outcomes and school improvement through school inspections* (S. 189–215). Amsterdam: Amsterdam University Press.
- Donaldson, M. L. & Johnson, S. M. (2010). The price of misassignment: The role of teaching assignments in Teach for America teachers' exit from low-income schools and the teaching profession. *Educational Evaluation and Policy Analysis*, 32 (2), 299–323.
- Eddy Spicer, D., Ehren, M., Bangpan, M., Khatwa, M. (2014). *Under what conditions do inspection, monitoring and assessment improve system efficiency, service delivery and learning outcomes for the poorest and most marginalised? A realist synthesis of school accountability in low- and middle-income countries. Protocol*. London: EPPI-Centre, Social Science Research Centre, Institute of Education, University of London. Verfügbar unter: <http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=x-3hzguBXhY%3D&tabid=3174> [Zugriff am 17.03.2015]
- Ehren, M. C. M., Altrichter, H., McNamara, G. & O'Hara, J. (2013). Impact of school inspections on school improvement; describing assumptions on causal mechanisms in six European countries. *Educational Assessment, Evaluation and Accountability*, 25 (1), 3–43. Verfügbar unter: <http://dx.doi.org/10.1007/s11092-012-9156-4> [Zugriff am 17.03.2015]
- Ehren, M. C. M. & Dijkstra, A. B. (2014). Evaluating Social Outcomes through School Inspections. In: A. B. Dijkstra & P. I. de la Motte (Hg.). *Social Outcomes of Education; The assessment of social outcomes and school improvement through school inspections* (S. 51–73). Amsterdam: Amsterdam University Press.

- Gray, J., Jesson, D., Goldstein, H., Hedges, K. & Rasbash, J. (1995). A multi-level analysis of school improvement: Changes in school's performance over time. *School Effectiveness and School Improvement*, 6, 97–114.
- Hamilton, L. S. & Koretz, D. M. (2002). Tests and their use in test-based accountability systems. In L. S. Hamilton, B. M. Stecher & S. P. Klein (Hg.). *Making sense of Test-based Accountability in Education*. Santa Monica: Rand cooperation. Verfügbar unter: http://www.rand.org/pubs/monograph_reports/MR1554/ [Zugriff am 17.03.2015]
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: an update. *Educational Evaluation and Policy Analysis*, 19 (2), 141–164.
- Hanushek, E. A. (1986). The Economics of Schooling. Production and Efficiency in Public Schools. *Journal of Economic Literature*, 24 (3), 1141–1177
- Hattie, J. (2009). *Visible Learning*. Abingdon: Routledge.
- Hopkins, D., Stringfield, S., Harries, A., Stoll, L. & Mackay, T. (2014). School and System Improvement. *School Effectiveness and School Improvement*, 25, 257–281.
- Levine, D. K. & Lezotte, L. W. (1990). *Unusually Effective Schools: A Review and Analysis of Research and Practice*. Madison, Wise: Nat. Centre for Effective Schools Research and Development.
- Louis, K. S., Dretzke, B. & Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement*, 21 (3), 315–336.
- Marzano, R. (2003). *What works in schools. Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Mourshed, M., Chijioke, C. & Barber, M. (2010). *How the world's most improved school systems keep getting better*. London: McKinsey. Verfügbar unter: <http://mckinseyon-society.com/how-the-worlds-most-improved-school-systems-keep-getting-better/> [Zugriff am 17.03.2015]
- Muijs, D., Creemers, B., Kyriakides, L., Van der Werf, G., Timperley, H. & Earl, L. (2014). Teaching Effectiveness. A State of the Art Review. *School Effectiveness and School Improvement*, 24, 231–256.
- Popham, W. J. (2003). *Test better, teach better. The instructional role of assessment*. Alexandria, VA: USAL ASCD.
- Purkey, S. C. & Smith, M. S. (1983). Effective schools: a review. *The Elementary School Journal*, 83, 427–452.
- Reynolds, D., Sammons, P., De Fraine, B., Townsend, T., Van Damme, J., Teddlie, C. & Stringfield, S. (2014). Educational Effectiveness Research (EER): A state of the Art Review. *School Effectiveness and School Improvement*, 25 (2), 197–230.
- Sammons, P., Hillman, J. & Mortimore, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research*. London: OFSTED.
- Scheerens, J. (1992). *Effective Schooling, Research, Theory and Practice*. London: Cassell.
- Scheerens, J. (2007) *Conceptual framework for the PISA 2009 background questionnaires*. Internal report, University of Twente.

- Scheerens, J. (2013). *What is effective schooling? A review of current thought and practice*. Paper for the International Baccalaureate Organization; Washington DC.
- Scheerens, J. (2014) School, teaching and system effectiveness: some comments on three state of the art reviews. *School effectiveness and School Improvement*, 25, 282–290.
- Scheerens, J. & Creemers, B. P. M. (1989). Conceptualizing school effectiveness. Development in school effectiveness research. *Special issue of the International Journal of Educational Research*, 13 (7), 691–706.
- Scheerens, J., Glas, C. A. W. & Thomas, S. M. (2003). *Educational evaluation, assessment, and monitoring: a systemic approach*. Lisse: Swets & Zeitlinger.
- Scheerens, J., Luyten, H., Van den Bergh, S. M. & Glas C. A. W. (2015). Exploration of direct and indirect associations of system level policy amenable variables and reading literacy performance. *Educational Research and Evaluation*. (im Erscheinen)
- Scheerens, J., Luyten, H., Steen, R. & Luyten-de Thouars, Y. (2007). *Review and meta-analyses of school and teaching effectiveness*. Enschede: University of Twente, Department of Educational Organisation and Management.
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77 (4), 454–499.
- Thomas, S. M., Peng, W. J. & Gray, J. (2010). Modeling patterns of improvement over time: value added trends in English secondary schools. *Oxford Review of Education*, 33, 261–295.
- Tolofari, S. (2005) „New public management and education“. *Policy Futures in Education*, 3 (1), 75–89. Verfügbar unter: <http://firgoa.usc.es/drupal/files/tolofari.pdf> [Zugriff am 17.03.2015]
- Van Bruggen, J. C. (2010). *Inspectorates of Education in Europe: some comparative remarks about their tasks and work*. Standing International Conference of Inspectorates of Education in Europe (SICI). Verfügbar unter: <http://www.sici-inspectorates.eu/Members/Comparative-Analyses/Inspectorates-of-Education-in-Europe> [Zugriff 17.03.2015]
- Vermeer, N. & Van der Steeg, M. (2011). *Onderwijsprestaties Nederland in Internationaal Perspectief*. CPB Achtergronddocument bij CPB Policy Brief 05, 2011. Den Haag: CPB.
- Visscher, A. & Ehren, M. (2011). *De eenvoud en complexiteit van Opbrengstgericht Werken*. Verfügbar unter: <http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2011/07/13/de-eenvoud-en-complexiteit-van-opbrengstgericht-werken.html> [Zugriff 30.03.2015]

Anhang

Detaillierteres Indikatorenset für Output, Input und schulische Prozesse
(Scheerens et al. 2011)

Tabelle 1: Überblick über Indikatoren für Bildungs-Outcome

Hauptkategorien von Outcome-Indikatoren	Unterkategorien	Technische Angaben
Output-Indikatoren	Leistungsmessungen fächergestützt Literacy (Lesen, Mathematik, Naturwissenschaft) Kompetenzen (z. B. das Lernen lernen)	Effektmessung des <i>Value Added</i> ; Wachstumskurven Bewertungsmethodik (von Multiple-Choice-Tests bis hin zu authentischen Leistungsmes- sungen) Kriteriums- bzw. norm- orientierte Testverfahren
Indikatoren für Out- come/Zielerreichung	Messungen der Zielerreichung Schulabschlussquote Anteil der Schülerinnen und Schüler, die ihren Schulab- schluss ohne Wiederholung einer Klasse erlangt haben Quote der Abbrecher Quote der Wiederholer	Auswahlorientierte Schulricht- linien, statistisch kontrolliert

Tabelle 2: Indikatoren für finanzielle und materielle Ausstattung auf Schulebene

Finanzielle und materielle Ausstattung auf Schulebene
<ul style="list-style-type: none"> • Anteil des Schulbudgets, der nicht aus öffentlichen Quellen stammt • Schulgebäude und Ausstattung • Ausstattung der Klassenräume (Möbiliar, Computer usw.) • Schulartikel wie Bleistifte und Papier, Schultafel, Flipchart • Verfügbarkeit von Lehrbüchern in den wichtigsten Unterrichtsfächern • Grundausstattung wie getrennte Toiletten für Jungen und Mädchen, Wasser, Strom, Heizung, Telefon, Verfügbarkeit von zusätzlichen Leistungen in Bezug auf Ernährung, Gesundheit und Beförderung

Tabelle 3: Überblick über Beispiele für Prozessindikatoren für Funktionsweisen von Schulen

Prozessindikatoren auf Schulebene
<p><i>Einbeziehung der Umgebung</i></p> <ul style="list-style-type: none"> • Der Grad der tatsächlichen Einbeziehung von Eltern in unterschiedliche Schulaktivitäten (Lehr- und Lernprozess, außerschulische Aktivitäten, unterstützende Aktivitäten) • Der prozentuale Anteil des gesamten jährlichen Schulhaushaltes, der aus Quellen vor Ort finanziert wird • Das Maß der Mitsprachemöglichkeiten, die Schulträger in Bezug auf die Arbeitsbedingungen von Lehrerinnen und Lehrern haben [mögliche Operationalisierungen im EDUCO-Projekt, El Salvador]
<p><i>Finanzielle und personelle Mittel der Schule</i></p> <ul style="list-style-type: none"> • durchschnittliche Berufserfahrung der Lehrerinnen und Lehrer einer Schule in Jahren • zahlenmäßiges Schüler-Lehrer-Verhältnis an einer Schule • durchschnittliche Klassenstärke pro Schule • Anteil der Lehrerinnen und Lehrer mit regulärer Ausbildung pro Schule • Personelle Ausstattung der „Spitze“ einer Schule (Schulleiter und stellvertretender Schulleiter pro 100 Schülerinnen und Schüler)
<p><i>Leistungsorientierte Ausrichtung</i></p> <ul style="list-style-type: none"> • Ob Schulen Leistungsstandards formulieren oder nicht • Wie intensiv Schulen den (bildungsbezogenen) Werdegang ihrer Schülerinnen und Schüler verfolgen, nachdem sie die Schule verlassen haben • Ob Schulen Leistungsergebnisse an interessierte Zielgruppen weitergeben oder nicht
<p><i>Schulleitungshandeln</i></p> <ul style="list-style-type: none"> • Zeit, die Schulleiter auf pädagogische Arbeit verwenden (im Vergleich zu Verwaltungsangelegenheiten oder anderen Aufgaben) • Ob Schulleiter die Leistung von Lehrerinnen und Lehrern bewerten oder nicht • Zeit, die bei Konferenzen auf unterrichtsbezogene Themen verwendet wird
<p><i>Kontinuität und Konsens im Kollegium</i></p> <ul style="list-style-type: none"> • Ausmaß personeller Veränderungen in einem bestimmten Zeitraum • Existenz oder Fehlen fächerbezogener Arbeitsgruppen oder Abteilungen (weiterführende Schulen) • Häufigkeit und Dauer formeller und informeller Zusammenkünfte des Kollegiums
<p><i>Geregeltes und sicheres Klima</i></p> <ul style="list-style-type: none"> • Statistisches Erfassen von häufigem Fehlen oder Schulschwänzen • Beurteilung der Schuldisziplin durch Schulleiter, Lehrerinnen/Lehrer und Schülerinnen/Schüler

Effiziente Nutzung der Zeit

- Gesamte Unterrichtszeit und Zeit pro Fach
- Durchschnittlicher Zeitverlust pro Unterrichtsstunde (durch organisatorische Notwendigkeiten, Raum- oder Gebäudewechsel, Störungen)
- Prozentualer Anteil ausgefallener Stunden pro Jahr

Lerngelegenheiten

- Beurteilung durch Lehrerinnen und Lehrer oder Schülerinnen und Schüler, ob jedes Item in einem Leistungstest unterrichtet wurde oder nicht

Evaluation der Lernfortschritte der Schülerinnen und Schüler

- Häufigkeit von curriculumspezifischen Test in den einzelnen Klassenstufen
- Häufigkeit von standardisierten Leistungstests
- Testergebnisse und ihre Verwendung durch die Lehrerinnen und Lehrer

Bewertung der Lehrerqualität

- Beurteilung der Unterrichtsqualität durch Peers (andere Lehrerinnen und Lehrer)
- Beurteilung der Unterrichtsqualität durch Schülerinnen und Schüler

Tabelle 4: Überblick über effektive Unterrichts- und Lernvariablen

Effektive Unterrichtsvariablen
<p><i>Hauptsächliche Unterrichtsfaktoren</i></p> <ul style="list-style-type: none"> • Lerngelegenheiten • Struktur und stützende Gerüste (kognitive Struktur) • Anregende Einbeziehung (Motivationsstruktur) • Aspekte des Klimas: <ul style="list-style-type: none"> • aufgabenorientiert • gegenseitiger Respekt • geordnet, sicher • Monitoring und Fragen • Feedback und Bestätigung • Modellhaftes Vermitteln von Lern- und Selbstregulierungsstrategien • „Authentische“ Anwendungen • Adaptiver Unterricht
<p><i>Lernstrategien von Schülerinnen und Schülern</i></p> <ul style="list-style-type: none"> • sichtbar: <ul style="list-style-type: none"> • Aktive Lernzeit • Arbeit mit Materialien • Kooperatives Lernen • mental: <ul style="list-style-type: none"> • Selbstregulierung • Selbstkontrolle • Meta-kognitive „Handlungen“ • Lernstile

Zusammenhang von Schul- und Unterrichtsqualität und schulischen Rahmenbedingungen

Holger Gärtner

Zusammenfassung

Dieser Beitrag untersucht in einer Sekundäranalyse von Hamburger Schulinspektionsdaten Zusammenhänge zwischen den Rahmenbedingungen, unter denen Schulen arbeiten, und den innerschulischen Prozessen auf Schul- und Klassenebene. Als Indikator für die Rahmenbedingungen dient der Hamburger Sozialindex, der Merkmale der Schülerkomposition erfasst (kulturelles, soziales, ökonomisches Kapital, Migrationshintergrund). Mittelwertvergleiche zeigen, dass nur wenige Gesamtbewertungen der Schulinspektion stärker mit dem Sozialindex zusammenhängen. Stärkere Zusammenhänge zeigen sich zwischen Sozialindex und den Wahrnehmungen der Schülerinnen und Schüler von Aspekten der Schul- und Unterrichtsqualität. Positive Zusammenhänge lassen sich durch kompensatorische Maßnahmen von Schulen erklären, z.B. indem Schulen mit einer sozial benachteiligten Schülerschaft mehr Förderangebote bereitstellen. Die Ergebnisse tragen zur Diskussion über die Frage der Adjustierung von Inspektionsbewertungen bei.

Einleitung

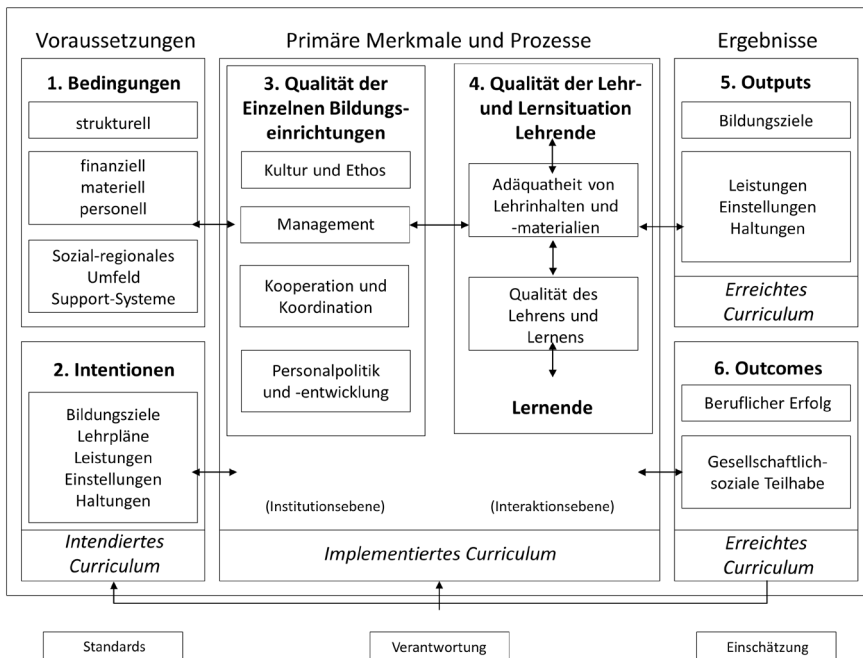
Thematisiert wird der Beitrag, den Inspektionsdaten leisten können, um grundlegende Forschungsfragen im Bereich der Schuleffektivitäts- und Qualitätsforschung zu beantworten (vgl. Wurster & Gärtner 2013). Dies geschieht anhand eines Beispiels, bei dem der Zusammenhang zwischen Rahmenbedingungen, unter denen Schulen arbeiten, und den innerschulischen Prozessen auf Schul- und Klassenebene, betrachtet wird. Analysiert werden Zusammenhänge zwischen dem Sozialindex einer Schule, den im Rahmen der Hamburger Schulinspektion getroffenen Qualitätsurteilen sowie den Wahrnehmungen zur Schul- und Unterrichtsqualität. Das 1. Kapitel (Theorie) skizziert Modellannahmen über die Auswirkungen schulischer Rahmenbedingungen auf die Unterrichts- und Managementprozesse. Das 2. Kapitel (Methode) beschreibt das methodische Vorgehen und die untersuchten Variablen, also den in Hamburg genutzten Sozialindex sowie die im Rahmen der Schulinspektion erhobenen Merkmale von Schul- und Unterrichtsqualität. Kapitel 3 (Ergebnisse) stellt Ergebnisse zu Zusammenhängen zwischen Rahmenbedingungen und Ausprägung beurteilter Aspekte von Schul- und Unterrichtsqualität in

den einzelnen Schularten dar, die im abschließenden Kapitel (Diskussion) ebenso diskutiert werden wie die Bedeutung von Inspektionsdaten für die empirische Bildungsforschung generell.

1. Theorie

1.1 Modellannahmen zum Zusammenhang von Input- und Prozessmerkmalen

Abbildung 1: Modell zu Qualität und Qualitätsentwicklung im Bildungsbereich



Quelle: Ditton 2007, S. 84

Modelle der Schulqualität und Schuleffektivität beschreiben relevante Einflussfaktoren für die Leistungsentwicklung von Schülerinnen und Schülern. Typischerweise werden vier Kategorien von Einflussfaktoren unterschieden (s. Abbildung 1): 1) Kontextfaktoren, unter denen Schulen arbeiten und welche sie selbst nicht ändern können, wie z. B. das räumliche Umfeld; 2) Inputfaktoren wie z. B. die personelle, räumliche oder finanzielle Ausstattung der Schule sowie gesellschaftliche Zielerwartungen in Form von Lehrplänen oder Prüfungsanforderungen; 3) Prozessmerkmale wie die Qualität des Unterrichts oder des Schulmanagements sowie

4) Outputfaktoren wie z.B. Schülerkompetenzen (vgl. Creemers & Kyriakides 2008; Ditton 2007; Scheerens & Bosker 1997). Diese Modelle implizieren, dass Schulen als organisationale Systeme funktionieren und dass ein Zusammenwirken der unterschiedlichen Faktoren den gewünschten Output auf Schülerebene optimieren kann (vgl. Wendt & Bos 2011).

In jüngeren empirischen Untersuchungen der Leistungsentwicklung von Schülerinnen und Schülern werden Leistungsergebnisse kontrolliert, also um Einflüsse von Faktoren „bereinigt“, auf die eine Schule oder Lehrkraft keinen oder kaum Einfluss haben (vgl. Arnold 1999; Baumert et al. 2006; Bellin 2009; Bos et al. 2010). Dies sind typischerweise Kontextfaktoren, also Informationen über Hintergrundmerkmale der Schüler/-innen, wie z.B. sprachliche Merkmale als Indikator für den Migrationshintergrund (z. B. deutsche oder nicht-deutsche Herkunfts- oder Familiensprache) oder deren sozioökonomischer Status (ökonomisches, kulturelles oder soziales Kapital der Schüler/-innen bzw. deren Eltern). Empirische Studien zur Leistungsentwicklung von Schülerinnen und Schülern zeigen eindeutig, dass diese individuellen Schülermerkmale mit der Schülerleistung in Zusammenhang stehen (vgl. Dumont et al. 2013; van Ewijk & Slegers 2010).

Wenn über den *individuellen* Einfluss der genannten Merkmale hinaus auf Schul- oder Klassenebene aggregierte Merkmale des Kontexts einen *zusätzlichen* Einfluss auf die Schülerleistung haben, spricht man von einem sogenannten *Kompositionseffekt* (vgl. Harker & Tymms 2004). Dies bedeutet, dass die Zusammensetzung (Komposition) der Schülerschaft einer Klasse oder Schule zu unterschiedlichen Lernbedingungen führen kann und damit zusätzlich zu den individuellen Lernvoraussetzungen der Schülerinnen und Schüler einen Einfluss auf die Schulleistung hat (vgl. Dumont et al. 2013). Je nach Schülerkomposition können Leistungsergebnisse von Schulen somit stärker oder schwächer ausfallen, ohne dass dies mit tatsächlichen Unterschieden in der Qualität der Schulen zusammenhängt. So zeigen z. B. Studien aus etlichen Ländern, dass Schülerinnen und Schüler, die von einer sozial privilegierteren Schülerschaft umgeben sind, höhere Schulleistungen aufweisen und mehr dazulernen, als Schülerinnen und Schüler mit vergleichbaren Voraussetzungen, die sich in einer Gruppe mit Schülern aus sozial eher schwachen Familien befinden (vgl. Dumont et al. 2013).

1.2 Erklärungen für Kompositionseffekte

Obwohl nationale wie internationale Untersuchungen regelmäßig Hinweise auf Kompositionseffekte liefern, gibt es hierfür relativ wenige theoretisch fundierte Erklärungsansätze (vgl. Dumont et al. 2013). Dies liegt z.T. an den Designs der durchgeführten Studien, die meist einem Black-Box-Ansatz unterliegen, d. h., Kontextmerkmale werden zwar bei der Erklärung von Schulleistungen kontrolliert, die

weiteren Variablen der oben skizzierten Modelle von Schulqualität werden jedoch nicht berücksichtigt (vgl. Hallinger & Heck 1996). Hierzu gehört insbesondere die Prozessebene, also die Frage des Zusammenhangs zwischen Kontextbedingungen, unter denen Schulen und Lehrkräfte arbeiten, und den Prozessen, die sie daran anpassend in den Klassen oder im Schulmanagement realisieren (vgl. van Ewijk & Sleegers 2010).

Hallinger und Heck (1996) kritisieren, bei empirischen Schulleistungsstudien den Zusammenhang zwischen Kontextbedingungen und schulischen Prozessen nicht zu untersuchen. In ihrem Review über die Wirkung von Schulleitungshandeln beschreiben sie verschiedene Modelle, wie sich Schulleitungshandeln auf Schulleistung auswirken kann. Neben einfachen Modellen, die einen direkten Einfluss des Leitungshandelns auf Schulleistung postulieren, benennen sie Mediations- und reziproke Modelle. Teilweise beinhalten diese Modelle auch Kontextfaktoren, die einen Einfluss auf das Schulleitungshandeln haben. So beeinflusst z. B. der sozioökonomische Hintergrund der Schülerschaft die Erwartungshaltung der Schulleitung sowie die Schwerpunktsetzung seiner Handlungen, u. a. im Bereich der unterrichtsbezogenen Führung (vgl. Hallinger & Murphy 1986).

Die letztgenannten Modelle werden in empirischen Studien jedoch kaum realisiert (vgl. Hallinger & Heck 1996), obwohl Hallinger und Murphy schon 1986 große Unterschiede in den Handlungsweisen von Schulleitungen und Lehrkräften in Schulen mit unterschiedlichem sozioökonomischem Status der Schülerschaft nachwiesen. Hierzu identifizierten sie aus allen kalifornischen Grundschulen (N=3.100) diejenigen, die über drei Jahre hinweg sowohl in der dritten als auch in der sechsten Klasse in allen Fachleistungen über den Erwartungswerten von Schulen mit vergleichbarer Schülerzusammensetzung lagen. Von den sechzehn so identifizierten Schulen untersuchten sie acht, deren Schülerschaft sich in ihrem sozioökonomischen Hintergrund deutlich unterschied. Durch eine sehr umfangreiche Diagnostik¹ konnten sie Unterschiede auf Prozessebene zwischen Schulen mit starker und schwacher sozioökonomischer Schülerschaft aufzeigen, die in Tabelle 1 im Überblick dokumentiert sind.

Die Ergebnisse von Hallinger und Murphy in Tabelle 1 zeigen, dass es einen engen Zusammenhang zwischen dem sozioökonomischen Hintergrund der Schülerschaft und den operativen Prozessen auf Schul- und Unterrichtsebene gibt. Starke Unterschiede ergeben sich insbesondere hinsichtlich der Breite des schulinternen Curriculums, der curricularen Schwerpunkt- und Zielsetzungen, der unterrichtsbezogenen Führung, des Einsatzes von Belohnungssystemen für Schüler/innen, den Erwartungs-

1 Diese Studie ist aufgrund ihrer Ergebnisse für den vorliegenden Beitrag relevant, da das Vorgehen der Forscher letztlich dem einer Inspektion entspricht. An den Schulen wurde Unterricht beobachtet, es wurden Befragungen mit Schülern, Eltern und Lehrkräften durchgeführt, ebenso Interviews mit diesen Gruppen und der Schulleitung.

haltungen an Schüler/innen sowie der Beziehung zwischen Schule und Elternhaus. Die effektiven Schulen stellen sich offensichtlich mit spezifischen pädagogischen Maßnahmen auf ihre Schülerschaft ein und generieren so höhere Leistungsergebnisse als dies in Schulen mit vergleichbarer Schülerkomposition der Fall ist.

Tabelle 1: Profile effektiver Schulen mit sozial schwacher bzw. starker Schülerschaft

Schulmerkmale	Schule mit sozial schwacher Schülerschaft	Schule mit sozial starker Schülerschaft
<i>Schulinternes Curriculum</i>		
Breite	eng	breit
Anforderungen	grundlegend	akademisch
Verknüpfung mit Unterricht	moderat	eng
<i>Lerngelegenheiten</i>		
Schwerpunktsetzungen	Basiskompetenzen	akademisch
Erwartungen an Hausaufgaben	gering bis moderat	hoch
<i>Leitbild</i>		
Art des Leitbildes	Beherrschen von Basis-kompetenzen	Beherrschen akademischer Leistungen
geteilte Überzeugung im Kollegium	hoch	hoch
<i>Unterrichtsbezogene Führung</i>		
Koordination des Curriculums	hoch	hoch
Überwachung des Unterrichts	hoch	gering bis moderat
Aufgabenorientierung	hoch	moderat
Beziehungsorientierung	gering bis moderat	moderat bis hoch
<i>Kooperation Schule-Eltern</i>		
Nähe zum Elternhaus	gering	hoch
Beteiligung der Eltern	gering	allgegenwärtig
Rolle der Schulleitung	Puffer	Realisierung von Schnittstellen
<i>Belohnung</i>		
Häufigkeit von Belohnungen	hoch	gering
Art der Belohnungen	extrinsisch, öffentlich	intrinsisch, privat
<i>Hohe Erwartungen</i>		
Quelle der Erwartungen	Schule	Schule und Elternhaus
aktuelle Erwartungen	hoch	hoch
zukünftige Erwartungen	moderat	hoch

Quelle: Hallinger & Murphy 1986, eigene Übersetzung

Zur Erklärung der gefundenen Unterschiede liefert die Arbeit von Dumont und Kollegen (2013) den aktuellen Stand über Wirkannahmen von Kompositionseffekten:

1. Schülerkomposition wirkt vermittelt über Anpassungsreaktionen der Lehrkraft,
2. Schülerkomposition wirkt durch Interaktion zwischen Schülerinnen und Schülern (*Peer Effect*),
3. Schülerkomposition wirkt vermittelt über unterschiedliche Ressourcen für Lernprozesse.

Schülerkomposition wirkt vermittelt über Anpassungsreaktionen der Lehrkraft

Diese Wirkannahme geht davon aus, dass Lehrkräfte ihr Verhalten und ihre Leistungserwartungen an die Schülerschaft, die sie unterrichten, anpassen. Leistungsschwache Klassen werden entsprechend weniger anspruchsvoll und langsamer unterrichtet, da Lehrkräfte an Klassen bestimmter Komposition (z.B. mit einem hohen Anteil an Schülern nicht-deutscher Herkunftssprache) geringere Leistungserwartungen stellen (vgl. Dumont et al. 2013; Harker & Tymms 2004; Harris 2010; Opdenakker & van Damme 2007).

Schülerkomposition wirkt durch Interaktion zwischen Schülerinnen und Schülern (Peer Effect)

Diese Wirkannahme geht davon aus, dass sich je nach Komposition unterschiedliche Interaktionen zwischen Schülerinnen und Schülern im Sinne einer sozialen Norm ergeben, die individuelle Lernprozesse positiv oder negativ beeinflussen können. Lerngruppen, die überwiegend aus Schülerinnen und Schülern sozial starker Elternhäuser bzw. mit einem hohen Leistungsniveau bestehen, schaffen eine starke Lern- und Leistungskultur, die für weitere Schüler/-innen ansteckend wirkt (vgl. Baumert et al. 2006). Lerngruppen, die überwiegend aus Schülerinnen und Schülern sozial schwacher Elternhäuser bzw. niedrigem Leistungsniveau bestehen, etablieren u. U. eine soziale Norm, in der ein leistungsorientiertes Verhalten u. a. als störend empfunden wird (vgl. Stamm 2008).

Schülerkomposition wirkt vermittelt über unterschiedliche Ressourcen für Lernprozesse

Die dritte Annahme geht davon aus, dass eine unterschiedliche Komposition von Schülerinnen und Schülern zu unterschiedlichen Ressourcenzuweisungen führt, die letztlich die individuellen Lernprozesse der Schüler/-innen beeinflusst. Eine Schule, die Schüler/-innen aus sozial starken Elternhäusern anzieht, kann z.B. über Fördervereine zusätzliche finanzielle Mittel einwerben und entsprechend in zusätzliche Ausstattung investieren (Labormittel; Instrumente, Ausflüge, etc.) (vgl. Harris 2010). Ebenso können Schulen an begünstigten Standorten leichter

gut qualifiziertes Personal rekrutieren als dies Schulen in sozialen Brennpunkten gelingt (vgl. Hallinger & Murphy 1986; Scafidi et al. 2007).

1.3 Schulinspektion und Kontextmerkmale

Seit Mitte der 2000er Jahre arbeiten Schulinspektorate in den Ländern als Form der externen Evaluation von Schulen im Routinebetrieb und liefern nicht nur für die Einzelschule wertvolle Informationen über Schul- und Unterrichtsqualität, sondern auch auf Systemebene. Beispielsweise werden in Bayern während eines Schuljahres ca. 770 Schulen extern evaluiert (ISB 2012). In Hessen wurden während der Erstinspektion von 1.815 Schulen in den Jahren 2006 bis 2010 über 400.000 Personen schriftlich befragt und über 40.000 Unterrichtssequenzen besucht (IQ Hessen 2012) (vgl. Gärtner & Wurster 2013). In Hamburg wurden im ersten Zyklus von 2006 bis 2013 alle 375 staatlichen Schulen inspiziert.

Durch diese kontinuierliche Erfassung entsteht eine Fülle an Informationen über Schul- und Unterrichtsqualität, da Schulinspektionen Vollerhebungen aller öffentlichen Schulen realisieren. Zudem gelingt ihnen, was vielen empirischen Schul- und Unterrichtsforschern verwehrt bleibt: eine direkte Beobachtung von Unterricht und somit wertvolle Einblicke ins Unterrichtsgeschehen. Neben der Erfassung und Bewertung relevanter Aspekte des Unterrichts werden ebenfalls Outputdaten sowie Prozessqualitäten auf Organisationsebene wie Schulmanagement oder Schulkultur erfasst (vgl. Döbert & Dederich 2008). In der Gesamtschau entsteht ein Bild, das sich gut mit dem Modell von Ditton (2007) beschreiben lässt (s. Abbildung 1).

Wendt und Bos (2011) stellen in ihrem Artikel die Frage, wie Schulinspektion auf Forschungsergebnisse zum Kompositionseffekt reagieren kann, z. B. „wie relevante Informationen zu den sozialen Kontextbedingungen von Schule systematisch in die Berichtslegung und Bewertung der Inspektorenteams einfließen können“ (S. 231). Diese Argumentation wird u. a. durch die dargestellte Studie von Hallinger und Murphy (1986) unterstützt. Betrachtet man jedoch die aktuelle Anlage von Inspektionsverfahren in Deutschland, so erkennt man die systematische Ausblendung von Kontextfaktoren, d.h. bewertet werden bislang nahezu ausschließlich schulische Prozessmerkmale (vgl. Wendt & Bos 2011), obwohl es in Schulen ein starkes Interesse daran gibt, auch die Voraussetzungen, unter denen sie arbeiten, bewerten zu lassen.

Van Ewijk und Slegers (2010) formulieren auf Grundlage ihrer Meta-Analyse zu Kompositionseffekten Empfehlungen für die Messung von Kontexteffekten, damit deren Einfluss weder über- noch unterschätzt wird. Die Autoren plädieren für den Gebrauch eines SES-Indikators, der mehrere relevante SES-Dimensionen umfasst. Wendt und Bos (2011) plädieren ebenfalls für einen umfassenden SES-Indikator und verorten mögliche Dimensionen anhand der Theorien von Bourdieu (1983) und

Coleman (1988). Das soziale Kapital meint im Verständnis dieser Autoren das Netzwerk sozialer Beziehungen, das Personen über die Zugehörigkeit zu einer Gemeinschaft und der damit verbundenen Rechte in ihrem Bildungserfolg unterstützt. Das ökonomische Kapital bildet die Bedeutung von Kapital ab, z. B. das Einkommen. Beim kulturellen Kapital unterscheidet Bourdieu Bildung, die sich eine Person durch die Investition von Zeit, durch kulturelle Güter oder durch den Erwerb von Titeln erworben hat. Darüber hinaus werden bei der Konstruktion des Sozialindex Migrationsmerkmale der Schülerinnen und Schüler herangezogen, da der Migrationshintergrund (z. B. erfasst über die Familiensprache) noch immer einen bedeutsamen Zusammenhang mit schulischer Leistung aufweist (vgl. Schulte et al. 2014).

1.4 Fragestellung

Die bisherigen Darstellungen haben verdeutlicht, dass empirische Schulleistungstudien zwar regelmäßig das Vorhandensein von Kompositionseffekten aufzeigen, dass in diesen Studien jedoch oft keinerlei Angaben über die in Schule und Unterricht ablaufenden Prozesse erhoben werden. Schulinspektionen dagegen legen den Schwerpunkt ihrer Diagnose auf Prozessqualitäten. Daher soll in dieser Studie erstmals explorativ anhand von Inspektionsdaten die Frage des Zusammenhangs zwischen Kontextbedingungen, unter denen Schulen arbeiten, und den in ihnen ablaufenden operativen Prozessen untersucht werden. Nach Dittons Qualitätsmodell (s. Abbildung 1) betrifft dies somit den Zusammenhang zwischen Block 1: Bedingungen (speziell sozioökonomischer Hintergrund der Schülerschaft) und den Blöcken 3 (Prozesse auf Institutionsebene) und 4 (Prozesse auf Unterrichtsebene). Als Operationalisierung von Prozessqualität dienen die Qualitätsbewertungen der Schulinspektion sowie einzelne Informationsquellen wie Schülerwahrnehmungen von Schul- und Unterrichtsqualität. Schulische Kontextmerkmale werden über den in Hamburg genutzten Sozialindex operationalisiert. Im Gegensatz zur skizzierten Literatur wird der Zusammenhang zwischen Kontextmerkmalen und Prozessqualität aufgrund der generellen Anlage von Schulinspektion als einer Organisationsdiagnose ausschließlich auf der Schulebene nachgegangen. Aussagen zum Zusammenhang zwischen Kontextbedingungen und Unterrichtsqualität auf Klassenebene können aufgrund von Inspektionsdaten nicht analysiert werden, da diese nur auf Schulebene vorliegen.

2. Methode

2.1 Stichprobe

Grundlage der folgenden Analysen sind die Inspektionsdaten aller Schulen (N=375), die im Rahmen des ersten Inspektionszyklus im Zeitraum von 2006 bis 2013 in Hamburg durchgeführt wurden. Dies betrifft 169 Grundschulen, 72 Schulen mit mehreren Bildungsgängen (MBG), 67 Gymnasien, 46 berufliche Schulen sowie 21 Sonderschulen. Die durchschnittliche Anzahl an Schülerinnen und Schülern beträgt pro Schule 456 (Spannbreite von 32 bis 1.820), die mittlere Kollegiumsgröße 46 (Spannbreite von 6 bis 166).

2.2 Instrumente

Zentraler Bestandteil im Inspektionsverfahren ist die Einschätzung der Unterrichtsqualität durch die Beobachtung von Unterrichtssequenzen (vgl. Pietsch 2011). Hierfür führen zwei Inspektorinnen und Inspektoren, ein Schulformexperte/eine Schulformexpertin (Schulleitung derselben Schulform) sowie bei berufsbildenden Schulen ein Wirtschaftsvertreter/eine Wirtschaftsvertreterin zwischen 40 (an Grundschulen) und 100 (an berufsbildenden Schulen) Unterrichtsbeobachtungen im Rahmen des mehrtägigen Schulbesuchs durch. Dort bewertet das Team die Qualität einzelner, fachübergreifender Unterrichtsmerkmale anhand standardisierter Beobachtungsbögen. Zusätzlich werden alle Schülerinnen und Schüler ab der dritten Jahrgangsstufe, ihre Eltern sowie die Lehrkräfte der Schule mittels Online-Fragebögen zu bestimmten Qualitätsmerkmalen der Schule befragt. Weiterhin werden ausgewählte Vertreter/-innen dieser Akteursgruppen ebenso wie die Schulleitung und das technische und Verwaltungspersonal mittels teilstandardisierter Interviews um vertiefende Einsichten in das Schulgeschehen gebeten. Darüber hinaus soll die Einsichtnahme in Dokumente wie das Schulprogramm Auskunft über den bisher erreichten Stand der schulischen Qualitätsentwicklung geben.

Nach Abschluss des Schulbesuches werten die Inspektionsteams die erhobenen Daten aus und melden ihre Befunde der Einzelschule zurück. Diese Rückmeldungen beziehen sich auf die in Tabelle 2 dargestellten Qualitätsbereiche und sind jeweils vierfach gestuft: 1=schwach: Die Schule erfüllt keines oder nahezu keines der Teilkriterien des Qualitätsbereichs; 2=eher schwach als stark: Die Schule weist in diesem Qualitätsbereich mehr Schwächen als Stärken auf; 3=eher stark als schwach: Die Schule weist in diesem Qualitätsbereich mehr Stärken als Schwächen auf; 4=stark: Die Schule erfüllt alle oder nahezu alle Teilkriterien dieses Qualitätsbereichs optimal oder gut. Auf diese Bewertungen greift die folgende Sekundäranalyse zurück (vgl. Leist et al. 2009; Leist et al. 2011). Aus den Inspektionsberichten geht hervor,

welche Informationsquellen das Inspektionsteam zur Bewertung eines Qualitätsaspektes heranzieht und zusammenfasst. Diese Zusammenfassung beruht auf einer Einschätzung durch das Inspektionsteam und ist nicht durch einen Algorithmus vorgegeben (z. B. wie stark einzelne Informationsquellen gewichtet werden sollen).

Tabelle 2: Bewertete Qualitätsaspekte der Schulinspektion Hamburg

Qualitätsaspekt	<i>MW</i>	<i>SD</i>	Beispielinhalt
1.1 Führung wahrnehmen	3.06	.59	Die Schulleitung vertritt überzeugend eine pädagogische Vision.
1.2 Personal entwickeln	2.49	.60	Die Schulleitung führt regelmäßig Unterrichtsbesuche durch.
1.3 Finanz- und Sachmittel gezielt einsetzen	3.13	.59	Vorhandene Gelder werden sinnvoll eingesetzt.
1.4 Profil entwickeln und Rechenschaft ablegen	3.15	.67	Ergebnisse wichtiger Maßnahmen der Schulentwicklung werden systematisch dokumentiert und öffentlich gemacht.
2.1 Das schuleigene Curriculum gestalten	2.23	.65	Kollegiale oder fachliche Supervision findet systematisch statt
2.2 Unterrichten, Lernen, Erziehen	2.71	.52	Lernprozesse bauen sinnvoll aufeinander auf.
2.3 Organisatorische Rahmenbedingungen sichern	3.20	.50	Die Unterrichtsorganisation ermöglicht und fördert Kooperation im Kollegium.
2.4 Leistung beurteilen	2.31	.56	Leistungserwartungen, -kriterien und Verfahren der Leistungsbeurteilung sind Schülerinnen und Schülern sowie Eltern bekannt.
2.5 Prozesse und Ergebnisse evaluieren	2.13	.66	Die Entwicklungsschwerpunkte des Schulprogramms werden in regelmäßigen Zeitabständen evaluiert.
2.6 Förderkonzepte entwickeln	2.56	.58	Die Lehrkräfte setzen diagnostische Verfahren ein, um Lernvoraussetzungen und Lernprozesse der Schülerinnen und Schüler zu beurteilen.
2.7 Beratungsangebote gestalten	2.97	.65	Elternsprechtage und individuelle Elternberatungen werden regelmäßig und effizient durchgeführt.
2.8 Die Schulgemeinschaft beteiligen	2.83	.51	Schülerinnen und Schüler sowie Eltern gestalten das Schulleben aktiv mit.
3.1 Zufriedenes Personal, Schüler, Eltern, Betriebe	3.29	.53	Die Nachfrage von Schülerinnen, Schülern und Eltern zeugt von hoher Attraktivität der Schule.

Quelle: Behörde für Bildung und Sport Hamburg (2006)

Berechnungen von Mittelwert (*MW*) und Standardabweichung (*SD*) aufgrund der Daten der ersten Runde an Schulinspektionen (N=375 Schulen aller Schularten); Bewertungsskala 1 bis 4.

Des Weiteren werden im Folgenden die Schülerwahrnehmungen zur Schul- und Unterrichtsqualität separat analysiert. Die Schülerangaben haben hierbei den Vorteil, dass sie aufgrund der hohen Fallzahlen an Befragten sehr reliabel sind (vgl. Leist et al. 2009; Wurster & Gärtner 2013) und zudem die Analyse hinsichtlich einiger Subdimensionen zulassen (s. Tabelle 3).

Tabelle 3: Skalen des Schülerfragebogens zu Schul- und Unterrichtsqualität

Qualitätsaspekt	N _{items}	MW	SD	ICC2	Beispielitem
Skala Unterrichten, Lernen, Erziehen	12	2.88	.28	.98	Im Unterricht lernen wir, wie wir besser lernen können.
Skala Leistungen beurteilen	6	2.79	.34	.98	An unserer Schule sagen uns die Lehrerinnen und Lehrer, was man tun muss, um sich zu verbessern.
Skala Die Schulgemeinschaft beteiligen	3	2.69	.31	.97	An unserer Schule wird unsere Schülermeinung bei Entscheidungen berücksichtigt.
Gesamtskala Bildung und Erziehung	29	2.87	.31	.99	An unserer Schule haben wir Regeln vereinbart, wie wir miteinander umgehen wollen.
Skala Zufriedenes Personal, zufriedene Schülerinnen und Schüler, Eltern und Betriebe	8	3.05	.35	.98	Was ich an meiner Schule gut finde: dass Lehrerinnen und Lehrer und Schülerinnen und Schüler freundlich und respektvoll miteinander umgehen.

Quelle: Leist et al. 2009

Berechnung auf Basis der Schulwerte; ICC2=Intraklassenkorrelation als Angabe der Reliabilität des auf Schulebene aggregierten Wertes; Antwortskala: 1=trifft nicht zu bis 4=trifft voll und ganz zu.

Als Operationalisierung des sozioökonomischen Hintergrundes der Schülerschaft einer Schule wird der Hamburger Sozialindex genutzt. Der Sozialindex ist ein auf Schulebene verdichteter Hinweis auf Herkunftsmerkmale, die die Schülerinnen und Schüler einer Schule mitbringen. In Hamburg gibt es seit 1996 einen Sozialindex für Grundschulen und Schulen mit Sekundarstufe I. Dieser basiert u. a. auf der Kapitaltheorie von Bourdieu (1982, 1983) und erfasst ein Bündel soziokultureller Merkmale, die sich drei Dimensionen zuordnen lassen: soziales, ökonomisches und kulturelles Kapital.

In Form repräsentativer Datenerhebungen wird der Hamburger Sozialindex regelmäßig aktualisiert. Jede Schule wird einer von sechs Kategorien zugeordnet:

1. stark belastete soziale Lage der Schülerschaft
2. eher stark belastete soziale Lage der Schülerschaft
3. tendenziell belastete soziale Lage der Schülerschaft

4. tendenziell bevorzugte soziale Lage der Schülerschaft
5. eher bevorzugte soziale Lage der Schülerschaft
6. bevorzugte soziale Lage der Schülerschaft

2.3 Datenaufbereitung

Die folgenden Analysen beruhen auf Mittelwertsvergleichen der Qualitätsbewertungen zwischen Schulen mit unterschiedlichem Sozialindex. Da der Sozialindex für Sonder- und berufliche Schulen nicht vorliegt, beziehen sich die folgenden Ergebnisse ausschließlich auf Grundschulen, Schulen mit mehreren Bildungsgängen und Gymnasien.

3. Ergebnisse

Tabelle 4 stellt zunächst die Verteilung des Sozialindex über alle Schulen bzw. getrennt nach Schularten dar. Hierbei wird deutlich, dass sich insbesondere die Grundschulen stark hinsichtlich des sozioökonomischen Hintergrundes der Schülerschaft unterscheiden. Alle Kategorien des Sozialindex werden nahezu gleich häufig vergeben. Dies stellt sich in den Schulen mit Sekundarstufe I anders dar. Der Sozialindex der Schulen mit mehreren Bildungsgängen rangiert primär zwischen 2 und 4, während derjenige der Gymnasien zwischen 4 und 6 rangiert. Für die folgenden Analysen bedeutet die geringere Variabilität des Sozialindex in diesen Schularten, dass mögliche Zusammenhänge zu Prozessmerkmalen tendenziell geringer ausfallen.

Tabelle 4: Verteilung Sozialindex nach Schulform und Gesamt

Sozialindex	N _{Gesamt}	N _{Grundschule}	N _{MBG}	N _{Gymnasium}
1	32	28	4	0
2	55	29	23	3
3	49	20	24	5
4	66	33	12	21
5	51	29	1	21
6	38	23	2	13
Gesamt	291	162	66	63

MBG=Schule mit mehreren Bildungsgängen

Die Tabellen 5a und 5b stellen Mittelwerte und Standardabweichungen der Qualitätsbewertungen nach Schulart und Sozialindex dar. Zudem wird die Größe der Unterschiede zwischen den einzelnen Kategorien des Sozialindex je Schulform in Form der Effektgröße η^2 angegeben. Der Einteilung Cohens (1988) folgend, kann ein η^2 von .01 als kleiner Effekt, von 0.06 als mittlerer und von über .14 als großer Effekt gelten. Mehrheitlich bestehen bei den Qualitätsbewertungen nur geringe Unterschiede zwischen den Kategorien des Sozialindex (insgesamt 22 kleine Effekte, 13 mittlere und 3 große).

Tabelle 5a: Mittelwerte und Standardabweichungen der Gesamtbewertungen nach Schulart und Sozialindex

Schul- form	Sozial- index	1.1		1.2		1.3		1.4		2.1		2.2	
		<i>MW</i>	<i>SD</i>	<i>MW</i>	<i>SD</i>	<i>MW</i>	<i>SD</i>	<i>MW</i>	<i>SD</i>	<i>MW</i>	<i>SD</i>	<i>MW</i>	<i>SD</i>
G	1	3.1	.6	2.4	.5	3.2	.5	3.1	.7	2.1	.5	2.8	.5
	2	3.2	.6	2.6	.6	3.2	.6	3.3	.6	2.1	.6	3.0	.3
	3	3.1	.7	2.7	.7	3.2	.6	3.2	.6	2.2	.6	2.8	.6
	4	3.0	.5	2.4	.5	3.3	.5	3.2	.7	2.0	.7	2.8	.5
	5	3.0	.6	2.3	.5	3.2	.6	3.1	.6	2.0	.6	2.7	.5
	6	3.0	.6	2.6	.7	3.0	.4	3.0	.8	2.0	.4	2.8	.5
MBG	1	2.8	.5	2.5	.6	3.0	.8	2.8	.5	2.0	0	2.5	.6
	2	3.0	.7	2.5	.6	3.2	.6	3.0	.8	2.3	.6	2.7	.6
	3	2.9	.7	2.5	.7	2.9	.6	3.1	.8	2.5	.7	2.5	.6
	4	3.1	.7	2.8	.8	2.8	.4	3.1	.7	2.3	.7	2.6	.5
	5	3.0		2.0		3.0		4.0		2.0		3.0	
	6	3.0	0	2.0	0	3.5	.7	3.5	.7	2.0	0	2.0	0
GYM	1												
	2	3.0	0	2.0	0	2.7	.6	3.0	1.0	2.0	0	3.0	0
	3	3.4	.5	2.6	.5	3.2	.4	3.8	.4	3.0	.7	2.6	.5
	4	3.0	.5	2.5	.5	2.9	.7	3.3	.6	2.3	.7	2.6	.5
	5	3.1	.5	2.5	.5	2.8	.5	3.3	.5	2.4	.6	2.4	.5
	6	3.1	.6	2.5	.7	3.0	.6	3.5	.5	2.5	.5	2.5	.5
		<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>
G		.02	.77	.07	.06	.03	.42	.02	.79	.02	.72	.02	.58
MBG		.02	.96	.06	.63	.08	.40	.05	.70	.05	.66	.05	.67
GYM		.04	.64	.05	.60	.04	.63	.08	.31	.10	.18	.08	.27

G=Grundschule; GYM=Gymnasium; MBG=Schule mit mehreren Bildungsgängen; *eta*²>.10 fett

In jeder Schulart besteht ein großer Effekt des Sozialindex auf die Qualitätsbewertung: in der Grundschule und im Gymnasium auf *2.6 Förderkonzepte entwickeln* und in Schulen mit mehreren Bildungsgängen auf *2.4 Leistungen beurteilen*.

Die Ausprägung vorhandener Förderkonzepte für sowohl leistungsschwache als auch leistungsstarke Schülerinnen und Schüler hängt in Grundschulen und Gymnasien negativ mit dem Sozialindex zusammen. So ist die Bewertung von *2.6 Förderkonzepte entwickeln* in beiden Schularten in den Schulen, die in einer schwierigen sozialen Lage arbeiten, besser als in Schulen, deren Schülerinnen und Schüler aus einer privilegiierteren sozialen Lage kommen. In Grundschulen mit Sozialindex 1 und 2 rangiert die mittlere Qualitätsbewertung von 2.6 zwischen 2.9 und 3.0, während sie in Grundschulen mit Sozialindex 5 und 6 zwischen 2.4 und 2.5 verläuft.

Tabelle 5b: Mittelwerte und Standardabweichungen der Gesamtbewertungen nach Schulart und Sozialindex

Schulform	Sozialindex	2.3		2.4		2.5		2.6		2.7		2.8		3.1	
		MW	SD	MW	SD	MW	SD	MW	SD	MW	SD	MW	SD	MW	SD
G	1	3.4	.5	2.3	.5	2.1	.5	2.9	.4	3.1	.7	2.7	.6	3.5	.5
	2	3.4	.6	2.3	.6	2.3	.8	3.0	.5	2.9	.7	2.8	.6	3.6	.5
	3	3.2	.6	2.4	.7	2.3	.8	2.6	.5	2.8	.7	2.9	.5	3.4	.6
	4	3.2	.6	2.2	.5	1.9	.7	2.7	.5	2.8	.8	2.9	.5	3.5	.6
	5	3.2	.4	2.4	.6	2.0	.6	2.5	.5	2.7	.7	2.9	.5	3.5	.5
	6	3.1	.5	2.4	.5	2.2	.7	2.4	.5	2.7	.6	3.0	.4	3.4	.6
MBG	1	3.3	.5	1.8	.5	2.0	0	2.5	.6	3.3	.5	2.8	.5	3.0	0
	2	3.2	.6	2.2	.6	2.1	.8	2.4	.6	3.1	.5	2.7	.4	3.0	.5
	3	3.1	.4	2.7	.6	2.2	.7	2.6	.6	3.0	.6	3.0	.6	3.2	.5
	4	3.1	.3	2.4	.5	2.1	.5	2.4	.5	3.1	.5	2.8	.6	3.1	.3
	5	4.0		3.0		2.0		3.0		3.0		4.0		4.0	
	6	3.0	0	2.5	.7	1.0	0	2.0	0	2.0	0	3.0	0	3.0	0
GYM	1														
	2	3.0	0	2.3	.6	2.0	0	2.7	.6	3.0	0	2.7	.6	3.3	.6
	3	3.4	.5	2.2	.4	2.2	.4	3.0	0.0	3.6	.5	3.0	.7	3.2	.8
	4	3.2	.4	2.3	.5	2.2	.5	2.8	.4	3.1	.5	3.0	.4	3.1	.5
	5	3.0	.5	2.3	.7	2.3	.7	2.6	.5	3.3	.5	3.0	.3	3.0	.4
	6	3.2	.4	2.0	.4	2.3	.6	2.3	.5	3.2	.6	2.9	.3	3.2	.4
		<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>
	G	.04	.2	.02	.7	.04	.2	.15	<.01	.04	.3	.05	.2	.02	.8
	MBG	.07	.4	.21	.01	.09	.3	.07	.4	.12	.2	.11	.2	.09	.3
	GYM	.05	.5	.06	.4	.02	.8	.17	.03	.08	.3	.04	.7	.02	.9

G=Grundschule; GYM=Gymnasium; MBG=Schule mit mehreren Bildungsgängen; *eta*²>.10 fett

Ähnliche Unterschiede, in ihrem Ausmaß jedoch deutlich geringer, treten in der Grundschule bei folgenden Qualitätsaspekten auf: *2.3 organisatorische Rahmenbedingungen schaffen*; *2.7 Beratungsangebote gestalten* sowie *2.8 die Schulgemeinschaft beteiligen*. Meist sind diese Zusammenhänge negativ, d.h. je bevorzugter die soziale Lage der Schülerschaft einer Schule, umso schwächer fallen die Bewertungen der Schulinspektion aus. So wird die Schaffung organisationaler Rahmenbedingungen (2.3) schwächer bewertet, je stärker der sozioökonomische Hintergrund der Schülerschaft ist. Hierunter fallen z.B. Grundsätze der Unterrichtsverteilung und Vertretungsregelungen, eine Unterrichtsorganisation, die Kooperationen des Kollegiums sowie fächerübergreifendes, fächerverbindendes und projektorientiertes Lernen der Schülerinnen und Schüler zulässt. Ebenfalls werden die vorhandenen Beratungsangebote (2.7) in Grundschulen mit privilegierter Schülerschaft schwächer beurteilt als an Grundschulen mit niedrigem Sozialindex. Die Mitwirkung von Schülerinnen und Schülern sowie von Eltern am Schulleben (2.8) wird dagegen umso besser bewertet, je höher der Sozialindex der Schülerschaft.

An Gymnasien gibt es neben dem bedeutsamen Zusammenhang zwischen Sozialindex und Bewertung des Kriteriums *2.6 Förderkonzepte entwickeln* einen weiteren, der etwas geringer ausfällt und aufgrund der geringen Fallzahl nicht signifikant wird: Die *Unterrichtsbewertung* (2.2) fällt an Schulen mit privilegierterer Schülerschaft schwächer aus als an Schulen mit niedrigem Sozialindex.

Anders verhält sich die Situation bzgl. des dritten großen Effektes: die Beurteilung des Merkmals *2.4 Leistungen beurteilen* in Schulen mit mehreren Bildungsgängen. Hierunter fallen Aspekte wie die Transparenz der Leistungserwartungen und -kriterien gegenüber Schülerinnen, Schülern und Eltern, die Erörterung der Leistungsbeurteilung mit Schülerinnen und Schülern sowie die Auswertung von Leistungsergebnissen zur weiteren Unterrichtsentwicklung. Je privilegierter die Schülerschaft einer Schule mit mehreren Bildungsgängen, umso besser wird dieser Qualitätsaspekt beurteilt (von MW=1.8 bei Sozialindex 1 bis 2.5 bei Sozialindex 6). Etwas geringer und nicht signifikant fällt der Zusammenhang zwischen Sozialindex und dem vorhandenen *Beratungsangebot* (2.7) aus: hier werden jedoch wieder die Schulen mit privilegierter Schülerschaft etwas schwächer beurteilt als Schulen mit eher belasteter Schülerschaft. Insgesamt sind die Ergebnisse der Schulen mit mehreren Bildungsgängen vorsichtig zu interpretieren, da etliche Kategorien des Sozialindex nur wenige Schulen umfassen.

Tabelle 6 fokussiert im Folgenden auf eine Informationsquelle, welche Inspektorinnen und Inspektoren nutzen, um eine Gesamtbeurteilung zu erstellen: die Schülerbefragung. Die Tabelle stellt die Ergebnisse für die vorhandenen Skalen der Schülerbefragung dar. Die Größe der Standardabweichungen verdeutlicht, dass sich die mittleren Schülereinschätzungen zwischen den Schulen innerhalb einer Schulart nur sehr wenig unterscheiden. Ein Grund für diese geringe Variabilität dürfte in der großen Stichprobe pro Schule liegen. Die vorhandenen Unterschiede

können jedoch zu einem größeren Anteil durch den Sozialindex erklärt werden, als dies bei den Gesamtbewertungen der Fall war (bis zu 22 %). Mehrheitlich bestehen bei den Schülereinschätzungen große Unterschiede zwischen den Kategorien des Sozialindex (insgesamt 2 kleine, 4 mittlere und 9 große Effekte).

Tabelle 6: Mittelwerte und Standardabweichungen der Schüleraussagen nach Schulart und Sozialindex

Schul- form	Sozial- index	Skala Unterrich- ten, Lernen, Erziehen		Skala Leistungen beurteilen		Skala Schulge- meinschaft beteiligen		Gesamt- skala Bildung und Erziehung		Skala Zufrieden- heit	
		<i>MW</i>	<i>SD</i>	<i>MW</i>	<i>SD</i>	<i>MW</i>	<i>SD</i>	<i>MW</i>	<i>SD</i>	<i>MW</i>	<i>SD</i>
G	1	3.2	.2	3.1	.2	2.8	.3	3.2	.2	3.4	.2
	2	3.1	.2	3.1	.2	2.9	.3	3.2	.2	3.4	.2
	3	3.0	.2	2.9	.2	2.8	.3	3.0	.2	3.3	.2
	4	3.0	.2	2.9	.2	2.7	.3	3.0	.2	3.3	.2
	5	3.0	.2	2.9	.2	2.7	.3	3.0	.2	3.3	.2
	6	3.0	.2	2.9	.2	2.7	.3	3.0	.2	3.3	.2
MBG	1	2.7	.1	2.8	.2	2.5	.2	2.8	.2	2.8	.3
	2	2.7	.2	2.7	.2	2.6	.2	2.7	.2	2.8	.2
	3	2.6	.2	2.7	.2	2.6	.2	2.7	.2	2.8	.1
	4	2.6	.1	2.6	.1	2.6	.1	2.6	.1	2.8	.1
	5	2.6	.	2.6	.	2.8	.	2.7	.	3.2	.
	6	3.0	.2	2.8	.3	2.8	.1	3.0	.2	3.3	.2
GYM	1
	2	2.5	.	2.4	.0	2.4	.	2.5	.0	2.7	.2
	3	2.6	.1	2.5	.1	2.5	.2	2.6	.1	2.7	.3
	4	2.6	.1	2.5	.1	2.6	.1	2.6	.1	2.8	.1
	5	2.6	.1	2.5	.1	2.7	.1	2.6	.1	2.9	.1
	6	2.7	.1	2.5	.1	2.7	.1	2.6	.1	2.9	.1
		<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>	<i>eta</i> ²	<i>p</i>
G		.14	<.01	.18	<.01	.04	.4	.12	<.01	.04	.22
MBG		.22	.03	.13	.12	.09	.46	.18	.03	.22	<.01
GYM		.15	.09	.07	.41	.15	.08	.17	.03	.18	.02

G=Grundschule; GYM=Gymnasium; MBG=Schule mit mehreren Bildungsgängen; *eta*²>.10 fett

In den Grundschulen werden typische Merkmale guten Unterrichts (2.2 Unterrichten, Lernen und Erziehen), effektive Formen der Leistungserfassung und Rückmel-

dung (2.4 Leistungen beurteilen) sowie die Gesamtskala zu Bildung und Erziehung (2) umso schwächer beurteilt, je höher der Sozialindex der Schülerschaft ist.

Ein genau gegenteiliges Muster findet sich an Gymnasien, deren Schülerschaft größtenteils über einen starken sozioökonomischen Hintergrund verfügt (s. Tabelle 4). Hier finden sich durchweg positive Zusammenhänge zwischen Sozialindex und Schülerwahrnehmungen hinsichtlich typischer Merkmale guten Unterrichts (2.2 Unterrichten, Lernen, Erziehen), der Beteiligung am Schulleben (2.8 Schulgemeinschaft beteiligen), der Gesamtskala Bildung und Erziehung (2) sowie der Zufriedenheit mit der Schule (3.1 Zufriedenheit).

Schulen mit mehreren Bildungsgängen zeichnen sich nochmals durch ein anderes Muster aus. Hier sind die Unterschiede zwar größer, in ihrer Richtung jedoch nicht konsistent. Da in dieser Schulart etliche Kategorien des Sozialindex nur mit wenigen Schulen besetzt sind, können die gefundenen Unterschiede stark von den Ergebnissen einzelner Schulen beeinflusst sein.

4. Diskussion

Das vorliegende Kapitel behandelt anhand einer konkreten Fragestellung die Eignung von Schulinspektionsdaten für Untersuchungen der empirischen Bildungsforschung. Exemplarisch wird anhand von Daten der Hamburger Schulinspektion der Zusammenhang zwischen sozioökonomischem Status der Schülerschaft und Prozessqualitäten von Schulen untersucht.

Die Qualität der methodischen Zugangsweise der Hamburger Schulinspektion ist an etlichen Stellen dokumentiert (vgl. Pietsch 2010; Pietsch et al. 2009; Pietsch & Tosana 2008), sodass die Datenerfassung in diesem Artikel nur sehr überblickshaft dargestellt zu werden braucht.

Die Ergebnisse verdeutlichen, dass Inspektionsdaten Aussagen zur Art und zum Ausmaß des Zusammenhangs zwischen Schülerkomposition und innerschulischen Prozessen ermöglichen. Insbesondere erlaubt die breite Erfassung unterschiedlichster Qualitätsmerkmale von Schule durch die Schulinspektion die Exploration von Zusammenhängen, die bislang nicht im Rahmen der Forschung zu Kompositionseffekten untersucht wurden (vgl. Hallinger & Murphy 1986; Opdenakker & van Damme 2007).

4.1 Interpretation

Die Ergebnisse zum Zusammenhang von Sozialindex und Qualitätsbewertung im Rahmen der Schulinspektion zeigen, dass nur einige schulinterne Prozesse stärker mit dem sozioökonomischen Hintergrund der Schülerschaft assoziiert sind und

jede Schulart ein unterschiedliches Zusammenhangsmuster aufweist. Folgende Hypothesen werden für die gefundenen Zusammenhänge diskutiert:

Kompensatorische Maßnahmen: Die mehrheitlich negativen Zusammenhänge zwischen Sozialindex und Qualitätsbewertung können u. U. dadurch erklärt werden, dass es in Schulen mit einer sozial privilegierten Schülerschaft weniger Bedarf an kompensatorischen Maßnahmen gibt, die innerhalb des Inspektionsverfahrens erfasst und bewertet werden (Beratungsangebote an Schüler/-innen und Eltern, Förderangebote für Schüler/-innen, projektorientiertes Lernen).

Nähe zum Elternhaus: Die positiven Zusammenhänge zwischen Sozialindex und der Beteiligung von Schülerinnen und Schülern sowie Eltern am Schulleben stimmt mit den Erkenntnissen von Hallinger und Murphy (1986) überein, dass die Nähe zum Elternhaus bei Schulen mit sozial privilegierter Schülerschaft größer ist.

Geringe Fallzahlen: In Schulen mit mehreren Bildungsgängen und auch an Gymnasien treten z. T. zwar große Effekte auf; hier ist jedoch zu bedenken, dass in diesen Schularten manche Kategorien des Sozialindex nur mit wenigen Schulen besetzt sind. Daher können die gefundenen Unterschiede in diesen Schularten stark von den Ergebnissen einzelner Schulen beeinflusst sein.

Reliabilität der Messung: Die Analyse von Sozialindex und Schülerwahrnehmungen verdeutlicht, dass die Sozialindex-bedingten Unterschiede in den einzelnen Informationsquellen (hier Schülerwahrnehmungen) stärker ausfallen als Unterschiede in den Gesamtbewertungen. Dies erklärt sich zum einen durch die hohe Reliabilität der Schülerangaben, zum anderen dadurch, dass in die Gesamtbewertungen mehrere Informationsquellen eingehen (Beobachtungen, Fragebogen, Interviews, Dokumentenanalyse) und diese anschließend auf eine grobe vierstufige Bewertungsklassifikation vereinfacht werden.

Schulartspezifische Verteilung des Sozialindex: Bei der Interpretation der Zusammenhänge zwischen Schülerwahrnehmungen und Sozialindex werden schulartspezifische Muster deutlich. Die positiven Zusammenhänge zwischen Sozialindex und Schülerwahrnehmungen am Gymnasium sprechen für die Wirkannahmen 1 und 2, d. h. in leistungsstarken Schulen unterrichten Lehrkräfte möglicherweise anspruchsvoller und Störungen treten vermutlich seltener auf. Die allesamt negativen Zusammenhänge zwischen Sozialindex und Schülerwahrnehmungen in Grundschulen lassen sich möglicherweise auch mit kompensatorischen Maßnahmen der Schule erklären (zusätzliche Angebote im Unterricht, zusätzliche Förder- und Beratungsangebote), die von den Schülerinnen und Schülern wahrgenommen werden.

4.2 Bedeutung von Inspektionsdaten für die empirische Bildungsforschung

Die Analysen demonstrieren grundsätzlich die Nützlichkeit von Inspektionsdaten, sekundäranalytisch zur Beantwortung von Forschungsfragen der empirischen Bildungsforschung beizutragen (vgl. Wurster & Gärtner 2013). Das vorliegende Beispiel hat insbesondere das Potenzial der *einzelnen* Informationsquellen (hier Schülerwahrnehmungen) verdeutlicht. Fragebogen und Unterrichtsbeobachtungen der Hamburger Schulinspektion sind theoretisch fundiert und methodisch gut abgesichert. So liefern die einzelnen Zugänge aussagekräftige Einschätzungen von Schul- und Unterrichtsqualität aus Sicht von Eltern, Schülerinnen und Schülern, Lehrkräften und externen Beobachtern. Die vorliegende Analyse lässt sich ohne Weiteres auf die Beurteilung des Schulmanagements durch Lehrkräfte oder auf die Wahrnehmung der Schulqualität durch Eltern übertragen, um so umfangreichere Aussagen zu den Merkmalsbereichen 3 (Institutionsebene) und 4 (Interaktionsebene) des Modells von Dittton (2007) zu gewinnen. Für die empirische Bildungsforschung ist es von hohem Interesse, diese Informationsquellen im Rahmen von Sekundäranalysen zu nutzen (vgl. Wurster & Gärtner 2013).

4.3 Adjustierung von Inspektionsergebnissen

Ob Ergebnisse von Schulinspektionen, ähnlich wie Schulleistungsdaten, aufgrund von Kontextmerkmalen adjustiert werden sollen, kann nicht abschließend entschieden werden (vgl. Wendt & Bos 2011). Dazu sind die Ergebnisse zwischen den Schularten zu inkonsistent; teilweise liegen nur geringe Fallzahlen vor und müssen anhand weiterer Untersuchungen überprüft werden. Da die vorliegenden Ergebnisse nur wenige bedeutsame Zusammenhänge zwischen Sozialindex und den Gesamtbewertungen der Schulinspektion aufweisen, scheinen zumindest im aktuellen Beurteilungsprozess keine großen Benachteiligungen von Schulen mit hohem oder niedrigem Sozialindex aufzutreten. Im Gegensatz zur typischen Ergebnislage in Schulleistungsstudien, in denen sozial benachteiligte Schülergruppen tendenziell schwächer abschneiden als Schülerinnen und Schüler aus privilegierteren Elternhäusern, ist der Zusammenhang bezogen auf die Bewertungen der Schulinspektion tendenziell gegensätzlich. Schulen, die mit einer sozial eher benachteiligten Schülerschaft arbeiten, werden in Bezug auf einige Kriterien besser beurteilt als Schulen mit einer sehr privilegierten Schülerschaft. Dies hängt vermutlich mit verstärkten kompensatorischen Maßnahmen an Schulen in sozial benachteiligter Lage zusammen.

Die Frage, ob die Bewertung einer Schule erst dann angemessen erfolgen kann, wenn die Rahmenbedingungen (Umfeld der Schule, Ausstattung, Schülerkomposition ...) mit erfasst werden und die Schulinspektion ihren Diagnoseauftrag diesbe-

züglich erweitert, sollte jedoch auf Grundlage zukünftiger Untersuchungen weiter diskutiert werden. Weitere Untersuchungen zu dieser Fragestellung können vorhandene Datengrundlagen von Inspektoraten detaillierter analysieren, z. B. die Frage der Messinvarianz der eingesetzten Skalen in den unterschiedlichen Schularten, um die Voraussetzungen für anschließende Mittelwertsvergleiche besser abschätzen zu können, oder die Frage der Reliabilität der vorhandenen Informationsquellen. Daten aus anderen Bundesländern können möglicherweise in den einzelnen Untergruppen von Schulart und Sozialindex eine genügend große Fallzahl realisieren, sodass die hier gefundenen Unterschiede belastbarer repliziert werden können.

Literatur

- Arnold, K.-H. (1999). *Fairneß bei Schulsystemvergleichen: Diagnostische Konsequenzen von Schulleistungsstudien für die unterrichtliche Leistungsbewertung und binnenschulische Evaluation*. Münster: Waxmann.
- Baumert, J., Stanat, P. & Watermann, R. (2006). *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bellin, N. (2009). *Klassenkomposition, Migrationshintergrund und Leistung. Mehrebenenanalysen zum Sprach- und Leseverständnis von Grundschulern*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bos, W., Gröhlich, C., Guill, K., Ivanov, S., May, P., Nikolova, R., Scharenberg, K. & Wendt, H. (2010). *KESS 8. Kompetenzen und Einstellungen von Schülerinnen und Schülern – Jahrgangsstufe 8*. Hamburg: Behörde für Schule und Berufsbildung.
- Bourdieu, P. (1982). *Die feinen Unterschiede – Kritik der gesellschaftlichen Urteilskraft*. Frankfurt a. M.: Suhrkamp.
- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (Soziale Welt: Sonderband 2, S. 183–198). Göttingen: Schwartz & Co.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Erlbaum.
- Coleman, J. S. (1988). *Social capital in the creation of human capital*. *American Journal of Sociology*, 94 (1), 95–120.
- Creemers, B. P. M. & Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness*. New York: Routledge.
- Ditton, H. (2007). Schulqualität – Modelle zwischen Konstruktion, empirischen Befunden und Implementierung. In J. van Buer & C. Wagner (Hrsg.), *Qualität von Schule. Ein kritisches Handbuch* (S. 83–92). Frankfurt a. M.: Lang.
- Döbert, H. & Dederig, K. (Hrsg.). (2008). *Externe Evaluation von Schulen. Historische, rechtliche und vergleichende Aspekte*. Münster: Waxmann.

- Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. *Internationale und nationale Befunde. Psychologie in Erziehung und Unterricht*, 60, 163–183.
- Hallinger, P. & Heck, R. H. (1996). Exploring the principal's contribution to school effectiveness: 1980–1995. *School Effectiveness and School Improvement*, 9 (2), 157–191.
- Hallinger, P. & Murphy, J. F. (1986). The social context of effective schools. *American Journal of Education*, 94 (3), 328–355.
- Harker, R. & Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement*, 15 (2), 177–199.
- Harris, D. N. (2010). How do school peers influence student educational outcomes? Theory and evidence from economics and other social sciences. *Teachers College Record*, 112 (4), 1163–1197.
- IQ Hessen (2012). *Bilanzbericht der Schulinspektion*. Wiesbaden: Institut für Qualitätsentwicklung.
- ISB (2012). *Bildungsbericht Bayern 2012*. München: Staatsinstitut für Schulqualität und Bildungsforschung.
- Leist, S., Pietsch, M., Tosana, S. & Bardowiecks, S. (2009). *Item- und Skalenhandbuch der Schulinspektion Hamburg*. Hamburg: Behörde für Schule und Berufsbildung, IfB.
- Leist, S., Töpfer, T., Bardowiecks, S., Pietsch, M. & Tosana, S. (2011). *Handbuch zum Unterrichtsbeobachtungsbogen der Schulinspektion Hamburg*. Hamburg: Institut für Bildungsmonitoring.
- Opdenakker, M.-C. & van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal*, 33 (2), 179–206.
- Pietsch, M. (2010). Evaluation von Unterrichtsstandards. *Zeitschrift für Erziehungswissenschaft*, 13 (1), 121–148.
- Pietsch, M. (2011). *Nutzung und Nützlichkeit der Schulinspektion Hamburg. Ergebnisse der Hamburger Schulleitungsbefragung*. Hamburg: Behörde für Schule und Berufsbildung, Institut für Bildungsmonitoring.
- Pietsch, M. & Tosana, S. (2008). Beurteilereffekte bei der Messung von Unterrichtsqualität: Das Multifacetten-Rasch-Modell und die Generalisierbarkeitstheorie als Methoden in der externen Evaluation von Schulen. *Zeitschrift für Erziehungswissenschaft*, 11 (3), 430–452.
- Pietsch, M., Schnack, J. & Schulze, P. (2009). Unterricht zielgerichtet entwickeln. Die Hamburger Schulinspektion entwickelt ein Stufenmodell für die Qualität von Unterricht. *Pädagogik*, 61 (2), 38–43.
- Scafidi, B., Sjoquist, D. L. & Stinebrickner, T. R. (2007). Race, poverty, and teacher mobility. *Economics of Education Review*, 26, 145–159.
- Scheerens, J. (2000). *Improving school effectiveness (Vol. 68)*. Paris: International Institute for Educational Planning, UNESCO.

- Scheerens, J. & Bosker, R. J. (1997). *The Foundations of Educational Effectiveness*. Oxford: Pergamon.
- Schulte, K., Hartig, J. & Pietsch, M. (2014). Der Sozialindex für Hamburger Schulen. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine daten- und theoriegestützte Schulentwicklung* (S. 67–80). Münster: Waxmann.
- Stamm, M. (2008). *Die Psychologie des Schuleschwänzens*. Bern: Huber.
- van Ewijk, R. & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, 5 (1), 134–150.
- Wendt, H. & Bos, W. (2011). Indikatoren zur Kontextuierung von Inspektionsergebnissen. Bedeutung und Anforderungen. In S. Müller, M. Pietsch & W. Bos (Hrsg.), *Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht* (S. 217–236). Münster: Waxmann.
- Wurster, S. & Gärtner, H. (2013). Erfassung von Bildungsprozessen im Rahmen von Schulinspektion und deren potenzieller Nutzen für die empirische Bildungsforschung. *Unterrichtswissenschaft*, 41 (3), 217–235.

Die Qualität des Unterrichts an SINUS-Grundschulen

Eine Programmevaluation mithilfe von Daten der Schulinspektion Hamburg

Marcus Pietsch & Timo Ehmke

Zusammenfassung

Der folgende Beitrag zeigt auf, wie sich Daten aus Schulinspektionsverfahren gewinnbringend für die Evaluation von Projekten und Programmen im Bildungsbereich nutzen lassen. Mithilfe eines Propensity Score Matching wird untersucht, inwieweit das weitverbreitete Unterrichtsentwicklungsprojekt SINUS einen nachweisbaren Einfluss auf die Qualität des Unterrichts an Grundschulen in Hamburg hat. Die Analysen zeigen, dass sich SINUS-Schulen in der Bewertung durch die Inspektion ausschließlich im Qualitätsbereich „Unterrichten, Lernen, Erziehen“ von vergleichbaren Schulen unterscheiden. Weiterführende Analysen von Unterrichtsbeobachtungsdaten verdeutlichen, dass die Qualität des Unterrichts an SINUS-Grundschulen rund 16 Prozent höher ausfällt als an Schulen, die unter vergleichbaren Bedingungen arbeiten. Auch werden innerschulische Qualitätsunterschiede in Folge des Programms reduziert und Schwächen im Unterricht großflächig abgestellt. Abschließend wird die praktische Bedeutung dieser Befunde für die Schul- und Unterrichtsentwicklung diskutiert.

1. Einführung

Das Schul- und Unterrichtsentwicklungsprojekt SINUS gehört zu den erfolgreichsten und weitverbreitetsten Schulentwicklungsprogrammen der letzten Jahrzehnte in Deutschland (vgl. Prenzel et al. 2009). Das Akronym SINUS steht dabei für die Abkürzung „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“. Das Programm wurde 1998 in einer länderübergreifenden Initiative als ein Modellversuchsprogramm der Bund-Länder-Kommission (BLK) konzipiert und mit 180 Schulen der Sekundarstufe I in 15 Ländern gestartet (vgl. Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung 1997). Ziel des Programms ist es, den mathematisch-naturwissenschaftlichen Unterricht weiterzuentwickeln und die Arbeit von Lehrkräften weiter zu professionalisieren. Mit dem Nachfolgeprogramm „SINUS-Transfer“ fand ab 2003 eine beachtliche Dissemination des Programms statt: Rund 1.800 Schulen und ca. 6.000 Lehrkräfte

waren daran beteiligt (vgl. Fischer et al. 2013). Ab 2004 wurde der SINUS-Ansatz auch auf die Grundschule übertragen (vgl. Fischer et. al 2009). In zehn Modulen (siehe Abschnitt 2.1) arbeiteten Lehrkräfte von ca. 400 Grundschulen aus zehn Bundesländern daran, den Mathematikunterricht und den naturwissenschaftlichen Sachunterricht in der Primarstufe weiterzuentwickeln. In einer Erweiterungsphase ab 2009 wurden die Module noch einmal neu unter bestimmten thematischen Gesichtspunkten zusammengefasst und das Programm unter der Bezeichnung „SINUS an Grundschulen“ weitergeführt (vgl. Fischer et al. 2013). Das SINUS-Programm in der Sekundarstufe I wurde ab August 2007 nicht mehr zentral koordiniert, sondern durch die einzelnen Länder eigenverantwortlich weiter koordiniert. Ab 2013 gilt dies auch für das Programm „SINUS an Grundschulen“.

Um die Qualität der Arbeit in dem Programm zu sichern, wurden von Anfang an eine wissenschaftliche Begleitforschung installiert und zahlreiche Einzelstudien durchgeführt (vgl. Abschnitt 2.2). Evaluationsstudien im Sinne der Schuleffektivitätsforschung, bei denen weitere Vergleichsgruppen, Schülerleistungen und -einstellungen oder Unterrichtsmerkmale als Erfolgskriterien herangezogen werden, gab es allerdings bislang nur wenige. Hier setzt diese Studie an und nutzt die Daten der Hamburger Schulinspektion, um die Effektivität des Schulentwicklungsprogramms SINUS an Hamburger Schulen im Hinblick auf die Unterrichtsqualität abzuschätzen.

In diesem Beitrag wird zunächst der konzeptionelle Ansatz des SINUS-Projektes dargestellt, bevor dann Forschungsansätze und bisherige Befunde zur Effektivität des SINUS-Programms zusammengefasst und die Fragestellungen für diese Studie abgeleitet werden. Die Ergebnisse unserer Analysen werden dann für Beobachtungsmerkmale im Bereich Unterricht dargestellt und im Hinblick auf ihre Bedeutung für Schulentwicklungsprogramme und mögliche Konsequenzen diskutiert.

2. Theoretischer Hintergrund

2.1 SINUS als Schul- und Unterrichtsentwicklungsprojekt

Das BLK-Programm SINUS zielt darauf ab, den Unterricht im Hinblick auf konkrete Problemstellungen weiterzuentwickeln und zu verbessern, etwa im Hinblick auf die Aufgabenkultur oder das naturwissenschaftliche Experimentieren im Unterricht. Eine besondere Rolle spielt dabei die Kooperation zwischen den Lehrkräften: In Schulverbünden, den sogenannten Sets, arbeiten Lehrerinnen und Lehrer unter wissenschaftlicher Begleitung an vorgegebenen Themenbereichen (den Modulen). Reflexion und Evaluation des eigenen Unterrichts sind dabei zentrale Elemente. Die Schulsets werden von Koordinatorinnen und Koordinatoren betreut, die eng

auf Länder- und Bundesebene zusammenarbeiteten (vgl. Prenzel et al. 2009). Ausgangspunkt der SINUS-Arbeit in der Grundschule waren zehn empirisch ermittelte Module, die sich auf typische Herausforderungen des Mathematikunterrichts und naturwissenschaftlichen Sachunterrichts in der Grundschule beziehen. Es geht dabei um folgende Schwerpunkte (vgl. Demuth et al. 2011):

- „Gute Aufgaben“,
- „Entdecken, Erforschen, Erklären“,
- „Schülervorstellungen aufgreifen, grundlegende Ideen entwickeln“,
- „Lernschwierigkeiten erkennen – verständnisvolles Lernen fördern“,
- „Talente entdecken und unterstützen“,
- „Fächerübergreifend und fächerverbindend unterrichten“,
- „Interessen von Mädchen und Jungen aufgreifen und weiterentwickeln“,
- „Eigenständig lernen – Gemeinsam lernen“,
- „Lernen begleiten – Lernergebnisse beurteilen“ und
- „Übergänge gestalten“.

Diese Module werden im Fortsetzungsprojekt „SINUS an Grundschulen“ neu zusammengefasst und auf aktuelle Fragestellungen und Herausforderungen fokussiert (vgl. Fischer et al. 2011). Daraus ergeben sich fünf Arbeitsschwerpunkte:

- „Daten aus Unterrichtsevaluationen auswerten und für die Unterrichtsentwicklung nutzen“,
- „Bildungsstandards umsetzen (in Mathematik)“,
- „Wege zu den Naturwissenschaften eröffnen (im Sachunterricht)“,
- „Kinder mit besonderen Schwierigkeiten oder besonderen Begabungen unterstützen“ und
- „Übergänge gestalten“.

2.2 Forschungsansätze und Befunde zur Effektivität des SINUS-Programms

Bei Unterrichtsentwicklungsprogrammen wie SINUS sollen sich die Wirkungen auf der Ebene von Schulen und Lehrpersonen in einem veränderten Unterricht und in einer positiveren Kompetenzentwicklung der Lernenden zeigen (vgl. Fischer et al. 2013). Internationale Forschungen zeigen, dass wissenschaftsbasierte Schulentwicklungsmaßnahmen diesbezüglich kurzfristig kleine, aber dennoch beobachtbare Effekte nach sich ziehen, die sich im Zeitverlauf deutlich stabilisieren und mit zunehmender Programmdauer weiter anwachsen (vgl. Bormann et al. 2003). Deutschsprachige Forschungsergebnisse hierzu liegen bislang nur vereinzelt vor

und zu dem Programm SINUS existieren bisher erst zwei Studien, die die Ergebnisse von SINUS-Schulen hinsichtlich der fachlichen Leistung von Schülerinnen und Schülern und mit Blick auf Merkmale des Unterrichts mit anderen Schulen vergleichen, die nicht am SINUS-Programm teilgenommen haben.

In der ersten Studie wurden im Rahmen einer summativen Evaluation der ersten SINUS-Welle die Wirkungen des Programms untersucht, indem die Test- und Fragebogenergebnisse von Schülerinnen und Schülern an SINUS-Schulen mit denen einer repräsentativen Schulstichprobe aus PISA 2003 verglichen wurden (vgl. Prenzel et al. 2005). Die Ergebnisse des Vergleichs zwischen SINUS-Schulen und PISA-Schulen, die nicht an dem Programm teilgenommen haben, zeigten, dass SINUS im Verlauf der ersten Programmzeit auf allen untersuchten Ebenen Wirkung entfaltet hat. Dies betraf die erfolgreiche Umsetzung der Projekthalte auf Seiten der Lehrkräfte und die positive Wahrnehmung des Unterrichts auf Seiten der Schülerschaft. Hinsichtlich des Ziels, die Kompetenzen sowie die Interessen und Haltungen der Schülerinnen und Schüler in Mathematik und den Naturwissenschaften zu fördern, zeigten sich klare Vorteile zugunsten von SINUS bei Schulen mit mehreren Bildungsgängen und Integrierten Gesamtschulen. Hier berichteten die SINUS-Schülerinnen und -Schüler über ein höheres Interesse und mehr Freude an Mathematik.

Die Studie von Dalehefte et al. (2014) untersuchte die Wirksamkeit des Programms SINUS an Grundschulen und analysierte, inwieweit sich die mathematischen Kompetenzen von Schülerinnen und Schülern sowie das Weiterbildungs- und Kooperationsverhalten von Mathematiklehrkräften an SINUS-Grundschulen von den Ergebnissen der für Deutschland repräsentativen „Trends in International Mathematics and Science Study“ (TIMSS) 2011 unterscheiden. Der Vergleich ergab, dass Mathematiklehrkräfte an SINUS-Grundschulen in allen betrachteten Bereichen häufiger an Fortbildungen teilnehmen als Lehrkräfte, die nicht an SINUS-Schulen unterrichten. Bei Vergleich der mathematischen Kompetenzen von Schülerinnen und Schülern an SINUS-Schulen vs. anderen Schulen, ergab sich eine Effektstärke von $d = .18$ zugunsten des SINUS-Modellprogramms. Diese entspricht in dieser Altersstufe etwa dem Lernzuwachs von einem Drittel eines Schuljahres (vgl. Hill et al. 2008). Ein Vergleich der Perzentilbänder der mathematischen Kompetenz zwischen den beiden Gruppen ergab, dass von SINUS an Grundschulen insbesondere die schwächeren Schülerinnen und Schüler profitieren.

Die bisherige SINUS-Begleitforschung hat darüber hinaus unter anderem die Rolle der Schulleitung (vgl. Fischer 2005) und Aspekte der Kooperation von Lehrkräften (vgl. Fischer & Prenzel 2007; Ostermeier 2005) und ihre Arbeit mit dem Fachgruppen-Portfolio und anderen Dokumentationsmethoden untersucht (vgl. Stadler & Meentzen 2010; Meentzen 2009; Trepke et al. 2013). Die bislang vorliegenden Befunde deuten dabei darauf hin, dass sich SINUS-Schulen von anderen Schulen im Bezug auf effektivitätsgenerierende Merkmale vor allem mit Blick auf

den Unterricht unterscheiden könnten. So berichten Dalehefte et al. (2014), dass sich SINUS-Lehrkräfte zwar hinsichtlich unterrichtsrelevanter Fortbildungen, an denen sie teilnehmen, deutlich von anderen Lehrkräften unterscheiden, sich ihr Kooperationsverhalten vor Ort jedoch nicht von dem in Kollegien unterscheidet, deren Schulen nicht am Programm teilnehmen. Gleichwohl konnte in dieser Studie nur das Fortbildungsverhalten untersucht werden. Ob dieses jedoch mit einer veränderten Unterrichtspraxis einhergeht, ist derzeit unklar.

Zusammenfassend kann festgehalten werden, dass die genannten Studien durchaus Evidenz für die Wirksamkeit des SINUS-Programms belegen. Eine Studie, die jedoch gezielt den Nachweis von kausalen Effekten von SINUS-Schulen im Vergleich zu Nicht-SINUS-Schulen durchführt und damit mit einem Vergleichsgruppen-Design im Sinne der Kausalitätsforschung (vgl. Morgan & Winship 2007) arbeitet, steht jedoch noch aus.

3. Forschungsfragen und Hypothesen

Mit der aktuellen Studie soll die Wirksamkeit von „SINUS an Grundschulen“ in Hamburg untersucht werden, indem mittels der Methode des *Propensity Score Matchings* eine Vergleichsgruppe von Schulen gebildet wird, die nicht an dem Schulentwicklungsprogramm teilgenommen hat und als post-hoc generierte Kontrollgruppe für Vergleiche herangezogen werden kann. Konkret wird folgenden zwei Fragestellungen nachgegangen:

- Inwieweit unterscheiden sich SINUS-Grundschulen und Grundschulen, die nicht am SINUS-Programm teilgenommen haben, hinsichtlich unterrichtlicher, erzieherischer und organisatorischer Prozessmerkmale?
- Lassen sich zwischen SINUS-Grundschulen und Grundschulen, die nicht am SINUS-Programm teilgenommen haben, Unterschiede in der Unterrichtsqualität feststellen?

Da das Schulentwicklungsprogramm SINUS an Grundschulen explizit auf die Verbesserung des mathematisch-naturwissenschaftlichen Unterrichts abzielt, sind hier tendenziell positive Effekte zu erwarten. Die bisherigen Ergebnisse der SINUS-Begleitforschung belegen – wie gezeigt – positive Effekte auf Schülerleistungen, die den Rückschluss auf eine effektivere Unterrichtspraxis an SINUS-Schulen nahelegen. Allerdings weiß man aus der Schuleffektivitätsforschung, dass hinsichtlich der Veränderung von Unterrichtsmerkmalen durch Schulentwicklungsprogramme eher kleine Effektstärken zu erwarten sind (vgl. Seidel & Shavelson 2007).

4. Methodisches Vorgehen

Um die Wirksamkeit des Projektes SINUS auf schulische Prozessmerkmale an Hamburger Schulen zu evaluieren, werden im Folgenden Schulen, die an dem Projekt teilnahmen, mit solchen verglichen, die unter ähnlichen Bedingungen arbeiten, jedoch nicht an SINUS teilgenommen haben. Zum Einsatz kommt hier ein *Propensity Score Matching*. Das heißt, es werden kausale Effekte nach der Kontrolle von Variablen, die sowohl die abhängige Variable als auch den Treatmentstatus beeinflussen, geschätzt. Hierfür werden in einem ersten Schritt Paare von beobachteten Einheiten gebildet, die sich hinsichtlich der kontrollierten Variablen möglichst ähnlich sind, aber unterschiedlichen Treatmentgruppen angehören. Diese gepaarten Einheiten werden dann im zweiten Schritt zur Schätzung der Wirkung des Treatments verwendet. So werden nur Einheiten miteinander verglichen, die sich zwar im Treatmentstatus unterscheiden, aber in Hinblick auf potenziell beeinflussende Variablen möglichst ähnlich sind. Als abhängige Variable wird die Qualität des Unterrichts genutzt, die im Rahmen der Schulinspektion Hamburg mittels 20-minütiger Beobachtungen erfasst wird.

4.1 Datengrundlage

4.1.1 *Messung von Unterrichtsqualität im Rahmen der Schulinspektion Hamburg*

Die Daten zur Messung von Unterrichtsqualität wurden mittels Beobachtung von Unterrichtssequenzen erhoben. Für diese Unterrichtsbeobachtungen standen in der Regel zwei bis drei Tagen pro Schule und je Sequenz 20 Minuten Beobachtungszeit zur Verfügung. Die Auswahl der zu besuchenden Unterrichtssequenzen erfolgte jeweils vor dem Schulbesuch in Form einer stratifizierten Zufallsstichprobe, wobei Beobachtungssequenzen je Schulstunde gezogen wurden. Um dieses Vorgehen zu ermöglichen, reichten die Schulen kurz vor dem Schulbesuch ihre Stundenpläne bei der Schulinspektion Hamburg ein. Aus den vorliegenden Zeiteinheiten – von der ersten bis letzten Schulstunde – wurden dann für je zwei Unterrichtsstunden dreimal so viele Sequenzen zufällig gezogen, wie Inspektorinnen und Inspektoren am Schulbesuch beteiligt waren (in der Regel drei), sodass jeder Schulinspektor resp. jede Schulinspektorin innerhalb von zwei Schulstunden drei Sequenzen beobachtete. Die Unterrichtsstunden wurden hierfür in einen ersten und einen zweiten Teil eingeteilt, sodass die Mitarbeiterinnen und Mitarbeiter der Schulinspektion Hamburg entweder zu Beginn oder zum Abschluss (bei Doppelstunden ggf. auch in der Mitte) einer Unterrichtsstunde Informationen zur Unterrichtsqualität erfassten. Im Rahmen der Stichprobenziehung und Berichtlegung geht die Schulinspektion

Hamburg davon aus, dass die Qualität von Unterricht nicht ausschließlich von der Lehrperson abhängig ist, sondern als Angebot-Nutzungs-Beziehung definiert werden muss, die durch Lehrkräfte und Schülerschaft koproduziert wird. Dies bedeutet für die Stichprobenziehung praktisch, dass Lehrkräfte im Rahmen einer Schulinspektion nicht als stichprobenrelevantes Stratum betrachtet werden, da einzelne Lehr-Lern-Settings beobachtet werden, die eine jeweils spezifische Interaktion aus Lehrkraft, Schülerschaft, Fach, Lage der Unterrichtssequenz im Tagesablauf und weiteren Merkmalen darstellen. Entsprechend können einzelne Lehrkräfte auch häufiger, jedoch in verschiedenen Kontexten und von verschiedenen Inspektionsmitgliedern gesehen werden. Die Zuweisung der Beobachter zu den zu beobachtenden Unterrichtssequenzen erfolgte ebenso zufällig wie die Ziehung der Unterrichtssequenzen. Rund 10 Prozent aller Sequenzen, die je Schule beobachtet wurden, wurden darüber hinaus aus Gründen der Qualitätssicherung durch zwei Inspektionsmitglieder gleichzeitig eingesehen. An zu inspizierenden Grundschulen werden in der Regel mindestens 40 Sequenzen betrachtet, um zu gewährleisten, dass die Aussagen zur Qualität des Unterrichts einer Schulwoche auf Ebene der Einzelschule mit 95-prozentiger Wahrscheinlichkeit bei einer ungefähr 90-prozentigen Präzision ermöglicht werden.

Zur Bestimmung der Unterrichtsqualität werden die einzelnen Sequenzen von den Inspektorinnen und Inspektoren jeweils anhand eines Bewertungsbogens beurteilt, der 30 Kriterien umfasst. Die Kriterien dienen als Indikatoren für Qualitätsmerkmale des Hamburger Orientierungsrahmen Schulqualität (Behörde für Bildung und Sport 2006) und orientieren sich primär an den Merkmalen guten Unterrichts nach Helmke (2006). Die eingesetzten Items sollen die Messung von Unterrichtsgelingensbedingungen auf Basis einer Angebot-Nutzungs-Beziehung ermöglichen. Dabei decken die eingesetzten Items, wie Stralla (2009) im Rahmen einer vergleichenden Untersuchung aufzeigt, die national und international gängigen Kriterienkataloge zur Qualität von Unterricht differenziert ab. Die 30 Items wurden auf einer vierstufigen Ratingskala (Skalenniveau: ‚trifft nicht zu‘ bis ‚trifft zu‘) bewertet, wobei eine fünfte Kategorie markiert werden konnte, sofern die Unterrichtsbeobachter ein Merkmal für ‚nicht beobachtbar‘ hielten.

4.1.2 Umgang mit Daten zur Unterrichtsqualität der Schulinspektion Hamburg

Die Schulinspektion Hamburg meldet sowohl die Summenscores einzelner Items als auch den Wert einer Gesamtskala „Unterrichtsqualität“ an die inspizierten Schulen zurück (vgl. Pietsch et al. 2011). Verschiedene Analysen von Daten der Hamburger Schulinspektion haben jedoch gezeigt, dass die eindimensionale, IRT-basierte Skala verschiedene inhaltliche und methodische Vorteile gegenüber der

Nutzung einzelner Items hat: Einerseits zeigte sich, dass die Gesamtskala ein deutlich robusterer Indikator für die Unterrichtsqualität auf Schulebene ist als einzelne Items, da hier a) eine größere Beobachterübereinstimmung nachgewiesen werden kann und b) der durch die Ebene der Schule erklärte Varianzanteil höher ausfällt (vgl. Pietsch et al. 2011). Andererseits spielen bei einer derartigen Aggregation der Beobachtungsdaten selbst hohe Datenausfälle, die durch die Nicht-Beobachtbarkeit von Kriterien zustande kommen können, nahezu keine Rolle, da durch die Behandlung der fehlenden Werte mithilfe von Marginal-Maximum-Likelihood-Methoden potenzielle Verzerrungen in der Schätzung von Parametern auf Populationsebene ausgeglichen werden können (vgl. Pietsch 2011). Dabei liegt die interne Konsistenz der Skala bei 0.928 und es lässt sich nahezu keine schulformspezifische, jedoch eine auf die jeweils untersuchten Schulen zurückführende Varianz nachweisen (vgl. Pietsch 2010). Ebenfalls ist bekannt, dass die Unterrichtsqualität, die auf diese Weise im Rahmen der Inspektion erhoben wird, positiv mit Leistungsdaten zusammenhängt (vgl. IfBQ 2012), ein Indikator für externe Validität.

Eine Besonderheit des ersten Inspektionszyklus der Schulinspektion Hamburg lag darüber hinaus darin, dass die Inspektion bei der Modellierung von Unterrichtsqualität an die von Meyer (vgl. Meyer & Klapper 2006; Meyer 2008) initiierte Diskussion von Unterrichtsstandards anknüpfte. Um anschauliche Rückmeldungen geben zu können, wurde die metrische Skala in Abschnitte geteilt, die sich inhaltlich beschreiben lassen (vgl. Pietsch 2010). Die Schulinspektion Hamburg (vgl. IfBQ 2012) selbst unterscheidet dabei vier Stufen voneinander und definiert, dass Unterricht, in dem es ausschließlich darum geht, das Lernklima und pädagogische Strukturen zu sichern (Stufe 1), den Mindestanforderungen an lernwirksamen Unterricht nicht genügt, und benennt darüber hinaus die weiteren drei Stufen als Mindeststandard (Stufe 2), Regelstandard (Stufe 3) und Maximalstandard (Stufe 4). Dabei korrespondieren diese einzelnen Stufen mit den Basisdimensionen effektiven Unterrichts – Klassenführung, Schülerorientierung, kognitive Aktivierung und Lernautonomie (vgl. Klieme 2011). Den Befunden der Schulinspektion zufolge genügen rund 7 bis 8 Prozent aller Unterrichtsstunden an Hamburgs Schulen regelmäßig nicht dem Mindeststandard, sind also durch eine nicht gelingende Klassenführung gekennzeichnet, wohingegen im Mittel rund 25 Prozent aller Stunden den Maximalstandard gewährleisten und somit das im Hamburger Orientierungsrahmen Schulqualität definierte Ideal von Unterricht erreichen.

4.2 Beschreibung des Datensatzes

SINUS soll zu einer Verbesserung der Unterrichtskultur an Schulen führen. Im Fokus des Programms steht insofern nicht nur der einzelne Unterricht einer spezifischen Lehrkraft innerhalb einer Schule, sondern vielmehr die Weiterentwicklung

des Unterrichts auf Schulebene (vgl. Meyer 2011). Entsprechend nehmen Schulen an dem Programm teil, nicht einzelne Lehrerinnen und Lehrer. Diese Tatsache ermöglicht es, anhand von Daten der Schulinspektion Hamburg post-hoc herauszufinden, ob eine Schule am Programm teilgenommen hat oder nicht. So wird im Rahmen der Schulinspektion grundsätzlich erhoben, an welchen Maßnahmen und Programmen eine Schule in den drei Jahren vor einer Inspektion teilgenommen hat. Für jedes Jahr vor der Inspektion kann eine separate Angabe gemacht werden. Mithilfe dieser Daten wurden in einem ersten Schritt solche Schulen identifiziert, die in den drei Jahren vor einer Inspektion mindestens ein Jahr am SINUS-Projekt teilgenommen haben. Insgesamt nahmen 49 Grundschulen in Hamburg im Vorfeld einer Inspektion mindestens ein Jahr lang am Projekt SINUS teil. An diesen Schulen wurden zum Zeitpunkt der Inspektion 6.429 Schülerinnen und Schüler von 1.239 Lehrkräften unterrichtet. Im Vergleich hierzu nahmen 113 Schulen nicht an SINUS teil. An diesen Schulen wurden zum Zeitpunkt der Inspektion 13.698 Schülerinnen und Schüler von 2.608 Lehrerinnen und Lehrern unterrichtet. Insgesamt wurden an allen Grundschulen zusammen 8.361 Unterrichtssequenzen durch Schulinspektorinnen und Schulinspektoren eingesehen. Hiervon stammen 6.012 Unterrichtsbeobachtungen aus Schulen, die laut Eigenauskunft nicht an SINUS teilgenommen haben, und 2.349 aus SINUS-Schulen. Im Gesamtdatensatz liegt die Inter-Rater-Reliabilität der Beobachter, gemessen als Intra-Klassen-Korrelation (ICC), bei 0,92, was für eine hohe Konsistenz der Beobachter spricht. Dabei liegt die Unterrichtsqualität der SINUS-Schulen (Mittelwert=0,56) auf der Skala der Schulinspektion um rund 25 Prozent (Cohen's $d=0,18$) über derjenigen von Grundschulen ohne Teilnahme am Programm (Mittelwert=0,43), wobei der Unterschied hoch signifikant ist ($p<0,001$).

4.3 Analysemethode

Augenscheinlich unterscheidet sich somit die Qualität des Unterrichts an SINUS-Schulen von demjenigen an Nicht-SINUS-Schulen positiv und es gelingt entsprechend, Unterricht lernwirksamer zu gestalten. Gleichwohl kann nicht garantiert werden, dass sich SINUS-Schulen von den anderen Schulen mit Blick auf mögliche Kovariaten nicht systematisch unterscheiden, da SINUS-Schulen dem Programm nicht zufällig zugeordnet werden, sondern freiwillig am Programm teilnehmen. Insofern kann der berichtete Unterschied in der Unterrichtsqualität ebenso gut Selektionsverzerrungen geschuldet sein wie einem tatsächlichen Effekt des Programms. Die Kausalaussage „SINUS führt zu einer Effektivierung des Unterrichts“ ist daher nicht haltbar. Dieser Problematik kann auf verschiedenen Wegen begegnet werden, wobei es stets darum geht, kontrafaktische Kausalanalysen post-hoc mithilfe statistischer Verfahren zu ermöglichen (vgl. z. B. Legewie 2012).

Im vorliegenden Fall wird hierfür das sogenannte *Propensity Score Matching* genutzt (vgl. Rosenbaum & Rubin 1983). Die Grundidee dieses Verfahrens besteht darin, für alle Schulen der Treatmentgruppe (SINUS-Schulen) eine Kontrollgruppe zu ermitteln, die diesen Schulen in relevanten Merkmalen möglichst ähnlich ist. Auf diesem Wege soll sichergestellt werden, dass die zentrale Annahme der kontrafaktischen Kausalanalyse, die Annahme der konditionalen Unabhängigkeit (*Conditional Independence Assumption*, CIA) erfüllt wird. Dies bedeutet, dass nach der Kontrolle von Kovariaten die Verteilung der Analyseeinheiten über Treatment- und Kontrollgruppe möglichst zufällig ist.

Die Schätzung der *Propensity Scores* sowie das Matching wurden mithilfe des SPSS-Plugins *PS Matching* (vgl. Thoemmes 2014), das auf das R-Paket *match-it* zurückgreift, durchgeführt. Für das praktische Vorgehen existieren hier verschiedene Möglichkeiten (vgl. Thoemmes & Kim 2011), wobei für die nachfolgenden Analysen ein *Nearest Neighbor Matching* mit einer 1-zu-1-Zuordnung gewählt wurde. Entsprechend wurde jeder Treatment-Schule diejenige Schule mit dem ähnlichsten *Propensity Score* zugeordnet. Dabei wurde keine Begrenzung der Distanz der Scores berücksichtigt, also kein Caliper durchgeführt. Darüber hinaus wurde die Prozedur ohne Zurücklegen durchgeführt. Nicht zugeordnete Schulen ohne SINUS-Teilnahme wurden hingegen aus dem Datensatz entfernt, sodass – wie beabsichtigt – ebenso viele Schulen als Kontroll- wie als Treatmentgruppe im Analysedatensatz verblieben.

Als zentrale Hintergrundvariablen für das Matching wurden drei Bereiche definiert, die für eine Verletzung der CIA verantwortlich sein können, für die also ein Zusammenhang mit der zentralen Analysevariable Unterrichtsqualität zu erwarten ist: a) die soziale Zusammensetzung der Schülerschaft an der Schule, gemessen an deren Sozialindex, b) die Schulgröße, indiziert über die Anzahl der Schülerinnen und Schüler an der Schule sowie die Größe der jeweiligen Lehrerkollegien und c) das Inspektionsjahr. Die Auswahl dieser Variablen ist sowohl theoretisch als auch empirisch begründet. So ist ein nicht-linearer Zusammenhang zwischen der Unterrichtsqualität, wie sie die Schulinspektion Hamburg misst, und der sozialen Zusammensetzung der Schülerschaft an Schulen empirisch nachgewiesen (vgl. IfBM 2011). Darüber hinaus ist bekannt, dass die Schulgröße sowohl einen Einfluss auf die Lernergebnisse von Schülerinnen und Schülern als auch auf Innovations- und Entwicklungsmöglichkeiten von Schulen hat (vgl. Leithwood & Jantzi 2009). Und zu guter Letzt besteht für Schulen in Hamburg durch die zeitlich versetzte Inspektion das Problem, dass Schulen im zeitlichen Verlauf eine zunehmend höhere Wahrscheinlichkeit haben, Teil der jährlichen Inspektionsstichprobe zu werden (vgl. Pietsch et al. 2014). Die Antizipation dieser Tatsache könnte beispielsweise Schulen, die in den ersten Jahren der Inspektion noch nicht inspiziert wurden, dazu bringen, dass sie mit zunehmender Zeitdauer in Erwartung eines Schulinspektionsbesuches präventiv Maßnahmen einleiten, die zu einer Verbesserung der Unter-

richtsqualität führen, welche jedoch grundsätzlich unabhängig davon sind, ob eine Schule an SINUS teilgenommen hat oder nicht.

5. Befunde

5.1 Befunde zur Matchingqualität

Mithilfe des beschriebenen Verfahrens konnten den 49 Hamburger SINUS-Grundschulen 49 vergleichbare Schulen zugeordnet werden. Nach dem Matching stehen den 2.349 beobachteten Unterrichtssequenzen an SINUS-Schulen 2.533 Sequenzen an vergleichbaren Schulen gegenüber. Im Mittel wurden an Treatment-Schulen rund 48 Sequenzen von der Inspektion eingesehen, an Schulen der Kontrollgruppe 52. Die Güte des Matchings lässt sich dabei anhand verschiedener Indikatoren überprüfen: Einerseits sollten sich die beiden Gruppen hinsichtlich der kontrollierten Variablen nach dem Matching möglichst nicht mehr statistisch nachweisbar unterscheiden. Andererseits sollte der Standardisierte Bias, der Unterschied in den Mittelwerten der Kovariaten unter Berücksichtigung der Varianz dieser Merkmale (vgl. Rosenbaum & Rubin 1985), auf einen Wert von möglichst unter 5 Prozent reduziert worden sein (vgl. Caliendo & Kopeining 2008). Darüber hinaus muss die nicht-parametrische Identifikation des Treatmenteffektes durch eine Überlappung der *Propensity Scores* (*Common Support*) der beiden Gruppen garantiert werden. Denn haben Treatment- und Kontrollgruppe eine ähnliche Verteilung der *Propensity Scores*, bedeutet dies, dass sie sich auch mit Blick auf die kontrollierten Kovariaten ähnlich sind (vgl. Rubin 2001).

Die Untersuchung der Mittelwertunterschiede (vgl. Tab. 1) offenbart, dass sich die SINUS-Schulen vor dem Matching hinsichtlich der Kontrollvariablen statistisch nachweisbar von den anderen Schulen im Datensatz nur mit Blick auf das Inspektionsjahr unterscheiden ($p < 0.05$). Gleichwohl liegt der mittlere Standardisierte Bias vor dem Matching bei fast 30 und im Anschluss bei unter 3 Prozent; d. h. während die marginale Verteilung der Kovariaten vor dem Matching vergleichsweise hoch war, führt das Matching zu einer deutlichen Anpassung von Nicht-SINUS-Schulen an SINUS-Schulen mit Blick auf die kontrollierten Kovariaten.

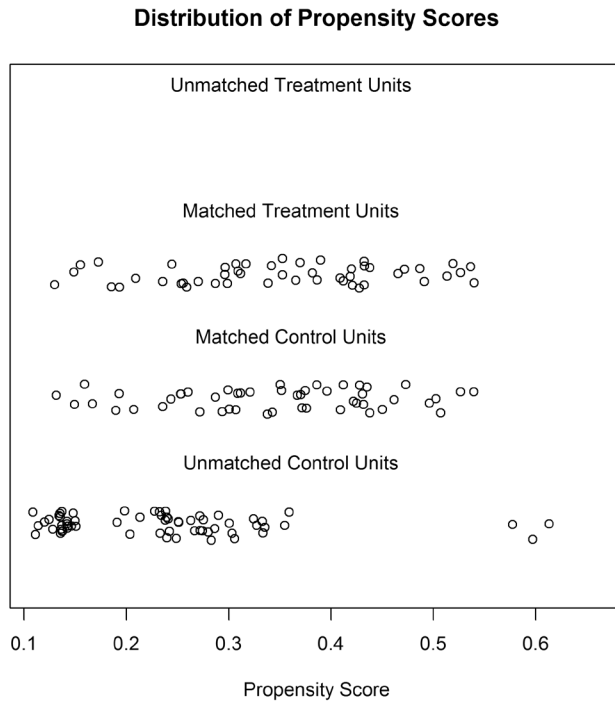
Tabelle 1: Übersicht über die Kontrollvariablen vor und nach dem Matching und den standardisierten Bias

	vor Matching			nach Matching		
	SINUS (N=49)	ohne SINUS (N=112)	Signi- fikanz	SINUS (N=49)	ohne SINUS (N=49)	Signi- fikanz
Anzahl Schüler/-innen	131	121	p>0.10	131	132	p>0.10
Anzahl Lehrer/-innen	25	23	p>0.10	25	26	p>0.10
Inspektionsjahr 2007 (in %)	4,1	12,5		4,1	2,0	
Inspektionsjahr 2008 (in %)	6,1	14,3		6,1	8,2	
Inspektionsjahr 2009 (in %)	10,2	16,1		10,2	16,3	
Inspektionsjahr 2010 (in %)	18,4	16,1	p<0.05	18,4	18,4	p>0.10
Inspektionsjahr 2011 (in %)	38,8	27,7		38,8	40,8	
Inspektionsjahr 2012 (in %)	12,2	13,4		12,2	14,3	
Inspektionsjahr 2013 (in %)	10,2	0,0		10,2	0,0	
Sozialindex I (in %)	8,2	18,8		8,2	8,2	
Sozialindex II (in %)	16,3	17,0		16,3	20,4	
Sozialindex III (in %)	6,1	13,4		6,1	4,1	
Sozialindex IV (in %)	28,6	14,3	p>0.10	28,6	22,4	p>0.10
Sozialindex V (in %)	16,3	18,8		16,3	14,3	
Sozialindex VI (in %)	16,3	13,4		16,3	20,4	
Sozialindex nicht vorhanden (in %)	8,2	4,5		8,2	10,2	
Mittlerer Standardisierter Bias		29,23			1,54	

Auch weitere Gütemaße lassen darauf schließen, dass das Matching zu einer guten Anpassung der Kontroll- an die Treatmentgruppe geführt hat. So zeigt Abbildung 1 dass selbst in den Randbereichen der *Propensity-Score*-Verteilung Matching-Partner für die SINUS-Schulen gefunden werden konnten.

Und letztlich deuten sowohl der globale χ^2 -Test ($p=0.93$) zur Prüfung der Stichprobenbalance nach Hansen und Bowers (2008) als auch die von Iacus, King und Porro (2009) zur Prüfung der Balanceoptimierung vorgeschlagene \mathcal{L}_1 -Statistik ($\Delta\mathcal{L}_1=0.06$) darauf hin, dass im vorliegenden Schuldatensatz für die gesamte Gruppe der SINUS-Schulen eine adäquat ausbalancierte Kontrollgruppe gefunden werden konnte und die Vergleichbarkeit von Treatment- und Kontrollgruppe durch das Matching hinsichtlich der genutzten Kovariaten merklich verbessert wurde.

Abbildung 1: Verteilung der Propensity Scores



5.2 Befunde zur Wirksamkeit des Programms SINUS an Grundschulen

5.2.1 Unterschiede in schulischen Prozessmerkmalen

Eine gute Schule zeichnet sich nach Ansicht der Hamburger Schulinspektion (vgl. IfBQ 2012) maßgeblich dadurch aus, dass unterrichtliche, erzieherische und organisatorische Prozesse gut gelingen. Welche Prozessmerkmale für eine gelingende Schul- und Unterrichtsentwicklung zentral sind, ist für Hamburger Schulen im Orientierungsrahmen Schulqualität (vgl. Behörde für Schule und Sport 2006) festgelegt. Der Orientierungsrahmen setzt Maßstäbe, an denen sich die Qualitätsentwicklung in Schulen orientieren soll. Auf diese Weise stellt er Kategorien für die Berichterstattung über Qualität bereit, die für die interne und externe Evaluation genutzt werden können. Der Orientierungsrahmen geht dabei davon aus, dass sich die Erfolge der Schule an ihren Wirkungen und Ergebnissen ablesen lassen. Ziel aller Qualitätsbemühungen einer Schule sind sichtbare Leistungen. Bedingungen dafür sind Verfahren der Führung und des Managements sowie Bildungs- und Er-

ziehungsprozesse, die für die Zielerreichung geeignet sind und deren Wirksamkeit empirisch nachgewiesen ist. Die Schulinspektion untersucht diese Verfahren und Prozesse. Die Schulinspektion Hamburg leitet ihre Bewertung der Qualität von Schule aus den 14 Qualitätsbereichen des Orientierungsrahmens ab, wobei Leistungen nicht bewertet werden. Die Bewertungen können dabei von „schwach“ über „eher schwach“ und „eher stark“ bis hin zu „stark“ reichen.

Tabelle 2: Bewertung der Qualitätsdimensionen an Schulen mit und ohne SINUS-Teilnahme (nach dem Matching)

Qualitätsdimension	SINUS	Nicht-SINUS	p
OR 1.1 Führung wahrnehmen	3,04	3,08	n.s.
OR 1.2 Personal entwickeln	2,60	2,40	n.s.
OR 1.3 Finanz- und Sachmittel gezielt einsetzen	3,27	3,12	n.s.
OR 1.4 Profil entwickeln und Rechenschaft ablegen	3,19	3,16	n.s.
OR 2.1 Das schuleigene Curriculum gestalten	2,08	1,98	n.s.
OR 2.2 Unterrichten, Lernen, Erziehen	2,96	2,76	<0.05
OR 2.3 Organisatorische Rahmenbedingungen sichern	3,35	3,20	n.s.
OR 2.4 Leistung beurteilen	2,42	2,38	n.s.
OR 2.5 Prozesse und Ergebnisse evaluieren	2,08	2,16	n.s.
OR 2.6 Förderkonzepte entwickeln	2,75	2,64	n.s.
OR 2.7 Beratungsangebote gestalten	2,83	2,82	n.s.
OR 2.8 Die Schulgemeinschaft beteiligen	2,88	2,84	n.s.
OR 3.1 Zufriedenes Personal, Schüler, Eltern, Betriebe	3,56	3,50	n.s.

OR = Orientierungsrahmen

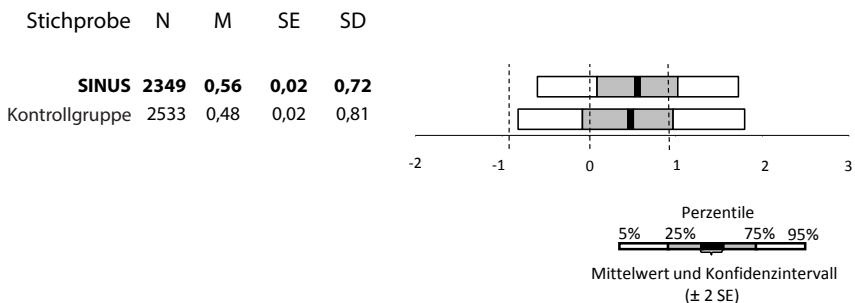
Diesbezüglich stellt sich die Frage, ob sich SINUS-Schulen von vergleichbaren Schulen nach Einschätzung der Schulinspektion systematisch in diesen Bereichen unterscheiden, wobei mit Blick auf die Qualität des Unterrichts insbesondere der Punkt 2.2 des Hamburger Orientierungsrahmens „Unterrichten, Lernen, Erziehen“ von Relevanz ist, da sich hier hinter die globale Einschätzung der Unterrichtsqualität an den inspizierten Schulen verbirgt, die durch die Inspektorinnen und Inspektoren der Schulinspektion auf Basis von Fragebogendaten, Interviews, Dokumentenanalysen und Unterrichtsbesuchen getroffen werden. Schaut man sich die Bewertungen der einzelnen Punkte an, so ist auffällig, dass sich Hamburger Grundschulen, die am Programm SINUS teilgenommen haben, ausschließlich in dem für das Programm relevanten Prozessbereich „Unterrichten, Lernen, Erziehen“ von vergleichbaren Schulen statistisch nachweisbar unterscheiden ($p < 0.05$, vgl. Tab. 2). In diesem Bereich liegt die Qualität an SINUS-Schulen nach Einschätzung der Inspektionsteams

um etwa 7 Prozent höher als an den Schulen der Kontrollgruppe. In allen anderen Bereichen, die die Schulinspektion einschätzt, unterscheiden sich Treatment- und Kontrollgruppe hingegen statistisch nicht nachweisbar voneinander ($p > 0.10$). Auf Ebene der schulischen Prozesse ist demnach die Qualität des Unterrichts das einzige Unterscheidungsmerkmal zwischen SINUS und Nicht-SINUS-Schulen.

5.2.2 Unterschiede in Merkmalen der Unterrichtsqualität

Da sich die Einschätzungen der Inspektorinnen und Inspektoren zu diesem Punkt großteils aus den Unterrichtsbeobachtungen speisen, ist zu erwarten, dass auch hier differentielle Effekte nachweisbar sind. Eine Erwartung, die sich bestätigt (vgl. Abb. 2).

Abbildung 2: Qualität des Unterrichts an Schulen mit und ohne SINUS-Teilnahme (nach dem Matching)



So unterscheidet sich die Qualität des Unterrichts an SINUS-Schulen ($M=0,56$) auch nach dem Matching noch statistisch signifikant ($p < 0.01$) von derjenigen an vergleichbaren Schulen ($M=0,48$), die nicht am Programm teilgenommen haben. Auffällig ist dabei vor allem auch, dass die Streuung der Qualität an den Programmschulen ($SD=0,72$) selbst nach Kontrolle der Schulgröße deutlich geringer ausfällt als an vergleichbaren Grundschulen ($SD=0,81$). An SINUS-Schulen ist demnach die Qualität des Unterrichts auf Schulebene homogener als an Schulen, die nicht an SINUS teilgenommen haben, jedoch unter vergleichbaren Bedingungen arbeiten. Betrachtet man Abbildung 2 genauer, so fällt darüber hinaus auf, dass sich die beiden Schulgruppen mit Blick auf die Verteilung vor allem im unteren Bereich unterscheiden. Extrem schwacher Unterricht ist an SINUS-Schulen demnach im Vergleich seltener zu finden als an vergleichbaren Schulen. Besonders deutlich wird dies, wenn man die Standards für Unterrichtsqualität der Schulinspektion anlegt.

Grundsätzlich verfehlen an allen Schulen des gematchten Datensatzes nur wenige Unterrichtssequenzen den Mindeststandard der Schulinspektion (N=141, 3%). Dabei reduziert die Teilnahme am SINUS-Programm das relative Risiko (RR), einen Unterricht an der Schule vorzufinden, bei dem es ausschließlich darum geht, das Lernklima und pädagogische Strukturen zu sichern – der also nicht lernwirksam ist –, um rund 56 Prozent bzw. erhöht die Chance (*Odds Ratio*, OR), zumindest den Mindeststandard an Unterrichtsqualität zu erreichen, um das 2,3-Fache. Mit Blick auf den Maximalstandard der Schulinspektion lassen sich hingegen keine bedeutsamen Unterschiede zwischen Treatment- und Kontrollgruppe feststellen.

Was dies praktisch bedeutet, veranschaulicht Tabelle 3. Dargestellt werden hier Merkmale der Unterrichtsqualität, die die Dimensionen effektiven Unterrichts, wie von Klieme (vgl. z.B. Klieme & Rackoczy 2008) beschrieben, charakterisieren. Bezogen auf dieses Modell, zeigt sich, dass die Klassenführung an allen in Hamburg inspizierten Grundschulen durchweg gut gelingt. Die Wahrscheinlichkeit, einen Unterricht zu finden, der im Rahmen der Schulinspektion diesbezüglich positiv eingeschätzt wird, beträgt an Schulen, die nicht am SINUS-Programm teilgenommen haben, 92 Prozent, an SINUS-Schulen sogar 94 Prozent ($p<0.001$).

Tabelle 3: Unterschiede in den Basisdimensionen effektiven Unterrichts an Schulen mit und ohne SINUS-Teilnahme (nach dem Matching)

	Klassenmanagement		Schülerorientierung		Kognitive Aktivierung		Lernautonomie	
Indikator-item	Die Arbeitsaufträge und Erklärungen sind angemessen, klar und präzise formuliert.		In den Unterricht sind der Erfahrungshorizont und/oder die Interessen der Schülerinnen und Schüler einbezogen.		Der Unterricht eröffnet Freiräume und ist nicht auf eine richtige Antwort fixiert.		Die Schülerinnen und Schüler werden angeregt/angeleitet, den Unterricht aktiv mitzugestalten, oder sie gestalten den Unterricht aktiv mit.	
	MW	p (MW>2.5)	MW	p (MW>2.5)	MW	p (MW>2.5)	MW	p (MW>2.5)
SINUS-Schulen	3,50	94%	3,00	73%	2,35	44%	1,93	28%
Kontrollgruppe	3,46	92%	2,97	70%	2,30	42%	1,90	27%

*** Mittelwertunterschied signifikant auf dem Niveau von 0.001; ns Mittelwertunterschied nicht signifikant

Auch hinsichtlich der Schülerorientierung resp. des unterstützenden Unterrichtsklimas zeigt sich, dass diese generell positiv wahrgenommen (Wahrscheinlichkeit einer positiven Bewertung 73 % vs. 70 %, $p<0.001$) wird. Das Gleiche gilt für den Bereich der kognitiven Aktivierung. Die Wahrscheinlichkeit, dass dieses Merkmal sich an SINUS-Grundschulen positiv von vergleichbaren Schulen ab-

hebt, liegt auch hier nachweisbar höher ($p < 0.001$), wobei sich jedoch zeigt, dass dieses Merkmal im Rahmen der Schulinspektion grundsätzlich mit unter 50-prozentiger Wahrscheinlichkeit positiv bewertet wird. Schaut man sich zu guter Letzt den Bereich der Lernautonomie an, so wird sichtbar, dass dieser im Rahmen der Unterrichtsbeobachtungen der Schulinspektion Hamburg selten positiv bewertet wird (Wahrscheinlichkeit einer positiven Bewertung 28 % vs. 27 %) und dass sich keinerlei Unterschiede zwischen SINUS- und vergleichbaren Grundschulen nachweisen lassen ($p > 0.100$).

6. Diskussion und Ausblick

Es ist bekannt, dass wissenschaftsbasierte Schulentwicklungsmaßnahmen meist nachhaltige Effekte auf Schülerleistungen haben. Auch für das Programm SINUS an Grundschulen konnte kürzlich Vergleichbares mithilfe von Daten aus TIMSS 2011 festgestellt werden. Gleichwohl ist derzeit nur wenig über die Wirkungsmechanismen – also die vermittelnden Faktoren, auf Basis derer das Programm wirksam wird – bekannt. Befunde aus vorhergehenden Studien lassen erwarten, dass SINUS seine Wirkung primär über die Unterrichtsgestaltung entfaltet und sich der Unterricht an SINUS-Grundschulen somit von demjenigen an anderen Schulen, die nicht am SINUS-Programm teilnehmen, unterscheidet.

Mithilfe von Daten der Hamburger Schulinspektion, die im Zeitraum von 2006 bis 2013 Daten zur Unterrichtsqualität an allen Hamburger Schulen gesammelt hat und die den Anspruch hat, hiermit die Qualität des Unterrichts auf Schulebene empirisch verlässlich abbilden zu können, wurde diese Annahme geprüft. Da die Teilnahme am SINUS-Programm nicht zufällig erfolgte, wurde ein Quasi-Experiment durchgeführt, in dem a) verschiedene Prozessmerkmale effektiver Schulen, die im Rahmen der Inspektion durch Schulinspektorinnen und Schulinspektoren eingeschätzt werden, und b) der Unterricht an Hamburger Schulen, die an SINUS teilgenommen haben bzw. teilnehmen, mit denjenigen an Schulen verglichen wurden, die nicht an SINUS teilgenommen haben, jedoch unter vergleichbaren Bedingungen arbeiten. Um mögliche Selektionseffekte zu minimieren, wurde hierfür im Rahmen eines *Propensity Score Matching* die soziale Zusammensetzung der Schülerschaften, die Schulgrößen sowie das Inspektionsjahr der Schulen kontrolliert.

Die Befunde des Gruppenvergleiches bestätigen die bislang vorliegenden Befunde zu den Wirkungsmechanismen des Programms SINUS an Grundschulen. So konnte in einem ersten Schritt gezeigt werden, dass sich SINUS-Schulen von vergleichbaren Schulen mit Blick auf deren prozessuale Effektivitätsmerkmale – im Hamburger Orientierungsrahmen Qualitätsbereiche genannt – nach Einschätzung der Inspektionsteams ausschließlich im Bereich der Unterrichtsgestal-

tung statistisch nachweisbar voneinander unterscheiden. Entsprechend ließen sich weder Unterschiede in der Führung durch Schulleitungen, den organisatorischen Rahmenbedingungen des Unterrichts, der Nutzung von Evaluationskonzepten, der Beteiligung von Eltern und Schülern sowie der Zufriedenheit der Schulbeteiligten mit ihren Schulen nachweisen. Grundsätzlich scheint SINUS somit vor allem das intendierte Ziel einer verbesserten Unterrichtsentwicklung zu erreichen.

Was dies konkret bedeutet, wurde in einem zweiten Schritt gezeigt, in dem Daten aus Unterrichtsbeobachtungen der Schulinspektion genutzt wurden. Anhand von jeweils mehr als 2.000 Beobachtungen wird sichtbar, dass sich die Qualität des Unterrichts an SINUS-Schulen nachweisbar positiv von demjenigen an vergleichbaren Schulen abhebt. Im Mittel liegt die Qualität des Unterrichts an SINUS-Schulen rund 16 Prozent über derjenigen an vergleichbaren Schulen. Auffällig ist dabei weiterhin, dass, selbst bei Kontrolle der Schulgröße, die Variation der Unterrichtsqualität innerhalb der Stichproben unterschiedlich groß ist. So fällt die Streuung innerhalb der SINUS-Stichprobe rund 13 Prozent geringer aus als in der Kontrollgruppe. Und zu guter Letzt wird sichtbar, dass SINUS darüber hinaus dazu führt, dass Schwächen im Unterricht eliminiert werden. Unterricht, der den Mindeststandard an Unterrichtsqualität, wie ihn die Schulinspektion Hamburg definiert, nicht erreicht, ist an SINUS-Schulen sehr selten zu vorzufinden und findet dort weniger als halb so häufig statt wie an vergleichbaren Schulen.

Der Nachweis dieser drei Punkte ist dabei von besonderer Bedeutung, da die wenigen Studien, die Merkmale der Schul- mit denen der Unterrichtseffektivitätsforschung zusammenbringen (vgl. Ellett & Teddlie 2003), zeigen, dass der Unterricht an effektiven Schulen sich in genau diesen drei Bereichen von demjenigen an weniger effektiven Schulen unterscheidet. Unterricht an effektiven Schulen ist demnach gekennzeichnet durch a) eine generell höhere Unterrichtsqualität, b) eine geringere Streuung der Qualität innerhalb der Schule und c) dem Nichtvorhandensein von extrem schwachen Unterricht. Dies sind allesamt nachgewiesene Merkmale von SINUS-Schulen, was im Umkehrschluss darauf hindeutet, dass es sich bei SINUS-Schulen um Schulen mit einer vergleichsweise höheren Effektivität handelt, an denen Schülerinnen und Schüler höchstwahrscheinlich bessere Lernergebnisse erzielen können als an anderen Schulen.

Zusammenfassend lässt sich entsprechend feststellen, dass SINUS an Grundschulen grundsätzlich zu einer Effektivierung der Unterrichtspraxis führt und dass an diesen Schulen infolgedessen bessere Lernergebnisse aufseiten der Schülerinnen und Schüler zu erwarten sein sollten. Der Primäreffekt liegt dabei darin, dass Schwächen im Unterricht auf Ebene der Schule flächendeckend abgestellt werden. In der Konsequenz ermöglicht das Programm den Lehrkräften an den teilnehmenden Schulen, sich verstärkt auf die Motivations- und die Leistungsentwicklung der Schülerinnen und Schülern zu konzentrieren und den Unterricht zugewandt, fehlerfreundlich und unterstützend sowie kognitiv herausfordernd zu gestalten.

Besonders bedeutsam wird dieser Effekt, wenn man ihn auf das Gesamtsystem überträgt: So finden an Hamburgs Grundschulen in den Klassenstufen eins bis vier wöchentlich ca. 72.500¹ Unterrichtsstunden statt, wovon den Befunden der Inspektion zufolge etwa 2.200 (3 %) unterhalb des Mindeststandards liegen dürften. SINUS erhöht nun die Chance, dass der durch die Schulinspektion definierte Mindeststandard von Unterrichtsqualität erreicht wird, um das 2,3-fache ($OR=2,3$, $d=0.46$, $r=0.23$) und hat entsprechend das Potenzial, den Anteil an Unterrichtsstunden pro Woche, in denen die Lernwirksamkeit nicht im Mittelpunkt steht, um rund ein Viertel zu senken. Das heißt, rund 500 Stunden pro Woche bzw. 20.000 Stunden pro Schuljahr, bei einem Schuljahr mit 40 Schulwochen, könnten durch eine flächendeckende Einführung des Programms an Grundschulen in Hamburg effektiver gestaltet und auf diesem Wege könnte die Wahrscheinlichkeit von Lernerfolgen aufseiten der Schülerinnen und Schüler erhöht werden.

Welche Folgen die Unterrichtspraxis an Hamburger SINUS-Schulen konkret für die Lernentwicklung von Schülerinnen und Schülern hat, kann die vorgelegte Studie leider nicht klären. Dies ist besonders bedauerlich, da Dalehefte et al. (2014) darauf verweisen, dass die von ihnen berichteten kausal interpretierten Befunde zur Wirksamkeit von SINUS auf die mathematische Kompetenz von Grundschülerinnen und Grundschülern nur bedingt haltbar seien, da kein experimentelles oder quasi-experimentelles Untersuchungsdesign genutzt wurde, das den Kompetenzzuwachs aufseiten der Schülerinnen und Schüler berücksichtigt.

Insofern hätte eine Verbindung von Prozess- und Ergebnisdaten im Rahmen der hier vorgelegten quasi-experimentellen Studie die Möglichkeit geboten, a) die Befunde aus TIMSS 2011 zu validieren und b) den Einfluss des veränderten Unterrichts an SINUS-Schulen auf die Kompetenzentwicklung von Schülerinnen und Schülern empirisch verlässlich abzuschätzen. Da flächendeckende Untersuchungen zur Lernentwicklung von Schülerinnen und Schülern in Hamburg erst zum Ende des ersten Inspektionszyklus eingeführt wurden, waren derartige Analysen hier jedoch nicht möglich. Zukünftig böte sich jedoch zur Klärung vergleichbarer Fragestellungen die Verbindung von Schulinspektionsdaten mit den ebenfalls in Hamburg vorhandenen Daten der KERMIT-Erhebungen (Kompetenzen ERMITteln) an, wobei auch in einem solchen Szenario etliche methodische Fallstricke zu berücksichtigen wären, um empirisch verlässliche Kausalaussagen zur Wirksamkeit von Schulentwicklungsmaßnahmen auf schulische Prozess- und Ergebnismerkmale zu gewährleisten.

1 Im Schuljahr 2013/2014 wurden an Hamburger Grundschulen 53.649 Schülerinnen und Schüler in den Klassenstufen eins bis vier in Klassen mit einer Größe von durchschnittlich 20 Personen unterrichtet. Die Anzahl der Grundstunden, die sie während der Grundschulzeit wöchentlich belegen müssen beträgt derzeit 108. Hieraus ergeben sich: $\frac{53.649}{20} \times \frac{108}{4} = 72.426$ Unterrichtsstunden pro Schulwoche an Hamburgs Grundschulen.

Literatur

- Behörde für Bildung und Sport (2006). *Orientierungsrahmen: Qualitätsentwicklung an Hamburger Schulen*. Hamburg: Behörde für Bildung und Sport.
- Borman, G., Hewes G., Overman, L. & Brown, S. (2003). Comprehensive School Reform and Achievement: A Meta-Analysis. *Review of Educational Research*, 73 (2), 125–230.
- Caliendo, M. & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22 (1), 31–72.
- Dalehefte, I.-M., Wendt, H., Köller, O., Wagner, H., Pietsch, M., Döring, B., Fischer, C. & Bos, W. (2014). Bilanz von neun Jahren SINUS an deutschen Grundschulen: Evaluation im Rahmen der TIMSS 2011-Erhebung. *Zeitschrift für Pädagogik*, 60 (2), 245–263.
- Demuth, R., Walther, G. & Prenzel, M. (2011). *Unterricht entwickeln mit SINUS. 10 Module für den Mathematik- und Sachunterricht in der Grundschule*. Seelze: Klett-Kallmeyer.
- Ellet, C.D. & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, 17 (1), 101–128.
- Fischer, C. (2005). Rolle von Schulleitungen bei innovativen Projekten. Aus Sicht der Grundschule. *SchulVerwaltung. Zeitschrift für Schulleitung, Schulaufsicht und Schulkultur Ausgabe Niedersachsen und Schleswig-Holstein*, 9, 251–253.
- Fischer, C., Kobarg, M., Dalehefte I.-M. & Trepke, F. (2013). Ein Unterrichtsentwicklungsprogramm wissenschaftlich begleiten. Anlage und Hintergründe des Forschungsdesigns. *Psychologie in Erziehung und Unterricht*, 60 (1), 26–31.
- Fischer, C. & Prenzel, M. (2007). Wie Lehrkräfte bei der Qualitätssicherung unterstützt werden können. Das Beispiel SINUS-Transfer Grundschule. *Journal für Schulentwicklung*, 11 (2), 23–39.
- Fischer, C., Rieck, K. & Lobemeier, K.R. (2008). Mit Logbüchern dokumentieren und reflektieren. Das Beispiel SINUS-Transfer Grundschule. In E. M. Lankes (Hrsg.). *Pädagogische Professionalität als Gegenstand empirischer Forschung* (S. 73–86). Münster: Waxmann.
- Fischer, C., Rieck, K. & Dedekind, B. (2009): SINUS-Transfer Grundschule. Lehrkräfte verändern ihren Mathematikunterricht und ihren naturwissenschaftlichen Sachunterricht an Grundschulen – (wie) geht das? *MNU PRIMAR*, 2, 44–49.
- Hansen, B.B. & Bowers, J. (2008). Covariate balance in simple, stratified and clustered comparative studies. *Statistical Science*, 23, 219–236.
- Helmke, A. (2006). Was wissen wir über guten Unterricht? *Pädagogik*, 58 (2), 42–45.
- Hill, C. J., Bloom, H. S., Black, A. R. & Lipsey, M. W., (2008). Empirical Benchmarks for Interpreting Effect Sizes in Research. *Child Development Perspectives*, 2 (3), 172–177.
- Iacus, S.M., King, G. & Porro, G. (2009). CEM: Coarsened exact matching software. *Journal of Statistical Software*, 30, 1–27.
- IfBM (2011). *Jahresbericht der Schulinspektion Hamburg. Schuljahr 2009–2010*. Hamburg: Institut für Bildungsmonitoring.

- IfBQ (2012). *Jahresbericht der Schulinspektion Hamburg. Schuljahr 2010–2011*. Hamburg: Institut für Bildungsmonitoring und Qualitätsentwicklung.
- Klieme, E. (2011). *Standards für die Unterrichtsqualität. Kann es das geben?* Vortrag an der Universität Zürich, 20. November 2011.
- Klieme, E. & Rackoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54 (2), 222–237.
- Legewie, J. (2012). Die Schätzung von kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der Schule. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 64 (1), 123 – 153.
- Leithwood, K. & Jantzi, D. (2009). A Review of Empirical Evidence about School Size Effects: A Policy Perspective. *Review of Educational Research*, 79 (1), 464–490.
- Meyer, H. (2008). Disput über aktuelle Probleme und Aufgaben der Didaktik. In M. A. Meyer, M. Prenzel & S. Hellekamps (Hrsg.), *Perspektiven der Didaktik (Zeitschrift für Erziehungswissenschaft: Sonderheft 9)*, S. 77–86). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Meyer, H. (2011). *Die Rolle der Schulleitung bei der Unterrichtsentwicklung*. Kiel: IPN.
- Meyer, H. & Klapper, A. (2006). Unterrichtsstandards für ein kompetenzorientiertes Lernen und Lehren. In R. Hinz & B. Schumacher (Hrsg.), *Auf den Anfang kommt es an: Kompetenzen entwickeln – Kompetenzen stärken* (S. 89–108). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Morgan, S. L. & Winship, C. (2007). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York: Cambridge University Press.
- Ostermeier, C. (2005). *Kooperative Qualitätsentwicklung in Schulnetzwerken: Eine empirische Studie am Beispiel des BLK-Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts (SINUS)“*. Münster: Waxmann.
- Pietsch, M. (2010). Evaluation von Unterrichtsstandards. *Zeitschrift für Erziehungswissenschaft*, 13 (1), 121–148.
- Pietsch, M. (2011). Fehlende Daten bei Unterrichtsbeobachtungen: Eine Sensitivitätsanalyse anhand von Daten der Schulinspektion Hamburg. *Empirische Pädagogik*, 25 (1), 47–87.
- Pietsch, M., Jahnke, N. & Mohr, I. (2014). Führt Schulinspektion zu besseren Schülerleistungen? Difference-in-Difference-Studien zu Effekten der Schulinspektion Hamburg auf Lernzuwächse und Leistungstrends. *Zeitschrift für Pädagogik*, 60 (3), 446–470.
- Pietsch, M., Schulze, P., Schnack, J. & Krause, M. (2011). Elaborierte Rückmeldungen zur Qualität von Unterricht. Über empirisch abgesicherte Bezugsnormen als Grundlage für die Weiterentwicklung von Unterricht und Schule. In S. Müller, M. Pietsch & W. Bos (Hrsg.), *Schulinspektionen in Deutschland – Eine Zwischenbilanz aus empirischer Sicht* (S. 193–216). Münster: Waxmann.
- Prenzel, M., Friedrich, A. & Stadler, M. (2009). *Von SINUS lernen. Wie Unterrichtsentwicklung gelingt*. Seelze: Klett-Kallmeyer.

- Prenzel, M., Carstensen, C. H., Senkbeil, M., Ostermeier, C. & Seidel, T. (2005). Wie schneiden SINUS-Schulen bei PISA ab? Ergebnisse der Evaluation eines Modellversuchsprogramms. *Zeitschrift für Erziehungswissenschaft*, 8 (4), 540–561.
- Rosenbaum, P. & Rubin, D. (1983). The Central Role of Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70 (1), 41–55.
- Rosenbaum, P. & Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39, 33–38.
- Seidel T. & Shavelson, R. J. (2007). Teaching effectiveness research in the last decade: Role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77 (4), 454–499.
- Rubin, D. (2001). Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services & Outcomes Research Methodology*, 2, 169–188.
- Stadler, M. & Meentzen, U. (2010). Wie Lehrkräfte bei der Reflexion über ihren Unterricht unterstützt werden können. Das Fachgruppenportfolio im Programm SINUS-Transfer. In F. H. Müller (Hrsg.), *Lehrerinnen und Lehrer lernen. Konzepte und Befunde zur Lehrerfortbildung* (S. 161–173). Münster: Waxmann.
- Stralla, M. (2009). *Die Unterrichtsbeobachtungen im Rahmen der deutschen Schulinspektion. Analyse des Kerninstruments zur Beurteilung der Schulqualität* (Unveröffentlichte Diplomarbeit). Berlin: Freie Universität Berlin.
- Thoemmes, F. (2014). *Propensity score matching in SPSS*. Verfügbar unter: <http://sourceforge.net/projects/psmspss/> [Zugriff am 30.03.2015]
- Thoemmes, F. & Kim, E. S. (2011). A Systematic Review of Propensity Score Methods in the Social Sciences. *Multivariate Behavioral Research*, 46, 90–118

Der Einfluss schulischer Prozesse auf die Lernentwicklung der Schülerinnen und Schüler an weiterführenden Schulen in Hamburg¹

Klaudia Schulte & Markus Lücken

Zusammenfassung

Zur Beantwortung der Frage, welche Merkmale schulischer Qualität einen Unterschied im Hinblick auf die Lernentwicklung machen, werden die zehn Schulen mit den höchsten Lernentwicklungen von Klasse 5 zu Klasse 7 in Hamburger Vergleichsarbeiten (KERMIT) und die zehn Schulen mit den niedrigsten Lernentwicklungen bezüglich der Unterschiede in der schulischen Prozessqualität verglichen. Die Schulen mit den höheren Lernentwicklungen schneiden in 7 der 13 durch die Schulinspektion erhobenen Qualitätsdimensionen besser ab als die Schulen der Gruppe mit niedrigeren Lernentwicklungen, der größte Effekt besteht bei dem Merkmal „Die Schulgemeinschaft beteiligen“. Diese Ergebnisse werden für die gesamte Stichprobe (41 Schulen bzw. 4.479 Schüler) unter Berücksichtigung der Kontextmerkmale (Hamburger Sozialindex) regressionsanalytisch überprüft. Dafür wird der Einfluss von acht Qualitätsmerkmalen sowie die soziale Belastung der Schule und die Schulform auf die Lernentwicklung in den Fächern Mathematik und Deutsch berücksichtigt. Drei Qualitätsbereiche sind signifikante Prädiktoren der Lernentwicklung, einer davon wirkt jedoch negativ. Nach Berücksichtigung des Sozialindex und der Schulform verschwinden diese Effekte.

1. Einführung

Spätestens seit den Konstanzer Beschlüssen 1997, den Ergebnissen von PISA 2000 und der daraus entstandenen Diskussion zur Bildung in Deutschland wurde eine empirische Wende eingeläutet, in der sozialwissenschaftlich gesicherte Befunde an Bedeutung für die Schulpolitik und die pädagogische Praxis gewinnen (vgl. Lange 2008). Dies resultierte in einem Paradigmenwechsel, der sich von einer Input- oder Ressourcenorientierung hin zu einer Outcome-Orientierung bewegt, in der die zu erreichenden Ziele bzw. Kompetenzen von Schülerinnen und Schülern fokussiert

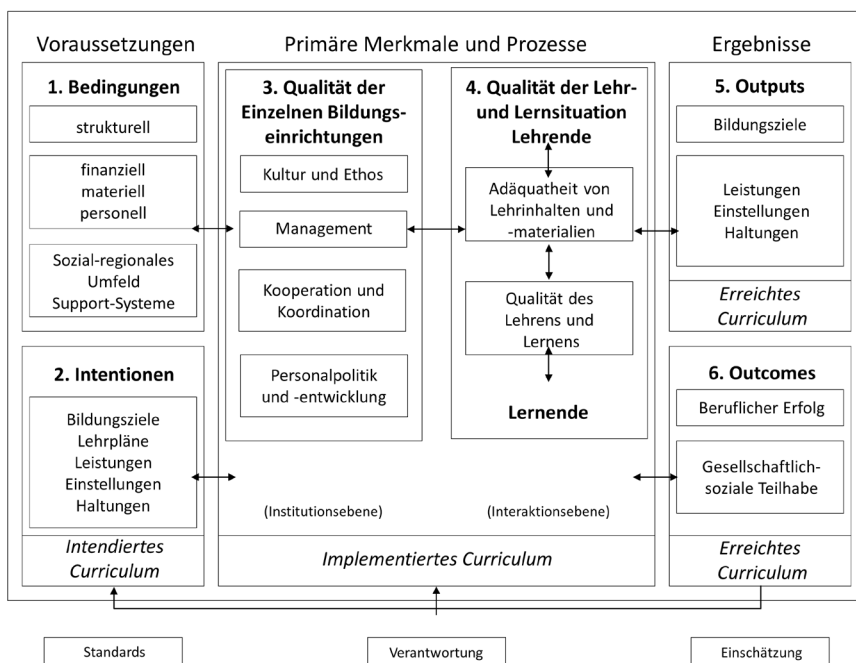
1 Wir danken Knut Schwippert, Marcus Pietsch und Felix Brümmer für die hilfreichen Hinweise zum Manuskript.

werden. Nicht die Ausbildung der Lehrkräfte, die Bereitstellung von Ressourcen und ein guter Lehrplan stehen im Vordergrund, sondern das Ziel, „kompetente“ Schülerinnen und Schüler auszubilden, ausgehend von verbindlichen Standards.

Dennoch reicht ein alleiniger Blick auf die Ergebnisse in der Schule, also auf die Leistungen der Schülerinnen und Schüler nicht aus. Um nachhaltige Schulentwicklung beschreiben und erklären zu können, muss untersucht werden, welcher Input, d. h. welche Ressourcen und welche Prozesse, mit den Leistungsergebnissen in einem (kausalen) Zusammenhang stehen.

In etablierten Modellen zur Schulqualität wird daher zwischen den Voraussetzungen, primären Merkmalen und Prozessen sowie schulischen Ergebnissen unterschieden und zwischen diesen Aspekten eine kausale und zeitliche Abfolge postuliert (vgl. Ditton 2007; siehe Abbildung 1 sowie den Beitrag von Gärtner in diesem Buch). Nach diesen Modellen sind Leistungen und Erfolge, die Schülerinnen und Schüler erreichen, sowohl durch die Qualitätsmerkmale und Prozesse der schulischen Einrichtung als auch durch die äußeren Rahmenbedingungen und Voraussetzungen bestimmt, unter denen Schulen arbeiten.

Abbildung 1: Modell zur Qualität und Qualitätsentwicklung im Bildungsbereich



In Hamburg liegen für diese drei Konzepte spezifische Daten vor: Zur Bestimmung der Voraussetzungen, unter denen Schulen arbeiten, kann in Hamburg der Sozialindex hinzugezogen werden. Mit diesem Index lässt sich die soziale Zusammensetzung der Schülerschaft und das sozial-regionale Umfeld der Schulen abbilden (s. Abschnitt 1.1). Die primären Merkmale und Prozesse der Schule werden in Hamburg im Rahmen der Schulinspektionen erfasst. Dort werden unterschiedliche schulische Prozessaspekte in den Blick genommen, zum Beispiel der Bereich des Unterrichtens (s. Abschnitt 1.2). Um Hinweise auf die Ergebnisse der Qualitätsentwicklung zu erhalten, können die Leistungen der Schülerinnen und Schüler genutzt werden, die in Hamburg im Rahmen von längsschnittlich angelegten Lernstandserhebungen regelhaft mit dem Instrument KERMIT (Kompetenzen ERMITteln; vgl. Lücken et al. 2014) erfasst werden (s. Abschnitt 1.3). Diese drei Bereiche der Schulqualität werden im Folgenden ausdifferenziert.

1.1 Voraussetzungen der Qualität im Bildungsbereich

Zur Operationalisierung eines Teils der Voraussetzungen, unter denen die Hamburger Schulen arbeiten, kann der Sozialindex der Schulen genutzt werden. Diesen gibt es in Hamburg seit 1996 für Grundschulen und weiterführende Schulen mit Sekundarstufe I. Der Sozialindex (vgl. Bos et al. 2006) beschreibt, basierend u. a. auf der Kapitaltheorie von Bourdieu (1982, 1983), die sozialen Rahmenbedingungen der Schulen auf einer Skala von 1 (sehr belastet) bis 6 (wenig belastet). Die Zuweisung der Schulen zu diesen sechs Belastungsgruppen hat Auswirkungen auf diversen Ebenen: Auf der einen Seite determiniert der Sozialindex unterschiedliche Ressourcenallokationen (z. B. kleinere Klassen oder mehr Sprachförderressourcen für Schulen mit niedrigeren Indizes). Auf der anderen Seite wird der Sozialindex in weiteren Zusammenhängen genutzt: bei der Bildung von sinnvollen Vergleichsstichproben im Rahmen von wissenschaftlichen Untersuchungen und bei der Berechnung von Vergleichswerten („fairer Vergleich“) für die schulbezogenen Ergebnissrückmeldungen im Rahmen der Kompetenzmessung an Schulen.

Die soziale Belastung von Schulen wird anhand von 24 Variablen mit einer konfirmatorischen Faktorenanalyse berechnet (vgl. Schulte et al. 2014). In der Tradition der bisherigen Sozialindizes liegt der Variablenzusammensetzung ein theoretisches Modell zugrunde, welches verschiedene Aspekte der sozialen Belastung voneinander unterscheidet:

- soziales Kapital
- kulturelles Kapital
- ökonomisches Kapital
- Migrationshinweise

Die drei erstgenannten Facetten orientieren sich an Bourdieus Ansatz der Kapitalarten (1982, 1983), welcher ressourcenorientiert Differenzen in bestehenden sozialen Ungleichheiten und deren Reproduktion in den Familien der Schülerinnen und Schüler aber auch im Bildungssystem (vgl. Bourdieu & Passeron 1971) aufzeigt. Das soziale Kapital, welches auch von Coleman (1988) beschrieben wird, meint das Netzwerk sozialer Beziehungen, welches Personen über die Zugehörigkeit zu einer Gemeinschaft und die damit verbundenen Pflichten in ihrem Bildungserfolg unterstützt. Bourdieu beschreibt das „Sozialkapital“ als „die Gesamtheit der aktuellen und potentiellen Ressourcen, die mit dem Besitz eines dauerhaften Netzes von mehr oder weniger institutionalisierten Beziehungen gegenseitigen Kennens oder Anerkennens verbunden sind“ (Bourdieu 1992, S. 63). Das ökonomische Kapital bildet in Bourdieus Theorie die Bedeutung von Kapital im finanziellen Sinne ab, z. B. das Einkommen. Dagegen unterscheidet Bourdieu bei dem kulturellen Kapital drei Zustände:

1. Inkorporierter Zustand: Bildung der jeweiligen Person, die sie sich durch die Investition von Zeit angeeignet hat.
2. Objektivierter Zustand: kulturelle Güter, wie z. B. Bücher oder Kunstgegenstände.
3. Institutionalisierte Zustand: die Erlangung von Titeln im Bildungsverlauf, z. B. der Erwerb eines Schulabschlusses.

Des Weiteren werden zur Konstruktion des Sozialindex Migrationshinweise der Schülerinnen und Schüler herangezogen (vgl. Bensen et al. 2010), da diese – selbst unter Kontrolle weiterer sozio-ökonomischer Hintergrundmerkmale – noch immer Hinweise auf bedeutsame Disparitäten liefern (vgl. Klieme et al. 2010; IQB-Ländervergleich: Pant et al. 2013; Stanat et al. 2012).

Darüber hinaus wurden Daten des Statistikamts Nord zur Berechnung des Sozialindex verwendet. Dabei handelt es sich um drei Variablen (z. B. die Arbeitslosenquote), die auf Ebene der statistischen Gebiete in Hamburg vorliegen, in denen die Schülerinnen und Schüler einer Schule wohnen: „Statistische Gebiete sind kleinräumige Gebietseinheiten, die nach städtebaulichen und sozialstrukturellen Homogenitätskriterien im Anschluss an die Volkszählung 1987 gebildet wurden.“ (Freie und Hansestadt Hamburg 2012, S. 22). In diesen Einheiten leben durchschnittlich 2.100 Personen. Diese amtlichen Daten, die für alle Schülerinnen und Schüler einer Schule vorliegen, können gerade auch bei hohen Datenausfällen oder möglichen Verzerrungen durch selektives Beantworten wertvolle Informationen liefern (vgl. Pietsch et al. 2006).

1.2 Primäre Merkmale und Prozesse der Qualitätsentwicklung

Zur Abbildung der primären Merkmale und Prozesse der Schulqualität können die Daten der Schulinspektion Hamburg genutzt werden. Die normativen Grundlagen von Schulinspektionen sind im Beitrag von Ehren und Scheerens in diesem Buch ausführlich beschrieben und werden daher an dieser Stelle nicht betrachtet.

Jede Schulinspektion folgt dem in Abbildung 2 dargestellten standardisierten Ablauf und wird durch ein multiprofessionelles Team vorgenommen. Daran beteiligt sind jeweils zwei Inspektorinnen bzw. Inspektoren sowie eine Schulformexpertin oder ein Schulformexperte sowie bei berufsbildenden Schulen eine Wirtschaftsvertreterin oder ein Wirtschaftsvertreter.

Abbildung 2: Ablauf des Inspektionsverfahrens



Nach der Erstinformation und einem Vorgespräch in der Schule werden strukturelle Informationen (z. B. Anzahl der Lehrkräfte) sowie die von der Schule vorgelegten Dokumente (z. B. Leitbild) erfasst. Vor dem eigentlichen Schulbesuch füllen die Schülerinnen und Schüler, deren Eltern sowie die Pädagoginnen und Pädagogen und (bei berufsbildenden Schulen) die Ausbildungsbetriebe größtenteils onlinebasierte Fragebögen aus. Die Ergebnisse der strukturellen Daten, der Dokumentenanalyse und der Fragebogenerhebungen werden vor dem Unterrichtsbesuch innerhalb des Inspektionsteams ausgewertet. Der eigentliche Schulbesuch beinhaltet Inter-

views mit verschiedenen Gruppen sowie die Beobachtung und Beurteilung von Unterrichtssequenzen durch das Inspektionsteam. Zwei bis drei Wochen danach stellt die zuständige Teamleiterin bzw. der zuständige Teamleiter die Ergebnisse der Schulleitung und der Schulöffentlichkeit vor. Der Bericht, welchen die Schule nach Abschluss der Inspektion erhält, basiert auf der durch das Inspektionsteam vorgenommenen Bewertung von 14 Qualitätsmerkmalen (siehe Tabelle 1). Diese 14 Qualitätsmerkmale waren im ersten Inspektionszyklus aller Hamburger Schulen inhaltlich im bis zum Jahr 2012 gültigen Orientierungsrahmen für Hamburger Schulen (BSB 2006) verankert. Der Orientierungsrahmen Schulqualität basiert auf verschiedenen wissenschaftlichen Studien über die Effektivität von Schule und Unterricht (z. B. Teddlie & Reynolds 2000).

Tabelle 1: Qualitätsmerkmale der Schulinspektion Hamburg

1. Führung und Management	2. Bildung und Erziehung	3. Wirkungen und Ergebnisse
1.1 Führung wahrnehmen	2.1 Das schuleigene Curriculum gestalten	3.1 Zufriedenes Personal, zufriedene Schüler/innen, Eltern und Betriebe
1.2 Personal entwickeln	2.2 Unterrichten, Lernen, Erziehen	3.2 Bildungslaufbahnen und Kompetenzen
1.3 Finanz- und Sachmittel gezielt einsetzen	2.3 Organisatorische Rahmenbedingungen sichern	
1.4 Profil entwickeln und Rechenschaft ablegen	2.4 Leistungen beurteilen	
	2.5 Prozesse und Ergebnisse evaluieren	
	2.6 Förderkonzepte entwickeln	
	2.7 Beratungskonzepte gestalten	
	2.8 Die Schulgemeinschaft beteiligen	

1.3 Wirkungen und Ergebnisse der Qualitätsentwicklung im Bildungsbereich

Um den Bereich der Ergebnisse bzw. des Outputs in dem Modell zur Qualität und Qualitätsentwicklung im Bildungsbereich nach Ditton (s. Abbildung 1) zu operationalisieren, steht in Hamburg ein umfangreiches Kompetenzermittlungssystem zur Verfügung. Dieses Instrument namens KERMIT erweitert die bereits bundesweit durchgeführten Vergleichsarbeiten (VERA) in den Klassenstufen 3 und 8 um weitere Erhebungen in den Klassenstufen 2, 5, 7 und 9 (Lücken et al. 2014). Durch eine wiederholte Ermittlung von Kompetenzen im Verlauf des Bildungswegs wird die Dokumentation der individuellen Lernentwicklung der Schülerinnen und Schüler mittels „echter Längsschnitte“ ermöglicht. Zusätzlich werden diese Ergebnisse mit sozialen und kriterialen Bezugsnormen verglichen, die sich auf die Hamburger Schulen und die von der Behörde für Schule und Berufsbildung (BSB) verabschiedeten Bildungspläne beziehen.

Mit den KERMIT-Erhebungen werden zentrale Leistungsindikatoren erfasst, die ebenfalls in den nationalen und internationalen Schulleistungsuntersuchungen (z. B. PISA, IQB-Ländervergleich) zur Untersuchung der Leistungsfähigkeit von Bildungssystemen herangezogen werden (vgl. Pant et al. 2013; Stanat 2012). Zum Einsatz kommen standardisierte, normierte und aufeinander abgestimmte Schulleistungstests für die Fächer und Testbereiche Deutsch Leseverstehen, Mathematik, Englisch Hörverstehen sowie die Naturwissenschaften. Auf diese Weise werden objektive diagnostische Informationen gewonnen, mit denen die aktuellen schulischen Leistungen der Hamburger Schülerinnen und Schüler sowie die Entwicklung ihrer Leistungen im Verlauf der Schulzeit zuverlässig abgebildet werden. Da die KERMIT-Erhebungen in den Jahrgangsstufen 5, 7 und 9 durch externe Testleiterinnen und Testleiter administriert werden und so die Durchführungsobjektivität weitestgehend gesichert ist, lassen sich gerade aus diesen Ergebnissen besonders aussagekräftige Erkenntnisse über das Hamburger Bildungssystem und die Leistungsentwicklung an den Schulen gewinnen. Alle KERMIT-Erhebungen orientieren sich an den bundesweit verbindlichen Bildungsstandards und dienen der Schul- und Unterrichtsentwicklung. Die Bildungsstandards sind für zentrale Schnittstellen im Bildungsweg formuliert: Sie liegen für das Ende der Grundschulzeit (Ende der 4. Klassen) sowie für das Ende der Sekundarstufe I vor. Bei den KERMIT-Erhebungen, die in Hamburg für die Jahrgänge 2, 5, 7, und 9 entwickelt werden, werden die Aufgaben verstärkt mit den Hamburger Bildungsplänen abgeglichen, um die curriculare Validität dieser Tests und damit die Aussagekraft über die Leistungsentwicklung Hamburger Schülerinnen und Schüler zu erhöhen.

1.4 Fragestellung

Anhand der Daten zu den Voraussetzungen, unter denen Schulen arbeiten, zu den Prozessen schulischer Qualität und zu den Ergebnissen der Schülerinnen und Schüler soll im vorliegenden Beitrag der Frage nachgegangen werden, welche Merkmale schulischer Qualität unter gegebenen Rahmenbedingungen einen Unterschied im Hinblick auf die durchschnittliche Lernentwicklung an Schulen in Hamburg machen. Dazu werden die Schulen mit den höchsten Lernentwicklungen und die Schulen mit den niedrigsten Lernentwicklungen im Hinblick auf mögliche Unterschiede in der schulischen Prozessqualität verglichen. Diese Ergebnisse werden sodann für die gesamte Stichprobe unter Zuhilfenahme regressionsanalytischer Modelle und unter Berücksichtigung der Kontextmerkmale der Schulen noch einmal genauer überprüft.

2. Methode

2.1 Stichprobe

Bei der analysierten Stichprobe handelt es sich um 41 weiterführende Schulen (20 Stadtteilschulen und 21 Gymnasien) aus Hamburg, an denen in den Jahren 2010–2012 eine Schulinspektion stattgefunden hat. Die Jahre 2010–2012 wurden ausgewählt, da für diesen Zeitraum vergleichbare Bewertungskriterien (sog. „Bewertungsprofile“) der Schulqualität im Rahmen der Schulinspektion vorliegen. In Tabelle 2 sind die Sozialindizes der analysierten Schulen zum Zeitpunkt der Datenerhebung aufgeführt. Für die Analysen wird der intervallskalierte Rohwert sozialer Belastung verwendet, welcher die Grundlage für die sechsstufige Einteilung des Sozialindex bildet (vgl. Schulte et al. 2014).

Tabelle 2: Sozialindizes der 41 Schulen, insgesamt und differenziert nach Schulformen

Sozialindex	Schulen insgesamt	Stadtteilschulen	Gymnasien
1	3	3	0
2	3	3	0
3	4	4	0
4	9	7	2
5	10	2	8
6	12	1	11

Genutzt wurden weiterhin die Daten von 4.479 Schülerinnen und Schülern, die im Jahr 2011 zum Anfang der 5. Klasse und im Jahr 2013 am Ende der 7. Klas-

se an Leistungsvergleichstests im Rahmen von Vergleichsarbeiten in Hamburg (KERMIT) teilgenommen haben. Die Konzentration auf diese Klassenstufen und die dazu gehörigen KERMIT-Erhebungen bietet den Vorteil, die Leistungsentwicklung der Schülerinnen und Schüler von Beginn des Besuchs der weiterführenden Schule in Klasse 5 über 2 Jahre bis zum Abschluss der Beobachtungsstufe Ende 6/Anfang Klasse 7 beobachten zu können.

2.2 Instrumente

2.2.1 *Lernleistung – Kompetenzen ermitteln (KERMIT)*

Zur Untersuchung des Zusammenhangs zwischen den Ausprägungen der Qualitätsmerkmale der Schulinspektion Hamburgs und der Entwicklung der Lernleistung an den Hamburger Schulen werden, wie bereits erwähnt, die Daten der Schulen genutzt, die 2011 im Jahrgang 5 und 2013 im Jahrgang 7 an KERMIT teilgenommen haben.

Bei KERMIT 5 und 7 verteilen sich die einzelnen Aufgaben der vier Testbereiche auf verschiedene Testheftversionen. Für die beiden Schulformen Stadtteilschule und Gymnasium gibt es pro KERMIT-Erhebung jeweils zwei Parallelversionen, um ein Abschreiben zu vermeiden. Die mittlere Lösungshäufigkeit bei allen vier Testheften liegt bei ca. 55 Prozent. Diese vier Testhefte sind für jeden Testbereich mit mindestens zehn gleichen Items verankert. Gleichzeitig wurde auch darauf geachtet, dass bei den Testheften genau so viele Anker verwendet wurden, um die einzelnen Testbereiche im Längsschnitt von Jahrgang 5 nach Jahrgang 7 abbilden zu können. In der Tabelle 3 sind die wichtigsten Kennwerte für die einzelnen Testbereiche für beide KERMIT-Erhebungen abgebildet.

Tabelle 3: Skalenskennwerte für die KERMIT-Erhebungen im Jahrgang 5 und 7

KERMIT 5 (2011)					
Testbereich	Gesamtanzahl Items	Items pro Testheft	Innere Konsistenz (Cronbachs Alpha)	Aufgabenschwierigkeit (in Prozent)	Trennschärfe
Deutsch Leseverstehen	35 Items	26–28 Items	.81	20–95	.23–.63
Mathematik	32 Items	26 Items	.80	20–91	.23–.58
Englisch Hörverstehen	21 Items	21 Items	.80	33–87	.28–.63
Naturwissenschaften	34 Items	22–23 Items	.79	20–88	.24–.57

KERMIT 7 (2013)					
Testbereich	Gesamt- anzahl Items	Items pro Testheft	Innere Konsistenz (Cronbachs Alpha)	Aufgaben- schwierigkeit (in Prozent)	Trennschärfe
Deutsch Leseverstehen	47 Items	29–30 Items	.83	23–90	.28–.57
Mathematik	95 Items	29–31 Items	.88	17–90	.18–.79
Englisch Hörverstehen	34 Items	23 Items	.83	20–91	.20–.56
Naturwissen- schaften	59 Items	23–24 Items	.80	20–87	.20–.75

Die Leistungen in den verschiedenen Testbereichen wurden zunächst zu einem Gesamtwert für die Fachleistung zusammengefasst. Des Weiteren wurde aus den Fachleistungsindices für die beiden Klassenstufen 5 und 7 (nach Transformation der Werte auf die gleiche Metrik) ein Differenzwert gebildet, der die Lernentwicklung von Klasse 5 nach 7 darstellt.

2.2.2 Schulische Prozesse – die Hamburger Schulinspektion

Der Bericht, welchen die Schule nach Abschluss der Inspektion erhält, basiert auf der Beurteilung der Schule durch das Inspektionsteam hinsichtlich der vorgegebenen Bewertungskriterien (siehe Tabelle 1). Dabei werden 13 der 14 übergeordneten Qualitätsmerkmale anhand verschiedener Kriterien bewertet (siehe Anhang). Der Bereich 3.2 „Bildungslaufbahnen und Kompetenzen“ wurde aus den Inspektionen ausgeklammert, da für diesen Bereich im ersten Zyklus der Hamburger Schulinspektion keine validen Daten vorlagen und somit keine Bewertung des Qualitätsmerkmals möglich war. Im zweiten Zyklus werden für die Bewertung dieses Merkmals Daten zu Abschlussnoten, Abschlussquoten und Übergängen genutzt (siehe den Beitrag von Diedrich in diesem Band). Die Kriterien wiederum werden durch einzelne Indikatoren operationalisiert. Die Indikatorenbewertung selbst wird auf der Grundlage einer Triangulation aller vorliegenden Daten vorgenommen, d.h. je nach Themenbereich fließen Eindrücke der Ergebnisse aus den Fragebögen, den Angaben aus den Interviews, den Unterrichtsbeobachtungen und den Schuldokumenten ein. Alle zu einem Indikator vorliegenden Daten werden aus den unterschiedlichen Perspektiven des Inspektionsteams integriert bewertet. Aus dem Mittelwert der Bewertungen der Indikatoren wird im letzten Schritt konsensual im Inspektionsteam die Entscheidung für eine Bewertung der Qualitätsmerkmale auf einer Skala von 1 (schwach) bis 4 (stark) auf der Grundlage aller Daten und

Eindrücke auf Schulebene getroffen. Bei der Bewertung zählen nicht nur die erhobenen Daten, sondern auch die Expertise des Inspektionsteams. D.h., es kann in Einzelfällen vorkommen, dass die Daten eine bestimmte Bewertung nahelegen, diese jedoch anders vorgenommen wird, da sich durch die Expertendiskussion im Inspektionsteam ein abweichendes Bild der Schule ergibt.

3. Ergebnisse

3.1 Schulgruppenvergleiche

Um zu überprüfen, ob sich bestimmte Schulen mit einer hohen bzw. niedrigen Lernentwicklung der zuvor gebildeten Fachleistungsindizes pro Jahrgang systematisch im Hinblick auf die Qualitätsmerkmale der Schulinspektion unterscheiden, wurden zunächst zwei Gruppen bestimmt: Ausgewählt wurden jeweils die 25 Prozent der Schulen, die die höchste und die niedrigste durchschnittliche Lernentwicklung aufweisen. Die durchschnittliche Lernentwicklung der so ausgewählten 20 Schulen liegt zwischen 42.35 und 185.12 Punkten (siehe Tabelle 4). Da es sich bei 20 Schulen um eine sehr kleine Stichprobengröße handelt, werden nichtparametrische, exakte Tests durchgeführt.

Der Unterschied der Lernentwicklung zwischen den beiden Gruppen ist signifikant ($W = 55,00$; $p = .00$; $r = .85$). Jedoch ist die Streuung der Lernentwicklung in der Gruppe mit einer höheren Lernentwicklung sehr viel größer als in der Gruppe mit einer weniger hohen Lernentwicklung.

Tabelle 4: Deskriptive Statistik der zwei Leistungsgruppen

	Anzahl Schulen	<i>MW</i> LE	<i>SD</i>	Minimum	Maximum
Gruppe 1 (niedrige Lernentwicklung)	10	67.68	9.24	42.35	73.31
Gruppe 2 (hohe Lernentwicklung)	10	108.35	27.06	96.36	185.12

MW = Mittelwert; *LE* = Lernentwicklung; *SD* = Standardabweichung

3.2 Differentielle Qualitätsprofile der Schulgruppen

In Tabelle 5 sind die Mittelwerte der Qualitätsmerkmale in den beiden Schulgruppen aufgeführt, sowie die Signifikanz und Effektstärke der Gruppenunterschiede. Es ist ersichtlich, dass Gruppe 1 – die Schulen mit der niedrigeren durchschnittlichen Lernentwicklung – in 7 der 13 Qualitätsmerkmale nominell niedrigere Mittelwerte

als Gruppe 2 aufweist. Den relativ größten signifikanten Unterschied weist dabei das Qualitätsmerkmal „Die Schulgemeinschaft beteiligen“ auf. Im Mittel gelingt es der zweiten Gruppe von Schulen also besser, die Schulgemeinschaft angemessen an schulischen Prozessen zu beteiligen. Des Weiteren zeigen sich mittelgroße und auf einem 10-Prozent-Niveau signifikante Tendenzen zum Vorteil der zweiten Gruppe auf den Qualitätsmerkmalen „Profil entwickeln und Rechenschaft ablegen“ sowie „Das schuleigene Curriculum gestalten“. Die erste Gruppe der Schulen mit geringerer Lernentwicklung zeigt bei 3 der 13 Qualitätsmerkmale höhere Werte; dabei handelt es sich in zwei Fällen um mittelgroße Effekte bezüglich der Qualitätsmerkmale „Personal entwickeln“ und „Finanz- und Sachmittel gezielt einsetzen“. Beide Unterschiede werden nicht signifikant.

Tabelle 5: Gruppenspezifische Mittelwerte der Beurteilung der Qualitätsmerkmale sowie Effektstärken der Gruppenunterschiede

Qualitätsmerkmal	Gruppe 1	Gruppe 2	<i>W</i>	<i>p</i> (eins.)	<i>r</i>
1.1 Führung wahrnehmen	3	3.1	101.00	.485	0.08
1.2 Personal entwickeln	2.5	2.3	92.50	.175	0.25
1.3 Finanz- und Sachmittel gezielt einsetzen	3	2.8	96.00	.248	0.19
1.4 Profil entwickeln und Rechenschaft ablegen	2.8	3.4	84.00	.072	0.38
2.1 Das schuleigene Curriculum gestalten	2	2.4	87.00	.098	0.38
2.2 Unterrichten, Lernen, Erziehen	2.4	2.3	100.00	.500	0.10
2.3 Organisatorische Rahmenbedingungen sichern	3.2	3.4	95.00	.314	0.21
2.4 Leistungen beurteilen	2.3	2.3	105.00	.686	0.00
2.5 Prozesse und Ergebnisse evaluieren	2.1	2.3	94.50	.152	0.22
2.6 Förderkonzepte entwickeln	2.6	2.6	105.00	.675	0.00
2.7 Beratungskonzepte gestalten	3	3	105.00	.737	0.00
2.8 Die Schulgemeinschaft beteiligen	2.7	3.2	83.00	.035	0.49
3.1 Zufriedenes Personal, zufriedene Schüler/innen, Eltern und Betriebe	3.2	3.3	100.00	.500	0.11

W = Teststatistik Wilcoxon-W, *p* = Signifikanzniveau, *r* = approximatives Maß der Effektstärke

3.3 Regressionsanalytische Überprüfung der Zusammenhänge

Entgegen der Erwartung liegen zwischen den beiden Schulgruppen nur in Bezug auf ein Qualitätsmerkmal ein systematischer Unterschied und in Bezug auf zwei Qualitätsmerkmale Tendenzen vor. Es ist jedoch nicht anzunehmen, dass dies die zugrunde liegenden Systematiken in den Daten erschöpfend beschreibt. Aus diesem Grund sollen in einem nächsten Schritt die Besonderheiten der Daten in Regressionsanalysen stärker berücksichtigt werden:

1. Die für die Berechnungen genutzten Daten liegen auf verschiedenen Ebenen vor: Sowohl das Kontextmerkmal Sozialindex als auch die Qualitätsmerkmale der Schulinspektion als Prozessvariablen beziehen sich auf die Schulebene. Dagegen handelt es sich bei den Leistungsdaten um Daten, die auf Ebene der Schülerinnen und Schüler vorliegen. Es kann daher notwendig sein, systematische Einflüsse der Schulen im Rahmen eines Mehrebenenendesigns zu berücksichtigen (siehe Beitrag von Schwippert in diesem Band). Einen Hinweis darauf liefern die Intraklassenkorrelationen der Leistungswerte. Die Intraklassenkorrelation beträgt $ICC = .29$ für die Leistung im Bereich Deutsch in Klasse 5 und $ICC = .33$ in Klasse 7. Für den Bereich Mathematik beträgt die Intraklassenkorrelation $ICC = .32$ in Klasse 5 und $ICC = .36$ in Klasse 7. Folglich werden diese Einflüsse der Schulen in den nachfolgenden Analysen durch die Anwendung von Modellen auf zwei Ebenen (Individual- und Schulebene) berücksichtigt.
2. Statt eines manifesten Differenzwertes der Leistungen in Klasse 5 und Klasse 7 wird ein latenter Differenzwert im Rahmen eines Strukturgleichungsmodells berechnet. Dadurch ist es möglich, Messfehler zu berücksichtigen.
3. Statt eines Fachleistungsindex, in den verschiedene Fächer eingehen und in welchem sich gegenläufige Effekte aufheben könnten, werden die Ergebnisse einzeln für die Kernfächer Deutsch und Mathematik betrachtet.

Darüber hinaus werden die folgenden Analysen für die gesamte Stichprobe durchgeführt. Für die Regressionsanalysen wird nur der Bereich 2 „Bildung und Erziehung“ analysiert, da für diesen Bereich direktere Effekte auf die Lernentwicklung zu erwarten sind. Für die Analysen wird die Software Mplus (vgl. Muthén & Muthén 2012) genutzt.

Mathematik

Im Bereich Mathematik werden zunächst alle Qualitätsmerkmale des Bereichs 2 in einem Regressionsmodell subsummiert (siehe Tabelle 6, Modell 1). Es zeigen sich drei bedeutsame Effekte für die Qualitätsmerkmale „Das schuleigene Curriculum gestalten“, „Unterrichten, Lernen, Erziehen“ sowie „Die Schulgemeinschaft beteiligen“.

Tabelle 6: Regressionsanalytische Vorhersage der Lernentwicklung im Bereich Mathematik von Klasse 5 zu Klasse 7 durch Qualitätsmerkmale der Schulinspektion, Sozialindex und Schulform (Befunde aus Mehrebenenanalysen)

Variable	Modell 1		Modell 2		Modell 3	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
2.1 Das schuleigene Curriculum gestalten	<i>0.329</i>	0.131	<i>0.331</i>	0.129	0.185	0.265
2.2 Unterrichten, Lernen, Erziehen	<i>-0.498</i>	0.178	<i>-0.507</i>	0.167	-0.321	0.399
2.3 Organisatorische Rahmenbedingungen sichern	0.143	0.151	0.132	0.189	0.038	0.217
2.4 Leistungen beurteilen	-0.069	0.142	-0.078	0.147	-0.001	0.110
2.5 Prozesse und Ergebnisse evaluieren	-0.210	0.135	-0.217	0.147	-0.128	0.147
2.6 Förderkonzepte entwickeln	0.073	0.110	0.094	0.168	0.025	0.091
2.7 Beratungskonzepte gestalten	-0.139	0.157	-0.139	0.166	-0.059	0.151
2.8 Die Schulgemeinschaft beteiligen	<i>0.358</i>	0.146	<i>0.354</i>	0.152	0.221	0.298
Sozialindex			-0.095	0.555	-0.318	0.324
Schulform (Stadtteilschule = 1)					<i>-0.506</i>	0.146
R^2	0.321		0.370		0.819	

B = Standardisierter Regressionskoeffizient; *SE* = Standardfehler; R^2 = Varianzaufklärung. Signifikante Effekte sind kursiv ($p > .05$).

Dabei handelt es sich bei dem Zusammenhang des Bereichs „Unterrichten, Lernen, Erziehen“ um einen negativen Effekt. Die Effekte bleiben nach Hinzunahme der sozialen Belastung der Schule bestehen (Modell 2). Die soziale Belastung der Schule ist dabei kein signifikanter Prädiktor der Lernentwicklung. Im dritten Modell, in dem die Schulform inkludiert ist, verschwinden, bis auf ebendiese, alle

Effekte der Qualitätsmerkmale. Der signifikant negative Vorhersagewert der Variable Schulform zeigt an, dass die Lernentwicklung an Gymnasien größer ist als an Stadtteilschulen. Die Varianzaufklärung steigt im Verlauf der drei Modelle von 32 Prozent auf knapp 82 Prozent.

Deutsch

Im Bereich Deutsch (Tabelle 7) zeigen sich im ersten Modell, in dem die Qualitätsmerkmale der Schulinspektion inkludiert sind, die gleichen Effekte wie im Bereich Mathematik: Signifikante Prädiktoren weisen die Bereiche „Das schuleigene Curriculum gestalten“, „Unterrichten, Lernen, Erziehen“ (negativ) sowie der Bereich „Die Schulgemeinschaft beteiligen“ auf.

Tabelle 7: Regressionsanalytische Vorhersage der Lernentwicklung im Bereich Deutsch von Klasse 5 zu Klasse 7 durch Qualitätsmerkmale der Schulinspektion, Sozialindex und Schulform (Befunde aus Mehrebenenanalysen)

Variable	Modell 1		Modell 2		Modell 3	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
2.1 Das schuleigene Curriculum gestalten	<i>0.603</i>	0.111	0.211	0.140	0.094	0.115
2.2 Unterrichten, Lernen, Erziehen	<i>-0.611</i>	0.132	-0.240	0.148	-0.112	0.144
2.3 Organisatorische Rahmenbedingungen sichern	0.165	0.201	0.007	0.122	-0.020	0.045
2.4 Leistungen beurteilen	0.011	0.159	-0.040	0.062	-0.014	0.064
2.5 Prozesse und Ergebnisse evaluieren	0.081	0.112	0.058	0.054	0.040	0.061
2.6 Förderkonzepte entwickeln	-0.064	0.117	0.062	0.066	0.019	0.037
2.7 Beratungskonzepte gestalten	-0.110	0.122	-0.034	0.050	-0.022	0.034
2.8 Die Schulgemeinschaft beteiligen	<i>0.543</i>	0.107	0.155	0.122	0.067	0.122
Sozialindex			<i>-0.564</i>	0.144	<i>-0.459</i>	0.075
Schulform (Stadtteilschule = 1)					-0.237	0.156
<i>R</i> ²	<i>0.597</i>		<i>0.496</i>		<i>0.491</i>	

B = Standardisierter Regressionskoeffizient; *SE* = Standardfehler; *R*² = Varianzaufklärung. Signifikante Effekte sind kursiv (*p* > .05).

Diese Effekte verschwinden jedoch bei Hinzunahme der sozialen Belastung der Schule als Prädiktor. Die Hinzunahme der Schulform als Prädiktor verringert die Effekte zusätzlich. Dabei ist die Schulform selbst, im Gegensatz zum Bereich Mathematik, kein signifikanter Prädiktor. Die soziale Belastung zeigt im zweiten und dritten Modell einen signifikanten Effekt. D.h., je belasteter die Rahmenbedingungen sind, unter denen die Schulen arbeiten, desto schwächer die Lernentwicklung der Schülerinnen und Schüler im Bereich Deutsch. Zu beachten ist jedoch, dass die Varianzaufklärung von Modell 1 zu Modell 3 sinkt. Dieser Umstand kann auf eine zu hohe Komplexität des Modells hindeuten.

4. Diskussion

Zur Klärung der Frage, welche Merkmale schulischer Qualität einen Unterschied im Hinblick auf die durchschnittlichen Lernentwicklungen der Schülerinnen und Schüler machen, wurden in einem ersten Schritt zwei Gruppen von Schulen gebildet, die sich hinsichtlich der Lernentwicklungen in einem Fachleistungsindex voneinander unterscheiden: die Schulen in Gruppe 1 bilden die 25 Prozent der Schulen, deren durchschnittliche Lernentwicklung am niedrigsten ist; in Gruppe 2 sind die 25 Prozent der Schulen, deren durchschnittliche Lernentwicklung am höchsten ist, verortet. Dabei ist zunächst auffällig, dass die Streuung der Leistungswerte der Schulgruppe mit durchschnittlich höheren Lernentwicklungen sehr viel größer ist als die Streuung der Leistungswerte der Schulgruppe mit weniger hohen Lernentwicklungen. Die Unterschiede im Bereich großer Lernentwicklungen zwischen Schulen sind also stärker ausgeprägt, während sich die zehn Schulen mit weniger hohen Lernentwicklungen nicht sehr stark unterscheiden. In Bezug auf die unterschiedlichen Bewertungen der schulischen Prozesse zeigen sich mehrheitliche Vorteile der zweiten Gruppe: Diese Gruppe schneidet in 7 der 13 betrachteten Qualitätsdimensionen besser ab als die erste Gruppe; dabei handelte es sich in einem Fall um einen signifikanten, mittelgroßen Effekt, in zwei Fällen sind signifikante Tendenzen erkennbar. Der relativ gesehen größte Effekt besteht in Bezug auf das Merkmal „Die Schulgemeinschaft beteiligen“. Es gibt demnach Hinweise darauf, dass es Schulen, die eine höhere Lernentwicklung bei den Schülerinnen und Schülern erreichen, gut gelingt, alle Schulbeteiligten angemessen an schulischen Prozessen zu beteiligen. Darüber hinaus schneiden die Schulen dieser Gruppe besser bei den Qualitätsmerkmalen „Profil entwickeln und Rechenschaft ablegen“ und „Das schuleigene Curriculum gestalten“ ab. Damit gelingt es den Schulen der zweiten Gruppe zum einen besser, das Schulprogramm klar zu gestalten und Prioritäten zu setzen sowie bei den verschiedenen Akteuren darüber Rechenschaft abzulegen. Die Schulen der zweiten Gruppe haben außerdem klarere Curricula formuliert und diese institutionell,

auch im Hinblick auf eine fortwährende Weiterentwicklung, im schulischen Alltag verankert.

Insgesamt handelt es sich bei den Analysen zu Gruppenunterschieden auf Schulebene jedoch um eine Stichprobe geringer Fallzahl, sodass die Interpretationsmöglichkeiten dieses Datenausschnitts begrenzt sind; die vorliegenden Unterschiede sollten daher nicht überinterpretiert werden. Obwohl die Effektstärke ein von der Stichprobengröße unabhängiges Maß ist, müssten in Bezug auf den Unterschied, den schulische Prozesse bei Lernleistungen erzeugen können, im zweiten Zyklus der Schulinspektion Hamburg weitere Untersuchungen in Richtung einer höheren Repräsentativität vorgenommen werden. Mit der Weiterentwicklung von KERMIT, dem Aufwachsen der KERMIT-Daten in Hamburg und dem Vorhandensein weiterer Daten zu Lernentwicklungen über mehrere Jahre können sich weitere Studien anschließen, die wiederum auch zur Validierung der Bewertungsgrundlage der Hamburg Schulinspektion hinsichtlich der Ergebnisse schulischer Prozesse dienen können.

In einem weiteren Schritt wurden die Ergebnisse noch einmal anhand der Gesamtstichprobe und unter Berücksichtigung methodischer Spezifika der Stichprobe regressionsanalytisch analysiert. Dabei wurden sowohl die den Daten zugrunde liegenden mehreren Ebenen, messfehlerbedingte Einflüsse, der soziale Kontext der Schule, institutionelle Effekte durch die Schulform sowie die zwei Kernfächer Mathematik und Deutsch getrennt berücksichtigt. Für die Analysen wurde der Bereich 2 der Schulinspektion „Bildung und Erziehung“ herangezogen, da in diesem Bereich die größten Effekte auf die Lernentwicklung zu vermuten sind. In mehreren Schritten wurden zunächst die acht Qualitätsmerkmale des Bereichs 2, die soziale Belastung der Schule sowie die Schulform (Gymnasium vs. Stadtteilschule) in Regressionsmodellen berücksichtigt. In beiden Fächern waren die Qualitätsmerkmale „Das schuleigene Curriculum gestalten“, „Unterrichten, Lernen, Erziehen“ sowie „Die Schulgemeinschaft beteiligen“ signifikante Prädiktoren der Lernentwicklung. Für das Merkmal „Unterrichten, Lernen, Erziehen“ zeigte sich ein erwartungswidrig negativer Regressionskoeffizient. Nach Berücksichtigung des Sozialindex blieben die Effekte im Bereich Mathematik bestehen, im Bereich Deutsch verschwanden die Effekte. Im jeweiligen dritten Modell, in dem zusätzlich die Schulform inkludiert war, verschwanden die signifikanten Effekte der Qualitätsmerkmale auch im Bereich Mathematik zugunsten der Schulform.

Damit bleiben, im Vergleich zu den gruppenanalytischen Berechnungen im ersten Teil die systematischen Unterschiede in Bezug auf die Qualitätsmerkmale „Die Schulgemeinschaft beteiligen“ sowie „Das schuleigene Curriculum gestalten“ zunächst bestehen. Der negative Einfluss des Bereichs „Unterrichten, Lernen, Erziehen“ war in den Gruppenanalysen nicht erkennbar.

Zur Erklärung der zum Teil erwartungswidrigen Zusammenhänge sowie den – nach Berücksichtigung des Sozialindex und der Schulform nicht mehr vor-

handenen – Effekten der Qualitätsmerkmale lassen sich verschiedene Vermutungen anstellen:

- Verschiedenste weitere Faktoren können einen Einfluss auf die Lernleistung haben, die im Rahmen der vorliegenden Studie nicht kontrolliert wurden, zum Beispiel außerschulische Lerngelegenheiten, die Ausstattung der Schulen oder die Qualifikation der Lehrkräfte. D.h., bei der Analyse von Durchschnittswerten von Schulen können sich teilweise individuelle sehr positive und sehr negative Lernentwicklungen überlagern.
- Durch die Schulform und den Sozialindex wird ein großer Teil der Varianz in den Daten zur Lernentwicklung erklärt, sodass der verbliebene Teil für weitere systematische Effekte der Qualitätsmerkmale nicht ausreicht. Dies müsste in Studien mit größeren Stichproben weiter untersucht werden.
- Erste Analysen, in denen Interaktionen zwischen der sozialen Belastung und den Qualitätsmerkmalen sowie die Leistungsheterogenität der Lernvoraussetzungen der Schülerinnen und Schüler berücksichtigt wurden, brachten keine verwendbaren Ergebnisse. Diesem Pfad sowie der Überprüfung indirekter Effekte, Moderations- und Mediationsvariablen müsste jedoch weiter nachgegangen werden.
- Die Kontrollvariablen „soziale Belastung der Schule“ sowie „Schulform“ sind generell, so auch in der vorliegenden Stichprobe, stark konfundiert (vgl. Tabelle 2). Das Verschwinden vorhandener Effekte zugunsten von institutionellen Effekten bzw. Effekten der sozialen Belastung zeigt sich auch in weiteren Studien: Bei Dumont et al. (2013) wurde eine hohe Konfundierung der Variablen „Schulform“ und in diesem Fall „Komposition der Schülerschaft“ festgestellt, die Effekte der Komposition verschwanden nach Berücksichtigung der Schulform. Die Autoren vermuten u.a. einen zu geringen zeitlichen Abstand zwischen den Erhebungen (5. und 6. Klasse). Diese Begründung lässt sich bei den vorliegenden Ergebnissen ausschließen, da es sich um einen Zeitraum von zwei Jahren handelte.
- Der Datenstruktur mit verschiedenen Schulformen angemessenere Verfahren wären beispielsweise Multigruppenanalysen, in denen zunächst die Invarianz bzw. Varianz der Datenstruktur überprüft würde, um in der Folge differentielle Effekte innerhalb der Schulformen zu analysieren. Dazu erforderlich wäre jedoch eine größere Stichprobe auf der zweiten Ebene der Modelle. Auch das sogenannte *Propensity Score Matching* unter Einbezug wichtiger Kontrollvariablen (vgl. Beitrag von Pietsch und Ehmke diesem Band) liefert vermutlich genauere Ergebnisse.

Der erwartungswidrige negative Zusammenhang zwischen dem Qualitätsmerkmal „Unterrichten, Lernen, Erziehen“ und der Lernentwicklung überrascht. Das Merk-

mal basiert vor allem auf den Beobachtungen der Unterrichtssequenzen, die das Inspektorenteam an den Schulen durchführt. Dem zugrunde liegt ein Modell der Unterrichtsqualität (vgl. Pietsch et al. 2009), welches auf Basis einer Raschskalierung vier Ebenen der Qualität unterscheidet:

1. Lernklima und pädagogische Strukturen sichern
2. Klassen effizient führen und Methoden variieren
3. Schüler motivieren, aktives Lernen und Wissenstransfer ermöglichen
4. Differenzieren, Schüler wirkungs- und kompetenzorientiert fördern

Je höher der Wert für die Unterrichtsqualität, desto eher wird also differenziert und wirkungsorientiert gefördert. Es lässt sich vermuten, dass Stufe 1, in der es vor allem um basale Aspekte des *Classroom Managements* und die Sicherung von Regeln geht, in den vorliegenden Analysen in dem Sinne Einfluss nimmt, dass an stärker belasteten Schulen auf die motivationalen und Lernvoraussetzungen der Schülerinnen und Schüler reagiert und der Unterricht dementsprechend angepasst wird. Dies geschieht durch die starke Übernahme von Erziehungsaufgaben im Unterricht. Ein Hinweis auf einen solchen zugrunde liegenden Mechanismus liefert die Reduktion des negativen Effekts der Unterrichtsqualität auf die Lernentwicklung durch die Hinzunahme kompositioneller Merkmale wie der Schulform und des Sozialindex. Um in diesem Rahmen belastbare Aussagen treffen zu können, müssten jedoch weiterführende Analysen im Sinne der oben beschriebenen Multi-gruppenmodelle berechnet werden.

Insgesamt zeigt die vorliegende Studie, dass die Zusammenhänge zwischen Merkmalen der schulischen Prozessqualität und der Lernentwicklung der Schülerinnen und Schüler nicht so einfach geartet sind, wie ursprünglich zu vermuten wäre. So lässt sich die schlichte, aber plausible kausale Annahme „Je besser der Unterricht, desto größer die Lernentwicklung der Schülerinnen und Schüler“ hier nicht bestätigen. Vielmehr müssen, wie bereits dargelegt, weitere Einflüsse auf die Lernleistung in Folgestudien Berücksichtigung finden und die Komplexität der Datenstruktur unter Hinzunahme von Interaktionseffekten beachtet werden. Dazu sollten das Untersuchungsdesign und die Stichprobe, insbesondere die Anzahl der Schulinspektionsdaten, angereichert werden. Dennoch zeigen die hier gefundenen Effekte erste Hinweise darauf, dass einzelne Faktoren der schulischen Prozessqualität – selbst bei dieser „einfachen“ Verknüpfung der Schulqualitätsdaten mit den Leistungsdaten der Schülerinnen und Schüler – auf die Lernentwicklung durchschlagen können. Dieser Ansatz soll und wird in zukünftigen Studien weiter verfolgt werden.

Literatur

- BBS (2006). *Orientierungsrahmen: Qualitätsentwicklung an Hamburger Schulen*. Hamburg: Behörde für Bildung und Sport.
- Berkemeyer, N. & van Holt, N. (2012). Leistungsrückmeldungen im Längsschnitt. Erste Erfahrungen mit dem Schüler-Monitoring-System (SMS). In A. Wacker, U. Maier & J. Wissinger (Hrsg.), *Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung* (S. 109–130). Wiesbaden: Springer VS.
- Bonsen, M., Bos, W., Gröhlich, C., Harney, B., Imhäuser, K., Makles, A., Schräpler, J.-P., Terpoorten, T., Weishaupt, H. & Wendt, H. (2010). *Zur Konstruktion von Sozialindizes. Ein Beitrag zur Analyse sozialräumlicher Benachteiligung von Schulen als Voraussetzung für qualitative Schulentwicklung* (Bildungsforschung, Bd. 31). Bonn: Bundesministerium für Bildung und Forschung.
- Bos, W., Pietsch, M., Gröhlich, C. & Janke, N. (2006). Ein Belastungsindex für Schulen als Grundlage der Ressourcenzuweisung am Beispiel von KESS 4. In W. Bos, H.G. Holtappels, H. Pfeiffer, H.-G. Rolff & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven* (Bd. 14, S. 149–160). Weinheim: Juventa.
- Bourdieu, P. (1982). *Die feinen Unterschiede – Kritik der gesellschaftlichen Urteilskraft*. Frankfurt a. M.: Suhrkamp.
- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (Soziale Welt, Sonderbd. 2, S. 183–198). Göttingen: Schwartz & Co.
- Bourdieu, P. (1992). Ökonomisches, soziales und kulturelles Kapital. In ders. (Hrsg.), *Die verborgenen Mechanismen der Macht* (S. 49–75). Hamburg: VSA Verlag.
- Bourdieu, P. & Passeron, J.-C. (1971). *Die Illusion der Chancengleichheit: Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs*. Klett: Stuttgart.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94 (1), 95–120.
- Ditton, H. (2007). Schulqualität – Modelle zwischen Konstruktion, empirischen Befunden und Implementierung. In J. van Buer & C. Wagner (Hrsg.), *Qualität von Schule. Ein kritisches Handbuch* (S. 83–92). Frankfurt a. M.: Lang.
- Dumont, H., Neumann, M., Nagy, G., Becker, M., Rose, N. & Trautwein, U. (2013). Einfluss der Klassenkomposition auf die Leistungsentwicklung in Haupt- und Realschulen in Baden-Württemberg. *Psychologie in Erziehung und Unterricht*, 60, 198–213.
- Freie und Hansestadt Hamburg (Hrsg.) (2012). *Sozialmonitoring Integrierte Stadtteilentwicklung – Bericht 2012*. Hamburg: Behörde für Stadtteilentwicklung und Umwelt.
- Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., Prenzel, M., Schneider, W. & Stanat, P. (Hrsg.) (2010). *PISA 2009. Bilanz nach einem Jahrzehnt*. Münster: Waxmann.

- Kühle, B. & Ackeren, I. van (2012). Wirkungen externer Evaluationsformen für eine evidenzbasierte Schul- und Unterrichtsentwicklung. In M. Ratermann & S. Stöbe-Blossey (Hrsg.), *Governance von Schul- und Elementarbildung – Vergleichende Betrachtungen und Ansätze der Vernetzung* (S. 45–62). Wiesbaden: Springer VS.
- Kuhn, H.-J. (2014). Anspruch, Wirklichkeit und Perspektiven der Gesamtstrategie der KMK zum Bildungsmonitoring. *Die Deutsche Schule*, 106 (4), 414–426.
- Lange, H. (2008). Vom Messen und Handeln: „empirische Wende“ der Bildungspolitik? *Recht der Jugend und des Bildungswesens*, 56 (1), 7–15.
- Lücken, M., Thonke, F., Pohlmann, P., Hofmann, H., Golecki, R., Rosendahl, J., Benzing, M. & Poerschke, J. (2014) KERMIT – Kompetenzen ermitteln. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine daten- und theoriegestützte Schulentwicklung. Konzeption und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)* (S. 127–154). Münster: Waxmann.
- Muthén, L. K. & Muthén, B. O. (2012). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (2013). *IQB-Ländervergleich 2012 – Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Münster: Waxmann.
- Pietsch, M., Bensen, M. & Bos, W. (2006). Ein Index sozialer Belastung als Grundlage für Rückmeldungen von Leistungsergebnissen an Schulen und Klassen und für ‚faire Vergleiche‘ von Grundschulen in Hamburg. In W. Bos & M. Pietsch (Hrsg.), *KESS 4 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen* (S. 225–245). Münster: Waxmann.
- Pietsch, M., Schnack, J. & Schulze, P. (2009). Unterricht zielgerichtet entwickeln. Die Hamburger Schulinspektion entwickelt ein Stufenmodell für die Qualität von Unterricht. *Pädagogik*, 61 (2), 38–43.
- Schulte, K., Hartig, J. & Pietsch, M. (2014). Der Sozialindex für Hamburger Schulen. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine daten- und theoriegestützte Schulentwicklung. Konzeption und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)* (S. 67–80). Münster: Waxmann.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik*. Münster: Waxmann.
- Teddlie, C. & Reynolds, D. (2000). *The International Handbook of School Effectiveness Research*. Falmer Press: New York.

Anhang: Bewertungskriterien

1.1 Führung wahrnehmen

- Führungsverhalten: Ziele, Leitbild und QM-Handeln
- Qualitätsmanagement: Verfahren und Instrumente (im engeren Sinne)
- Führungsverhalten: Organisationsstruktur
- Kooperation mit externen Partnern

1.2 Personal entwickeln

- Neue Mitarbeiter
- Personal für herausgehobene Aufgaben
- Personalführung (allgemein)
- Begleitung der persönlichen Entwicklung des Einzelnen
- Fortbildungskonzept
- Teamarbeit – Intention der SL
- Strukturen zur Bearbeitung von Konflikten

1.3 Finanz- und Sachmittel gezielt einsetzen

- Gebäude, Räume, Gelände
- Verwaltung der Mittel

1.4 Profil entwickeln und Rechenschaft ablegen

- Profil
- Schulentwicklungsarbeit
- Öffentlichkeitsarbeit

2.1 Das schuleigene Curriculum gestalten

- Curricula

2.2 Unterrichten, Lernen, Erziehen

- Gestaltung von Lehr-Lernprozessen
- Übergreifende Aspekte bei der Gestaltung von Lehr-Lernprozessen
- Erziehung
- Teamarbeit und Kooperation im Kollegium

2.3 Organisatorische Rahmenbedingungen sichern

- Organisation des Regelunterrichts
- Organisation der Klassenzusammensetzung

2.4 Leistungen beurteilen

- Kriterien
- Transparenz
- Beteiligung

2.5 Prozesse und Ergebnisse evaluieren

- Evaluation im Hinblick auf Unterrichtsentwicklung
- Evaluation im Hinblick auf Schulentwicklung
- Evaluation weiterer relevanter Themen

2.6 Förderkonzepte entwickeln

- Rahmen
- Maßnahmen

2.7 Beratungsangebote gestalten

- Schulinterne Beratungsangebote
- Dokumentation und Rechenschaftslegung
- Information

2.8 Die Schulgemeinschaft beteiligen

- Beteiligung von Eltern
- Beteiligung des Elternrates
- Beteiligung der Schülerinnen und Schüler an Lernprozessen
- Beteiligung der Schülerinnen und Schüler am Schulleben
- Beteiligung des Schülerrates

3.1 Zufriedenes Personal, zufriedene Schüler, Eltern und Betriebe

Eltern als Stakeholder von Schule

Erkenntnisse über die Sichtweise von Eltern durch die Hamburger Schulinspektion

Susanne Peters

Zusammenfassung

Die Notwendigkeit der Wahrnehmung von Eltern als Stakeholdern von Schule lässt sich aus unterschiedlichen theoretischen Ansätzen herleiten: aus der Perspektive einer marktorientierten Steuerung des Bildungswesens, der Schulentwicklungsforschung und der systemorientierten Vertrauensforschung. Welche Erkenntnisse zur Elternperspektive auf Schule lassen sich aus Daten der Schulinspektion gewinnen? Inwieweit trägt Schulinspektion dazu bei, dass durch die Außensicht von Eltern auch blinde Flecke im System Schule entdeckt und diese im Rahmen von Schulentwicklungsprozessen bearbeitet werden? Ausgehend von unterschiedlichen theoretischen Ansätzen werden ausgewählte Daten der Hamburger Schulinspektion aus den Jahren 2007 bis 2014 dargestellt und interpretiert, die in schriftlichen Elternbefragungen und in Gruppeninterviews mit Eltern gewonnen wurden. Mit ihnen lassen sich Zusammenhänge zwischen „guter Schule“ und „guter Elternarbeit“ aufzeigen und Entwicklungsfelder für Schulen bestimmen. Darüber hinaus lässt sich Steuerungswissen für Verwaltung und Bildungspolitik bezüglich der Wahrnehmung von Eltern als Stakeholdern generieren.

1. Eltern als Stakeholder im Bildungssystem

Die Sichtweise von Eltern auf schulische Bildung ist ein Thema, das im Forschungskontext in Deutschland anders als im internationalen Raum bislang eine untergeordnete Rolle spielt (vgl. Sacher 2010). So wurde im Rahmen der OECD-Studie „PISA 2006 – Schulleistungen im internationalen Vergleich“ eine Begleitstudie durchgeführt, die sich mit der Rolle der Eltern beschäftigt und in deren Rahmen auch Eltern zu ihrer Zufriedenheit befragt wurden (vgl. OECD 2007, S. 271 ff.). Die Ergebnisse dieser Elternbefragungen zeigen, dass Eltern in Deutschland deutlich weniger von Schulen als Partner und Unterstützer von Bildungsprozessen wahrgenommen werden als in anderen Ländern (vgl. ebd., S. 275). Erst in letzter Zeit beginnt eine Auseinandersetzung mit der Sichtweise von Eltern auf Schule (vgl. Jäger-Flor, D. & Jäger, R. 2010; Killus & Tillmann 2011, 2012; Vodafone-

Stiftung Deutschland 2013). Eine stärkere Berücksichtigung von Eltern lässt sich aus unterschiedlichen Begründungszusammenhängen herleiten. Theoretische Ansatzpunkte zum Einfluss von Eltern auf Schule ergeben sich aus der Perspektive einer marktorientierten Steuerung, aus der Schulentwicklungsforschung vor allem im angloamerikanischen Raum sowie aus der systemtheoretisch ausgerichteten Vertrauensforschung. Ausgehend von diesen unterschiedlichen Ansätzen werden Daten der Hamburger Schulinspektion aus den Jahren 2007 bis 2014 dargestellt und interpretiert und Eltern als Stakeholder im Bildungssystem ins Bewusstsein gerückt. Hieraus lassen sich Erkenntnisse für zukünftige Schulentwicklungsprozesse und Steuerungswissen generieren.

1.1 Theoretische Annahmen aus Sicht der marktorientierten Systemsteuerung

Die globale Umsteuerung im Bildungswesen beinhaltet nach Fend (2008, S. 110 ff.) neben einer Outputorientierung eine Orientierung an Marktmechanismen wie Angebot und Nachfrage, in deren Rahmen auch in Deutschland Eltern als Stakeholder zunehmend an Bedeutung gewinnen. Fend stellt Überlegungen an, inwieweit Eltern zum Ziel der strategischen Überlegungen zu Investitionen gemacht werden sollten, etwa über Gutscheinsysteme, mit dem Eltern Bildung für ihre Kinder einkaufen (vgl. ebd., S. 131). Bereits mit den derzeit erweiterten Möglichkeiten von Eltern, Einfluss auf Schulentscheidungen zu nehmen, etwa über die Anwahl einzelner Schulen und/oder der Schulform (vgl. OECD 2007, S. 271), ändern sich die Herausforderungen für Schulen. Hirschman (1970) hat ein Modell entwickelt, mit welchem Entscheidungen von Konsumenten und deren Auswirkungen auf Organisationen beschrieben werden können, das von Matland erweitert worden ist (vgl. Matland 1995, S. 506 f.). Ausgehend von den Überlegungen Friedmans (1955), Erkenntnisse aus der Wirtschaft auf den Bildungsbereich und dessen Steuerung zu übertragen, etwa indem Schulentscheidungen von Eltern als Marktmechanismen genutzt werden, überträgt Matland das von Hirschman für Wirtschaftsunternehmen entwickelte Modell auf das Bildungssystem (vgl. Matland 1995, S. 507). Er unterscheidet die Strategien „Exit“ und „Voice“. *Exit* beinhaltet den Wechsel zwischen Produkten im Falle von Unzufriedenheit, der ein Unternehmen dazu zwingt, auf die Abwanderung seiner Kunden zu reagieren. Dies entspricht im Schulsystem der Steuerung über Schulwahlentscheidungen von Eltern, die gegebenenfalls aus den staatlichen Schulen abwandern und auf das Privatschulwesen ausweichen, um einer für sie unbefriedigenden Situation zu entkommen. *Voice* stellt die Nutzung von positiven und kritischen Rückmeldungen der Kunden für die Weiterentwicklung des Unternehmens dar. Auf das Schulsystem bezogen können Schulentwicklungsprozesse durch Feedback von Eltern und Schülerinnen und Schülern initiiert und

befördert werden. Nach Matland (ebd., S. 509) wird die Möglichkeit, die Schule zu wechseln bzw. in das Privatschulwesen abzuwandern, nur von sehr wenigen Eltern genutzt, wie sich in einer empirischen Studie gezeigt hat. Das *Exit-Voice*-Modell von Hirschman ist von Withey und Cooper (1989, zitiert nach Matland 1995, S. 510) um zwei eher auf Passivität angelegte Handlungsstrategien ergänzt worden: Loyalität (*Loyalty*) und Rückzug (*Neglect*). Loyalität wird dabei als Verbleib in der Organisation bei gleichzeitiger Unterstützung verstanden, Rückzug steht für den Verbleib in der Organisation bei gleichzeitiger Reduktion des Engagements auf ein Minimum (vgl. ebd., S. 510). Matland überträgt dieses erweiterte *Exit-Voice-Loyalty-Neglect*-Modell auf das Bildungssystem und die Schulentscheidungen, die von Eltern getroffen werden. Er analysiert, inwieweit aus dem Blickwinkel einer marktorientierten Systemsteuerung des Schulsystems die Nutzung unterschiedlicher elterlicher Strategien zu besseren Schulen führen, etwa durch bewusst getroffene Schulentscheidungen wie Abwanderung an Privatschulen (vgl. ebd., S. 511). Nach seinen Erkenntnissen ist *Exit*, also der Ausstieg aus dem öffentlichen Schulsystem für 99 Prozent der Eltern keine Option. Die meisten Eltern, die Schulentscheidungen innerhalb des öffentlichen Schulsystems treffen, nutzen als Strategien Loyalität und Rückzug, durch die jedoch kein Veränderungsdruck entsteht. Wenn die *Exit*-Strategie von Eltern nur selten genutzt wird und daher auf Schulen kaum Veränderungsdruck ausgeübt wird, dann ist aus seiner Sicht zu überlegen, wie die *Voice*-Strategie als Steuerungsinstrument im Bildungssystem genutzt werden kann, um an Schulen Veränderungsprozesse in Gang zu setzen und Rückmeldungen von Eltern in Schulentwicklungsprozesse einfließen zu lassen.

1.2 Erkenntnisse der Schulentwicklungsforschung zum Einfluss von Eltern auf den Bildungserfolg

Forschungsergebnisse aus dem angloamerikanischen Raum zeigen den Beitrag von Eltern zu einer verbesserten Qualität von Schule und Unterricht (sowohl auf Einzelschul- als auch auf Systemebene) und weisen auf die Bedeutung von Eltern als Stakeholdern von Schule hin (vgl. Epstein 2001; Leithwood et al. 2004; Fullan 2001). Die Einbeziehung von Eltern in die Gestaltung der Lernprozesse ist ausschlaggebend für die Steigerung von schulischen Leistungen der Schülerinnen und Schüler (vgl. Epstein 2001, S. 35; Fullan 2001). Ausgehend unter anderem von den Erkenntnissen Epsteins haben Hoover-Dempsey und Sandler (1995, 1997, 2005) ein Stufenmodell entwickelt, mit dem der Einfluss elterlichen Engagements auf die Leistungen der Schülerinnen und Schüler beschrieben werden kann. Es zeigt Schulen Ansatzpunkte für eine stärkere Einbeziehung von Eltern (vgl. Tabelle 1).

Tabelle 1: Das Stufenmodell zur elterlichen Einbeziehung in Schule

Stufe 5	Leistungen der Schülerin/ des Schülers	<ul style="list-style-type: none"> • Fähigkeiten und Kenntnisse • Gefühl der Selbstwirksamkeit bezüglich erfolgreichen Lernens
Stufe 4	Nutzungs- und Vermittlungsvariablen	<ul style="list-style-type: none"> • Passung zwischen elterlichem Engagement und schulischen Erwartungen
Stufe 3	Mechanismen, durch die elterliches Engagement die schulischen Leistungen des Kindes beeinflusst	<ul style="list-style-type: none"> • Modellieren, Verstärken, Instruktion als elterliche Unterstützung des Lernprozesses
Stufe 2	Die elterliche Entscheidung für eine Beteiligungsform	<p>Wird beeinflusst durch:</p> <ul style="list-style-type: none"> • Bestimmte Bereiche des elterlichen Könnens und Wissens • Anfragen, die mit der den Eltern zur Verfügung stehenden Zeit in Einklang sind (Familie, Berufstätigkeit) • Besondere Einladungen der Schule und konkrete Bitten um Mitwirkung
Stufe 1	Grundlegende Entscheidung von Eltern zur Beteiligung	<p>Wird beeinflusst durch:</p> <ul style="list-style-type: none"> • die elterliche Konstruktion der elterlichen Rolle • das elterliche Gefühl der Selbstwirksamkeit, das Kind zu unterstützen • Allgemeine Einladungen der Schule und die Bitte um Mitwirkung

Quelle: Hoover-Dempsey & Sandler 1997, S. 4

Hoover-Dempsey und Sandler (1995, S. 322) zeigen mit ihrem Modell auf, dass sowohl elterliche als auch schulische Faktoren Einfluss auf die elterliche Beteiligung und damit auf Schülerleistungen haben. Die Passung zwischen elterlichen Aktivitäten und den schulischen Erwartungen an elterliches Engagement ist dabei als Nutzungs- bzw. Vermittlungsvariable besonders relevant (s. Stufe 4 des Modells).

1.3 Erkenntnisse aus der Vertrauensforschung

Eine weitere theoretische Grundlegung zum Verhältnis Eltern und Schule findet sich in der aktuellen Vertrauensforschung unter Bezug auf systemtheoretische Überlegungen Luhmanns zum Vertrauen (2000). Die Bedeutung des Entstehens eines Vertrauensverhältnisses zwischen Klienten und Akteuren und die daraus resultierenden Konsequenzen für Organisationen und professionelle Akteure wurden für den Bereich der Sozialarbeit von Wagenblass (2004) und Tiefel & Zeller

(2012), für den schulischen Bereich zwischen Eltern und schulischen Akteuren im Bildungssystem von Peters (2014, 2015a, 2015b) aufgezeigt. Ausgehend von den erweiterten Entscheidungsmöglichkeiten von Eltern im Zuge der bildungspolitischen Umsteuerung verändert sich die Beziehung zwischen Eltern und Schule. Eltern geben der von ihnen gewählten Schule einen Vertrauensvorschuss, der von den schulischen Akteuren zunächst beantwortet werden muss. Kommunikation zwischen professionellen Akteuren und Eltern hat hierbei einen besonderen Stellenwert für das Entstehen einer Vertrauensbeziehung (vgl. Peters 2015b, S. 259 f.).

Im Rahmen einer in Hamburg durchgeführten qualitativen Forschungsstudie äußern Eltern Kritik an der professionellen Gestaltung der Kommunikation zwischen Schule und Eltern, und zwar bezogen auf die Transparenz, Partizipation und die Gestaltung des Interaktionsprozesses. Fehlende Transparenz manifestiert sich etwa in fehlenden Informationen über den schulischen Alltag oder über Lernziele. Der Aspekt Partizipation wird bedeutsam durch fehlende Einbeziehung der elterlichen Expertise in die Lernplanung. Die Gestaltung des Interaktionsprozesses wird als unprofessionell erlebt, etwa durch die Wahrnehmung fehlender Vorbereitung auf Gespräche oder das Signalisieren von Anstrengung und Überforderung durch die Lehrkräfte (vgl. ebd., S. 260). Durch die Art der Kommunikation zwischen Schule und Eltern entsteht neben einem informationsbasierten Vertrauen bzw. Misstrauen ein interaktionsbasiertes Vertrauen bzw. Misstrauen von Eltern (vgl. Tabelle 2).

Tabelle 2: Informationsbasiertes und interaktionsbasiertes Vertrauen

Informationsbasiertes Vertrauen durch Bereitstellen von Informationen	Interaktionsbasiertes Vertrauen durch Austausch von Informationen
<ul style="list-style-type: none"> • Transparenz durch Dokumentation der Lernentwicklung und Zielklarheit 	<ul style="list-style-type: none"> • Partizipation durch gemeinsame Gespräche über die Lernentwicklung
<ul style="list-style-type: none"> • Transparenz der Lernziele gegenüber den Eltern 	<ul style="list-style-type: none"> • Partizipation durch ein gemeinsames Verständnis von Lernen als gemeinsamem Anliegen von Eltern und schulischen Akteuren
<ul style="list-style-type: none"> • Transparenz durch Rückmeldung des aktuellen Lernstands an die Eltern 	<ul style="list-style-type: none"> • Professionelle Gestaltung des Interaktionsprozesses mit den Eltern durch die schulischen Akteure (Signalisieren von Problemlösekompetenz und Gesprächskompetenz)

Interaktionsbasiertes Vertrauen/Misstrauen gründet sich auf der Zufriedenheit bzw. Unzufriedenheit der Eltern mit den Kommunikationsprozessen zwischen ihnen und den schulischen Akteuren im Rahmen von Face-to-Face-Interaktionen, also

persönlichen Begegnungen (vgl. ebd., S. 261). Es stellt sich vor dem Hintergrund dieser Erkenntnisse die Frage, in welcher Weise an Hamburger Schulen informationsbasiertes und interaktionsbasiertes Vertrauen zwischen Eltern und Schule aufgebaut wird, wie die Schulen also die Eltern mit Informationen zum Lernen ihres Kindes versorgen und wie sie den Interaktionsprozess mit Eltern gestalten.

2. Die Sichtweise von Hamburger Eltern im Rahmen der Schulinspektion

2.1 Berücksichtigung von Eltern im Inspektionsverfahren

Die Hamburger Schulinspektion erhebt quantitative und qualitative Daten zur Perspektive von Eltern. Alle Hamburger Eltern haben die Möglichkeit, sich zum Zeitpunkt der Inspektion an der Schule ihres Kindes an einer Fragebogenerhebung zu beteiligen. Während des ersten Inspektionszyklus (2007 bis 2012) beinhaltete das Verfahren Fragen zu allen Bereichen des Orientierungsrahmens Schulqualität (BSB 2006). Mit Beginn des zweiten Zyklus wurde eine Straffung und Fokussierung vorgenommen. Es werden seitdem Daten zum Informationsfluss, der elterlichen Zufriedenheit mit dem Bildungs- und Betreuungsangebot, der Einbindung von Eltern in Lernprozesse, der Identifikation mit der Schule sowie den Partizipationsmöglichkeiten erhoben. Für alle Schulbeteiligten sind die Fragen transparent durch Veröffentlichung auf der Homepage der Schulinspektion (vgl. die Beiträge von Diedrich in diesem Band).

Während des Inspektionsverfahrens findet an den Schulen zusätzlich ein Gruppeninterview mit nach Möglichkeit jeweils acht Elternteilen statt, die von der Schule bestimmt werden. Dabei handelt es sich meist um schulische Gremienvertreterinnen und -vertreter. Oft nehmen aber auch einzelne Eltern daran teil, die keine Funktion in der Schule ausüben. Das Gruppeninterview wird auf bestimmte Themenbereiche fokussiert, die sich aus der Analyse von Dokumenten und Fragebogenergebnissen der einzelnen Schule ergeben. Die Teilnahme an den Gruppeninterviews schwankte im gesamten Inspektionszeitraum sehr stark. Teilweise kamen nur ein oder zwei Elternteile, teilweise wurde die empfohlene Zahl von acht Teilnehmerinnen und Teilnehmern überschritten. Die erwünschte Anzahl von acht Teilnehmerinnen und Teilnehmern wurde somit nicht immer erreicht.

Die Schulen erhalten von der Schulinspektion eine Rückmeldung zu den einzelnen Items der Fragebogenerhebung, aufgeschlüsselt nach den Prozentwerten zu den einzelnen Variablen und Konstrukten, sodass sich die Möglichkeit der eigenen Dateninterpretation über den Inspektionsprozess hinaus ergibt. Für die Auswertung der Daten wurden die Variablen im zweiten Inspektionszyklus zu Konstruk-

ten zusammengefasst: Information durch die Schule, Sicherung des Schulklimas, Breite des schulischen Angebots, Einbindung der Eltern in schulische Lehr- und Lernprozesse sowie Elternbindung.

2.2 Rückmeldungen von Eltern

Im ersten Zyklus (2007–2012) haben sich 36.075 Eltern an den Onlineerhebungen der Hamburger Schulinspektion beteiligt. Auch wenn die Rücklaufquote der Elternfragebögen von 2007 bis 2011 kontinuierlich angestiegen ist, wie im Jahresbericht 2010/2011 der Schulinspektion beschrieben wird (vgl. BSB/IfBQ 2013, S. 13), entspricht dies gemittelt nur einer Rücklaufquote von insgesamt 23,52 Prozent. Die Daten sind bei einer solchen Rücklaufquote in der Regel systematisch verzerrt und können nicht als repräsentativ für die Meinung aller Eltern einer Schule betrachtet werden. Die Beteiligung schwankte allerdings zwischen 73,5 und 3,6 Prozent, sodass die Ergebnisse einzelner Schulen als repräsentativ betrachtet werden können und an 10 Prozent der Schulen in die Bewertung einbezogen werden konnten (vgl. ebd.). Die Beteiligung war an Grundschulen und Gymnasien deutlich höher als an Haupt- und Realschulen bzw. Stadtteilschulen. Im zweiten Inspektionszyklus der Hamburger Schulinspektion, der im Herbst 2012 pilotiert wurde, ist die Datenerhebung bezüglich der Elternschaft modifiziert worden, um die Rücklaufquote zu erhöhen. Die Eltern erhalten seitdem die Fragebögen nicht mehr als Link zu einem online auszufüllenden Formular, sondern in Papierform (vgl. ebd.). Der Fragebogen wurde deutlich verkürzt und gleichzeitig um einige Fragen zur Einbindung der Eltern in Lehr-Lernprozesse ergänzt. Darüber hinaus sind die Fragebögen in verschiedene Sprachen übersetzt worden, um Eltern mit Migrationshintergrund die Teilnahme zu erleichtern. Mit der Veränderung der Art und Weise, in welcher die Eltern über die Befragung der Schulinspektion informiert wurden, und der Möglichkeit, den Fragebogen in Papierform auszufüllen, konnte die Beteiligung der Eltern deutlich gesteigert werden. Die Rücklaufquote der Elternbefragungen ist mit Beginn des zweiten Zyklus gestiegen und liegt seitdem bei rund 55 Prozent. Im Zeitraum von Herbst 2012 bis Frühjahr 2014 nahmen an den bis dahin inspizierten 49 Schulen insgesamt 16.120 Eltern an der Befragung teil.

Im Rahmen der Fragebogenerhebung sowohl des ersten als auch des zweiten Zyklus konnten einzelne Aussagen auf einer vierstufigen Skala von 1 bis 4 (trifft nicht zu/trifft eher nicht zu/trifft eher zu/trifft voll zu) zu den Bereichen *Information*, *Zufriedenheit mit der schulischen Arbeit*, *Elternbindung/Elternbeteiligung* bewertet werden. Ein Mittelwert über 2,5 bedeutet tendenziell Zustimmung, ein Mittelwert unter 2,5 bedeutet tendenziell eine Ablehnung der Aussage. Ein hoher Mittelwert signalisiert dementsprechend eine hohe Zustimmung zur Aussage, ein niedriger Mittelwert eine Ablehnung.

3. Darstellung von ausgewählten Ergebnissen aus den Fragebogenerhebungen

Wie äußern sich die Eltern zur Schule ihres Kindes bzw. ihrer Kinder? Um Aussagen über die Gesamtzufriedenheit der Eltern zu machen, wurde für die Befragten des ersten Zyklus der Mittelwert über alle Variablen hinweg berechnet. Die 36.075 Eltern, die im ersten Zyklus der Schulinspektion an der Online-Befragung teilgenommen haben, sind insgesamt überwiegend zufrieden mit der Schule ihres Kindes (MW/Gesamtstichprobe 3,0). An den im zweiten Zyklus bis Frühjahr 2014 inspierten 49 Schulen beurteilen die Eltern bei einer Rücklaufquote von 54,85 Prozent ihre Zufriedenheit in ähnlicher Weise wie im ersten Zyklus (MW 2,92). Für diesen Beitrag sind einzelne Ergebnisse ausgewählt worden, die sich auf Items beziehen, die in beiden Zyklen den Eltern vorgelegt wurden. Es lassen sich ähnliche Zufriedenheitswerte feststellen. Der Blick auf die nachfolgenden Tabellen zeigt jedoch, dass die Zufriedenheit der Eltern bezogen auf einzelne Aspekte variiert.¹

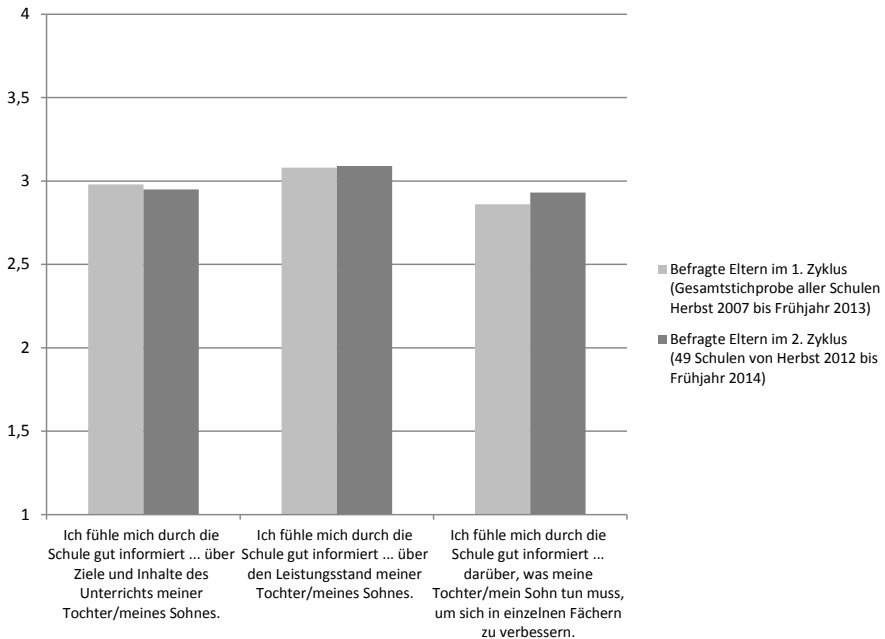
3.1 Wie zufrieden sind die Eltern mit den Informationen, die sie von der Schule erhalten?

Eine deutliche Mehrheit der Eltern hat die Erfahrung gemacht, dass sie die Möglichkeit haben, Kontakt zu den Lehrkräften aufzunehmen. Der Aussage, dass die Lehrkräfte für Rückfragen außerhalb von institutionalisierten Zusammenkünften wie Elternabenden und Elternsprechtagen ansprechbar sind, stimmen insgesamt 87,9 Prozent der Eltern zu. Eine Kontaktaufnahme durch die Lehrkräfte mit den Eltern bei Bedarf erleben 75,85 Prozent der Eltern. Dies bedeutet allerdings gleichzeitig, dass etwa ein Viertel der Eltern eine solche aktive Kontaktaufnahme der Lehrkräfte nicht oder eher nicht erlebt. Betrachtet man die Zufriedenheit der Eltern mit den Informationen, die sie von der Schule erhalten, so lassen sich folgende Aussagen treffen: Die Zufriedenheit der Eltern mit Informationen über Ziele und Inhalte des Unterrichts (MW 2,95), den Leistungsstand ihres Kindes (MW 3,09) und darüber, was der Sohn/die Tochter tun muss, um schulische Leistungen in einzelnen Fächern zu verbessern (MW 2,93), ist ähnlich hoch wie die Gesamtzufriedenheit mit der Schule (vgl. Abb. 1 unten). Gleichzeitig fühlen sich aber 26,5 Prozent nicht hinreichend über die Ziele und Inhalte des Unterrichts informiert, 22 Prozent nicht

1 Datenquelle: Daten der Schulinspektion Hamburg. Gesamtstichprobe aller Schulen im Zeitraum von Herbst 2007 bis Frühjahr 2013, (N=36.075) und Gesamtstichprobe aller inspizierten Schulen im Zeitraum von Herbst 2012 bis Frühjahr 2014, (N=16.120). Zusammenstellung durch Stephanie Graw und Lena Wegener.

hinreichend über den Leistungsstand ihres Kindes. 34 Prozent der Eltern beurteilen die Aussage „Ich fühle mich durch die Schule gut informiert ... darüber, was meine Tochter/mein Sohn tun muss, um sich in einzelnen Fächern zu verbessern“ als nicht bzw. eher nicht zutreffend.

Abbildung 1: Zufriedenheit der Eltern mit der Information durch die Schule im ersten und zweiten Zyklus der Schulinspektion



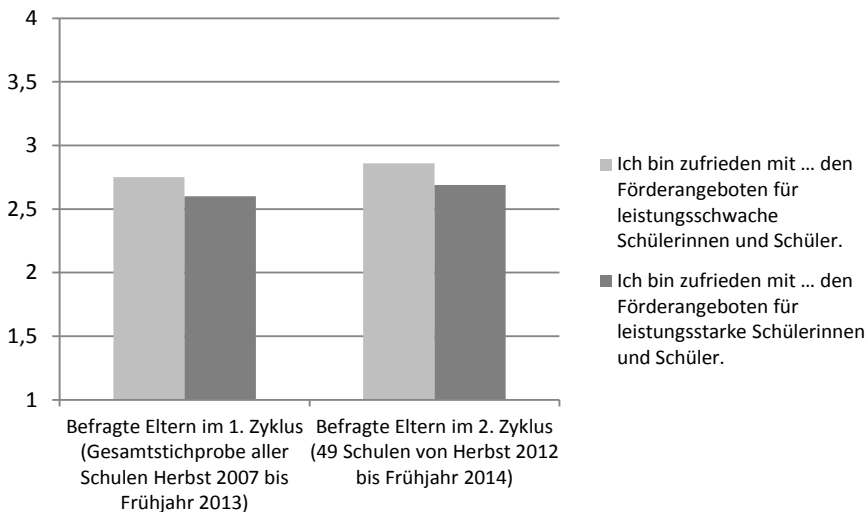
Datenquelle: Daten der Schulinspektion Hamburg. Gesamtstichprobe aller Schulen im Zeitraum von Herbst 2007 bis Frühjahr 2013 (N=36.075) und Gesamtstichprobe aller inspezierten Schulen im Zeitraum von Herbst 2012 bis Frühjahr 2014 (N=16.120).

Bezogen auf die in beiden Inspektionszyklen kontinuierlich abgefragten Items kann zusammenfassend festgehalten werden, dass die Zufriedenheit der Eltern mit den Informationen durch die Schule in beiden Inspektionszyklen auf einem ähnlichen Niveau angesiedelt ist. Gleichzeitig ist etwa ein Viertel der Eltern unzufrieden mit den Informationen, die sie von der Schule bezogen auf ihr Kind erhalten.

3.2 Wie zufrieden sind die Eltern mit der schulischen Arbeit?

Die Zufriedenheit der Eltern mit der schulischen Arbeit wird durch Fragen nach der Zufriedenheit mit den Förderangeboten für leistungsstärkere und leistungsschwächere Schülerinnen und Schüler abgefragt. Die Zufriedenheit mit den Förderangeboten sowohl für leistungsschwache Schülerinnen und Schüler (MW 2,75) als auch für leistungsstarke Schülerinnen und Schüler (MW 2,6) liegt im schwach positiven Bereich und damit unter der Gesamtzufriedenheit. In Bezug auf die Zufriedenheit mit der schulischen Arbeit bestätigt sich auch im zweiten Zyklus der Trend des ersten Zyklus, dass die Eltern mit der Förderung der leistungsschwachen Schülerinnen und Schüler zufriedener sind als mit der Förderung der leistungsstarken Schülerinnen und Schüler (vgl. Abb. 2).

Abbildung 2: Zufriedenheit der Eltern mit der schulischen Förderarbeit



Quelle: Daten der Schulinspektion Hamburg. Gesamtstichprobe aller inspizierten Schulen im Zeitraum von Herbst 2007 bis Frühjahr 2013 (N=36.075) und Ergebnisse der Elternbefragungen im zweiten Zyklus der Schulinspektion (49 Schulen von Herbst 2012 bis Frühjahr 2014, N=16.120).

Auch wenn die Elternschaft insgesamt eher zufrieden mit der Schule ist, so ist die Zufriedenheit mit der Förderung etwas geringer als die Zufriedenheit insgesamt. Gründe hierfür lassen sich im Jahresbericht der Schulinspektion finden: Förderung findet meist wenig institutionalisiert statt, hängt vom Engagement einzelner Lehrkräfte ab und zusätzliche Fördermaßnahmen werden häufig nicht auf den individuellen Förderbedarf abgestimmt (vgl. BSB/IfBQ 2013, S. 24).

3.3 Wie zufrieden sind die Eltern mit ihren Möglichkeiten, der Schule eine Rückmeldung zur Arbeit zu geben?

An Schulen gibt es unterschiedliche Formen der Rückmeldung, mit denen die Eltern der Schule Auskunft zu ihrer Zufriedenheit geben können. 83,1 Prozent der Eltern befragten Eltern gaben bei der Erhebung der Schulinspektion im ersten Zyklus an, im Rahmen von Elternabenden regelmäßig Gelegenheit zu haben, den Schulen Rückmeldung zu deren Arbeit zu geben. Nur 11,8 Prozent der befragten Eltern gaben an, schon einmal schriftlich von der Schule zu ihrer Zufriedenheit befragt worden zu sein. Wie bereits im ersten Zyklus werden auch zum Zeitpunkt des zweiten Zyklus Eltern sehr selten von ihrer Schule schriftlich über ihre Zufriedenheit mit deren Arbeit befragt. Lediglich 14,7 Prozent der Eltern beantworten die Aussage „Ich bin von der Schule schon einmal schriftlich über meine Zufriedenheit mit der Arbeit der Schule befragt worden“ mit trifft zu oder trifft eher zu. Die Befragung im Rahmen der Schulinspektion stellt demnach für viele Eltern die einzige Möglichkeit dar, der Schule auf indirektem Weg eine Rückmeldung zu ihrer Zufriedenheit zukommen zu lassen.

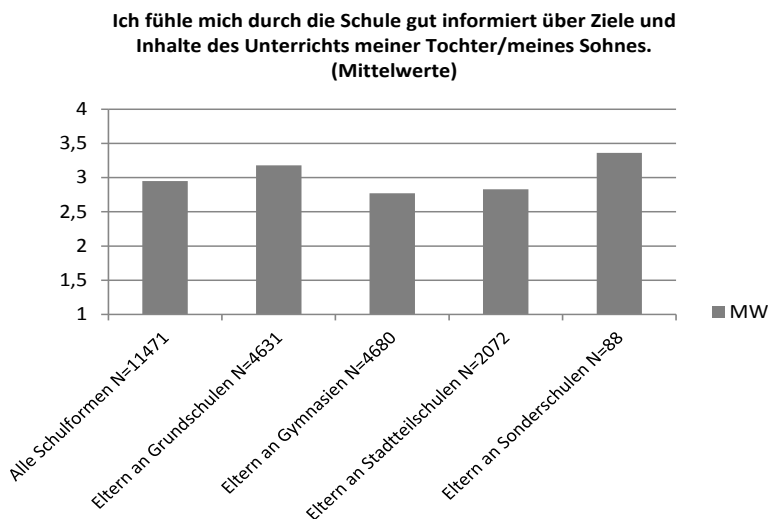
3.4 Wie werden Eltern in schulische Lernprozesse eingebunden?

Seit Beginn des zweiten Zyklus werden im Rahmen der Fragebogenerhebung einige zusätzliche Fragen gestellt, die sich auf die Einbeziehung von Eltern in schulische Lernprozesse beziehen. Die Daten der Schulinspektion Hamburg (Gesamtstichprobe der Schulen im zweiten Zyklus im Zeitraum von Herbst 2012 bis Frühjahr 2014, N=49 Schulen) zeigen eine deutlich geringe Einbeziehung der Eltern aller Schulformen in schulische Lernprozesse bzw. kaum aktive Aufforderung, diese Lernprozesse aktiv zu unterstützen. 72,6 Prozent der befragten Eltern geben an, bisher selten oder nie aufgefordert worden zu sein, Ausflüge und/oder Exkursionen zu begleiten. 83,6 Prozent geben an, dass sie bisher selten oder nie aufgefordert worden sind, individuelle Fördermaßnahmen (z. B. durch die Mitarbeit an Konzepten zur Leseförderung) aktiv zu unterstützen. Die Aussage „An der Schule meines Kindes wurde ich bisher dazu aufgefordert selber pädagogische Angebote (z. B. Kurse oder AGs) zu gestalten“ beantworten 88,1 Prozent mit selten oder nie. 92,6 Prozent äußern sich dahingehend, dass sie selten oder nie aufgefordert worden sind, den Schülerinnen und Schülern ihren Beruf vorzustellen.

3.5 Unterschiede elterlicher Zufriedenheit zwischen Schulen bzw. Schulformen

Sowohl im ersten als auch im zweiten Inspektionszyklus zeigen sich Unterschiede in der elterlichen Zufriedenheit bezogen auf die einzelnen Schulen. Ein Blick auf die 25 Schulen, mit denen die Eltern jeweils am zufriedensten bzw. am unzufriedensten sind, zeigt die großen Unterschiede elterlicher Zufriedenheit. Die 25 Schulen mit den zufriedensten Eltern weisen Mittelwerte zwischen 3,58 und 3,40 auf. Die Eltern sind hier also sehr zufrieden. Von diesen sind 6 Sonder-/Förderschulen mit unterschiedlichen Förderschwerpunkten und 19 Grundschulen, davon 6 Schulen mit integrativen Regelklassen und/oder Integrationsklassen. Die 25 Schulen, an denen die Eltern weniger zufrieden sind, weisen einen Mittelwert zwischen 2,74 und 2,61 auf. Die Eltern sind demnach aber auch an diesen Schulen tendenziell eher zufrieden als unzufrieden. Unter diesen Schulen sind 10 Gymnasien, 5 Grundschulen, 3 Stadtteilschulen, 2 Gesamtschulen und 5 Haupt- und Realschulen (die beiden letztgenannten Schulformen existierten in Hamburg nur bis 2010). Die Differenzierung der Zufriedenheit nach Schulformen deutet an, dass es schulformspezifische Unterschiede gibt. Bezogen auf eine Aussage der Fragebogenerhebung im zweiten Zyklus soll dies konkretisiert werden (vgl. Abb. 3).

Abbildung 3: Unterschiede der elterlichen Zufriedenheit nach Schulformen



Quelle: Daten der Hamburger Schulinspektion. Ergebnisse der Elternbefragungen im zweiten Zyklus der Schulinspektion (49 Schulen von Herbst 2012 bis Frühjahr 2014). Anzahl der Eltern, die sich zu diesem Item geäußert haben: N=11.471.

Zur Aussage, ob sie sich durch die Schule gut informiert über Ziele und Inhalte des Unterrichts ihrer Tochter/ihrer Sohnes fühlen, äußern sich die Eltern an Sonderschulen und die Eltern an Grundschulen positiver als der Durchschnitt der Eltern (MW 2,95). Eltern an Stadtteilschulen und an Gymnasien äußern sich etwas weniger positiv als der Durchschnitt aller Eltern. Zur Zufriedenheit mit den Förderangeboten für leistungsstarke Schülerinnen und Schüler lässt sich eine ähnliche Verteilung feststellen.

Zusammenfassend kann zur Elternbefragung der Schulinspektion festgehalten werden, dass sich die Daten des ersten Zyklus durch die Erhebung des zweiten Zyklus bestätigen. Bei einer deutlich höheren Rücklaufquote von 54,85 Prozent zeigen die Antworten der Eltern im Zeitraum des zweiten Zyklus ähnliche Zufriedenheitswerte bezogen auf die Informationen zum Unterricht, die Förderangebote als Teil der schulischen Arbeit und die Elternbeteiligung durch schriftliche Befragung. Betrachtet man die Antworten der Eltern insgesamt, so weisen diese auf bestimmte Themenfelder hin, die die Elternschaft durchgängig bewegen: die Förderung leistungsstarker und leistungsschwächerer Kinder sowie die Kommunikation zwischen Schule und Eltern. Gleichzeitig wird aus den Ergebnissen sichtbar, dass nur wenige Hamburger Schulen bislang die Zufriedenheit von Eltern in Eigeninitiative erfragt haben, um Schlüsse für die Schulentwicklung zu ziehen.

4. Gründe für elterliche Zufriedenheit/Unzufriedenheit

Hamburger Eltern schätzen viele Aspekte der auf die Elternzufriedenheit bezogenen Items eher positiv ein, beispielsweise die Kommunikation zwischen Schule und Familien, die Förderung und die Möglichkeiten, als Eltern zu partizipieren. Gleichwohl zeigen sich Unterschiede in der Zufriedenheit. Einige Schulen unterscheiden sich in Bezug auf ihre Zufriedenheit signifikant vom Durchschnitt. Diese Unterschiede sind nicht einer bestimmten Schulform zuzuordnen. Die Aussagen aus den im Rahmen der Schulinspektion geführten Elterninterviews an diesen Schulen zeigen Unterschiede der Zufriedenheit bzw. Unzufriedenheit vor allem in Bezug auf die Kommunikation zwischen Schule und Familie auf, insbesondere betreffend den Informationsfluss über die Anforderungen, die an die Schülerinnen und Schüler gestellt werden.

An dieser Stelle wird in den Aussagen von Eltern in den Interviews konkretisiert, woran sich die elterliche Zufriedenheit bzw. Unzufriedenheit bezogen auf die in beiden Zyklen im Rahmen der Fragebogenerhebung berücksichtigten Bereiche Information, Zufriedenheit mit der schulischen Arbeit, Elternbindung/Elternbeteiligung festmachen lässt. Hierfür wurden einige der Gruppeninterviews mit Eltern auf die Aspekte Förderung, Information, Kommunikation und Beteiligung hin betrachtet. Aus den 49 Schulen, die im zweiten Zyklus im Zeitraum von Herbst 2012

bis Frühjahr 2014 inspiziert worden waren, ist je eine Schule ausgewählt worden, an der die elterliche Zufriedenheit signifikant vom Mittelwert 2,92 nach oben bzw. unten abweicht, also je eine Schule, an der die Eltern deutlich zufriedener oder deutlich unzufriedener sind als an anderen Schulen. Zusätzlich wurde eine Schule ausgewählt, an der sich die Zufriedenheit im Bereich des Mittelwertes bewegt. Alle drei Schulen haben sich mit Aspekten der Förderung intensiv auseinandergesetzt und Strukturen implementiert, in denen individuelles Lernen ermöglicht und befördert wird.

Die entsprechenden Elterninterviews wurden ausgewertet, indem zu den von der Schulinspektion erfragten Bereichen mit einem qualitativen Kodierverfahren Aussagen gesucht, miteinander verglichen und auf Gemeinsamkeiten und Unterschiede hin betrachtet wurden. Interviewpassagen, die sich thematisch auf Themenbereiche des Elternfragebogens beziehen, wurden transkribiert. Im Transkript wurden grammatikalische und syntaktische Fehler belassen, die im Wesentlichen daraus resultieren, dass Deutsch für einige Eltern eine Fremdsprache darstellt. Die Darstellung der Auswertung konzentriert sich auf Ausschnitte, anhand derer Unterschiede zwischen Schulen sichtbar werden. Mit Beispielen von Aussagen einzelner Eltern sollen die Themen veranschaulicht werden, die für die elterliche Zufriedenheit dieser Schulen zentral sind. Sie können Anhaltspunkte dafür liefern, warum die Eltern an einigen Schulen zufriedener sind als im Durchschnitt der Schulen, bzw. dafür, in welchen Bereichen Eltern auch bei einer allgemeinen Zufriedenheit mit der Schule Entwicklungsbedarf sehen.

4.1 Exemplarische Aussagen von Eltern an einer Schule mit durchschnittlicher Zufriedenheit

Bei Schule A handelt es sich um eine Schule mit niedrigem Sozialindex, die jahrgangsübergreifendes Lernen eingeführt hat und deren Elternschaft durchschnittlich zufrieden ist (MW 2,91). Viele Eltern an dieser Schule haben eine andere Muttersprache als Deutsch. Mehrere Eltern im Interview beziehen sich auf negative Schulerfahrungen an anderen Schulen, die sie entweder in ihrer Kindheit oder als Eltern bereits gemacht haben, und ziehen Vergleiche zwischen den ihnen bekannten Schulen. Alle Interviewpartner sind sich einig über die Zufriedenheit mit dieser Schule.

M1: Wir wissen alle, dass die Kinder sind nie gleich. Selbst ich habe sechs Kinder, und die sind nie gleich. [...] Das ist hier viel besser, wo die gucken zum Beispiel, diese Kinder, die sind ganz langsam, die arbeiten in anderes Niveau, oder könnten die Kinder, die viel besser sind, die arbeiten zusammen, und was sie machen sehr hier, die Lehrer, die Lehrer sind einfach frei, du kannst jederzeit da gehen und sagst du, hey, ich bin nicht zufrieden, da gibt's überhaupt nicht Probleme. Dann schnell kommt mit einem

Kind ganz vorne und arbeitet mit dein Kind. Und ... die Lehrer sind hier für mich wie Familie. [...] Ich hab auch andere Schule, wo die Kinder sind. [...] Meine älteste, Gymnasium: Ich hatte mich nie wohl gefühlt – die Lehrer hatten einfach nur Distanz, ich bin der Lehrer – du bist die Mutter. Hier ist ganz anders. Die Lehrer sind mit Eltern einfach frei. Du kannst sprechen einfach frei und ich fühle mich einfach wohl.

V1: Und einfach mal so ein kleines Gespräch auf dem Schulhof, wenn man die Kinder abholt. Ich komm mal kurz. Du kannst auch mal einfach kurz was besprechen, zack, zack, zack. Sowas hätte es früher nie gegeben. Oder, was ich vorhin gesagt hab, die Lernentwicklungsgespräche kannte ich auch gar nicht. [...] Das war wirklich toll, weil man wirklich wusste um dieses Hand-in-Hand-Arbeiten. Das ist wirklich Wahnsinn, dass man auch als Eltern mitgeteilt bekommt, in regelmäßigen Abständen, Mensch, pass mal auf, [Name des Vaters], so und so, können wir unterstützen, ja.

M2: Und die Kinder unterschreiben.

V1: Und wir auch. Und das find ich toll.

(Schule A Elterninterview 2013, Minuten 9:33 bis 11:39)

Die Eltern beziehen sich auf die differenzierte und individualisierte Förderung, die an dieser Schule über ein auf Lernlandkarten basierendes Curriculum ermöglicht wird. Daneben ist die Kommunikation zwischen Schule und Eltern bedeutsam. Die Eltern erläutern, dass sie die jederzeitige Ansprechbarkeit der Lehrkräfte schätzen und das schnelle Bemühen um Lösungen („zack, zack, zack“). Hierzu trägt eine angstfreie Atmosphäre in den Gesprächen zwischen Eltern und schulischen Akteuren bei, die es ihnen ermöglicht, auch Kritik an der Schule zu üben. Die Lernentwicklungsgespräche (LEG) sind bedeutsam für die Eltern, da sie regelmäßig Informationen erhalten und außerdem auch Maßnahmen zur Unterstützung des Kindes abgesprochen werden. Die Verbindlichkeit der getroffenen Absprachen durch die Unterschrift von Kind und Eltern wird positiv gesehen. Die LEG werden als unterstützendes Element der Zusammenarbeit betrachtet und führen zum Eindruck von intensiver Kooperation von Schule und Familie in Form des „Hand-in-Hand-Arbeitens“.

4.2 Exemplarische Aussagen an einer Schule mit unterdurchschnittlicher Zufriedenheit

An Schule B ist die Elternzufriedenheit geringer als an der Schule aus dem vorherigen Beispiel und liegt auch signifikant niedriger (2,80) als der Mittelwert aller

Schulen (2,92). Auch an diesem Gespräch nehmen sowohl gewählte Elternvertreter als auch Eltern ohne Funktion teil. Die Erreichbarkeit der Fachlehrkräfte wird von den Interviewpartnern gleich zu Beginn als Problem beschrieben. Auf die Frage nach dem Umgang zwischen allen Schulbeteiligten an der Schule antworten die Eltern folgendermaßen:

M1: [...] Ich sag es jetzt mal ehrlich, wie ich es empfinde: Ich finde, es ist alles ein bisschen sehr schleppend. Ähm. Also. Wir im Elternrat haben natürlich immer einen direkten Draht zur Schulleitung. Ich persönlich bin auch eine Mutter, die direkten Kontakt zur Elternschaft wählt. [Unterbrechung durch zu spät kommende Mutter] Im Elternrat, wie gesagt, denk ich, ist eine Kommunikation schon vorhanden, weil wir einen direkten Kontakt zur Schulleitung haben. Da sind wir auch alle ganz ehrlich, schmeißen unsere Probleme auf den Tisch und ähm eigentlich ... wird seitens der Schulleitung auch sich immer bemüht, in irgendeiner Weise es zu glätten oder es zu lösen. Was nicht immer einfach ist und auch nicht immer gleich funktioniert. Aber ich persönlich finde es sehr schwer, Kontakt auch zu den Lehrern zu haben. Wir haben ja hier eine Plattform an der Schule, mit der wir kommunizieren, das ist IServ. Aber, also, ich hab es nicht erlebt, aber ich weiß es von anderen Eltern, die haben über IServ eine Mail an Lehrer geschrieben, weil irgendwie was vorgefallen war oder etcetera, die haben wochenlang keine Rückmeldung bekommen.

M2: Ich glaube, was fehlt, ist – ein Elternsprechtage. Das, was es inzwischen gibt, sind Lernentwicklungsgespräche. Wir haben alle Kontakt zu den Klassenlehrern. Die Eltern haben guten Kontakt zu den ...

M3: Eins ist inzwischen gestrichen worden, denk dran.

M2: Wir haben guten Kontakt zu den Klassenlehrern. Das läuft dann ja dann gut. Da gibt es viel Informationen. Aber wir haben mit den Fachlehrern überhaupt keinen Kontakt. Ich kenne die Lehrer zum Teil gar nicht. Ich weiß gar nicht, wer das ist. Ich kenn kein Gesicht, nichts. Und, jetzt ist ein Lernentwicklungsgespräch uns auch noch gestrichen worden. Jetzt gibt es nur noch eins. Also, da glaub ich, ist ein großes, großes Manko. Ein großes Loch tut sich auf. Wir können gar nicht mehr angreifen. Auch zwischen Eltern und Lehrern. Ich glaube zwischen Schulleitung und Eltern, wie du das auch sagst, da ist eine ganz tolle Kommunikation, ganz offen, sehr freundlich, sehr lustig auch. Wir haben auch Spaß in den Elternratssitzungen.

(Schule B, Elterninterview 2013, Minuten 8:11 bis 11:00)

Auf eine Nachfrage der Interviewerin konkretisieren die Interviewteilnehmer ihre Aussagen auf die Information der Elternschaft insgesamt:

M4: Wir haben da ein sehr schönes Buch, das nennt sich „Mein gelber Freund“, das wird auch in der 5. [...] Ich finde das sehr positiv. Das gefällt mir sehr gut, das könnte man überall einführen. Das geht wöchentlich hin und her, und wenn man die Lehrer hat auch, die das gut führen, das ist natürlich auch eine Disziplin der Lehrerschaft, haben wir im Prinzip einmal die Woche sofort ein Feedback, sind die Hausaufgaben gemacht, war es ordentliches Betragen in der Schule, Fehlzeiten, so, ich kann meine Sachen reinschreiben, krank, was weiß ich, wenn mir irgendwas auf dem Herzen liegt, Kontakte schließen und so weiter. Und das wird abgezeichnet und gegengezeichnet. Für mich ist das der kürzeste Weg und damit fahre ich in unserer Klasse super.

M2: Aber es läuft nicht in allen Klassen. Es läuft nicht in allen Klassen und man muss auch dazusagen, es wird längst nicht von allen ...

M4: Eltern genutzt. [...]

V6: Es haben alle Schüler. Alle Schüler dieser Schule.

[Alle reden durcheinander.]

M1: Es gibt aber Lehrer, die das nicht nutzen. Die nicht darauf achten, jetzt das gelbe Heft rausholen und dann reinschreiben. Das wird einfach ...

M4: Da sind die Lehrer einfach unterschiedlich. [...] Es ist an sich eine gute Sache. Wäre besser, wenn jeder darin arbeiten würde.

(Schule B, Elterninterview 2013, Minuten 12:18 bis 13:34)

Aus beiden Gesprächsabschnitten dieses Interviews wird die Unzufriedenheit der Eltern mit der Kommunikation zwischen Schule und Familie deutlich, die im Interview einen breiten Raum einnimmt. Der als schleppend beschriebene Umgang untereinander resultiert nicht aus fehlenden Strukturen der Kommunikation. Die eingeführten Lernentwicklungsgespräche, die Kommunikationsplattform und auch das „gelbe Heft“ weisen auf die Implementation von Kommunikationsstrukturen an dieser Schule hin, die von den Eltern für sinnvoll gehalten werden. Diese Maßnahmen zur Informationsweitergabe greifen aber aus Sicht der Eltern nicht, da „die Lehrer einfach unterschiedlich“ seien. Die Interviewpartnerinnen und -partner unterscheiden zwischen ihren Möglichkeiten, als Gremienvertreterinnen und -vertreter Informationen von der Schulleitung zu erhalten, und den übrigen Eltern der Schule, die diese Möglichkeit nicht haben. Sie weisen darauf hin, dass die eingeführten Kommunikationsstrukturen nur teilweise funktionieren, weil sie nicht von allen Lehrkräften der Schule gleichermaßen genutzt werden. Eine Mutter verwendet die

Formulierung „Wir können gar nicht mehr angreifen“, um zu beschreiben, welche Nachteile sich aus einem gestrichenen Lernentwicklungsgespräch ergeben, und um auszudrücken, dass sie kaum Gelegenheit hat Kritik zu äußern. Schule und Familie werden von ihr als Gegner betrachtet. Aus dem Wegfall eines zweiten Lernentwicklungsgesprächs pro Jahr schließen die Eltern auf eine Reduktion ihrer Möglichkeiten, Informationen über die Lernentwicklung ihres Kindes zu erhalten, und damit auch auf eine Verringerung der Möglichkeit das eigene Kind im Lernen zu unterstützen. Der Wunsch nach Kontakt zu den Fachlehrkräften wird deutlich vorgetragen. Die eingeführte Struktur einer Kommunikationsplattform ist nicht hinreichend, da konkrete Anfragen von Eltern nicht beantwortet werden. Es entsteht kein Dialog zwischen Lehrkräften und Eltern. Der aus ihrer Sicht „schleppende“ Umgang zwischen Schule und Familie wird dadurch befördert, dass nicht hinreichend für Verbindlichkeit in der Umsetzung der eingeführten Maßnahmen zur Sicherstellung des Informationsflusses gesorgt wird. Hieran wird indirekt Kritik am Handeln der Schulleitung sichtbar, in deren Verantwortung dies läge. Die Schulleitung wird zwar persönlich im Umgang geschätzt, gleichwohl drücken die Eltern, die durch Gremienarbeit Kontakt mit ihr haben, aus, dass auch mit der Schulleitung die Kommunikation nicht immer einfach ist. Die im obigen Gesprächsabschnitt ausgedrückte Lehrerabhängigkeit der Durchführung festgelegter Maßnahmen taucht im Interview immer wieder zu unterschiedlichen Fragen auf, wie zur Förderung oder zur Berufsorientierung. Kennzeichen dieser Schule ist eine den Eltern fehlende Sicherheit bezüglich eines verlässlichen und vergleichbaren Lehrerverhaltens. Eltern mit Geschwisterkindern berichten von höchst unterschiedlich agierenden Lehrkräften.

4.3 Exemplarische Aussagen von Eltern an einer Schule mit überdurchschnittlicher Zufriedenheit

Schule C erreicht in der Elternbefragung einen besonders hohen Mittelwert (2,97) in der Gesamtzufriedenheit mit der Schule. An Schule C werden die Aspekte Förderung, Information, Kommunikation und Beteiligung aktiv von den Eltern angesprochen. Dabei liegt der Schwerpunkt auf der Beschreibung der Aktivitäten von Eltern.

V1: [...] Die GBS, der Zubau, dessen hoffentlich baldige Fertigstellung. Wichtig ist uns auch noch die Präsentation sozusagen der Informationen, die wir an die Eltern dieser Schule weitergeben können an dieser Schule und dazu ist jetzt auf der einen Seite schon die Homepage neu gebaut worden der Schule und da haben wir die Hoffnung, dass wir noch ein paar Akzente setzen können, insofern, als dass wir die Informationsverbreitung da optimieren können, also einfach mehr Leute auch über das Internet ansprechen können, mehr Eltern ansprechen können, um die Information breiter zu streuen.

M 2: Dann hatten wir noch das Thema „Zweites Lernentwicklungsgespräch“. Da haben wir ja auch eine tolle Projektgruppe gegründet, in Zusammenarbeit mit der Schulleitung, dass wir noch ein zweites Gespräch oder Zusammenkunft hinbekommen können.

V1: Da sind wir auf Lösungssuche gewesen. Da hatten wir der Lehrerkonferenz einen Kompromiss vorgeschlagen, der vorsehen sollte, dass die Arbeitslast, die dann ja zusätzlich entsteht, mit in den Unterricht integriert wird, und dieser Vorschlag kam leider nicht positiv durch.

[...]

V 1: Es sind natürlich regelmäßige Punkte, die immer wieder auftauchen, sozusagen von Schuljahr zu Schuljahr, das ist klar, Fahrrad-TÜV, das muss ich auch erwähnen, ist klar, wo wir eben vom Elternrat oder vom Schulverein so eine Veranstaltung haben und die Fahrräder der Kinder auf technische Sicherheit überprüft werden.

V4: Wo auch die Verkehrswacht mit dabei ist. Organisiert auch die Verkehrswacht mit der Schule.

M3: Auch so dieses Fördern und Fordern der Kinder ist auch so Thema, Inklusion, das kommt auch immer wieder so ins Gespräch, man will den Kindern ja auch gerecht werden, es wird immer wieder aufgegriffen, dieses Thema. Auch was der Herr M. jetzt macht hier.

V4: Förderkonzept.

M3: Hat er auch super dargestellt beim letzten Elternvertretertreffen.

(Schule C, Elterninterview 2013, Minuten 0:31 bis 3:25)

An diesem Gesprächsausschnitt aus Schule C wird sichtbar, dass die Eltern sich als Kooperationspartner der Schule verstehen. Sie nennen als wichtige Themen des Elternrates die Umsetzung des Konzepts Ganztägige Bildung (GBS) und den Zubau eines weiteren Gebäudes. Sie sprechen von „wir“ und sehen sich in der Mitverantwortung für das Kommunikationsmanagement der Schule bezüglich dieser und anderer Themen. Sie sehen sich in der Verantwortung für die Schul-Homepage. Sie arbeiten in Projektgruppen gemeinsam mit den schulischen Akteuren an Themen. Sie organisieren selbstverantwortlich schulische Veranstaltungen, die aus ihrer Sicht bedeutsam für das Leben der Schülerinnen und Schüler sind, und tragen so zum Schulleben bei. Um das bislang zweite Lernentwicklungsgespräch pro Jahr zu erhalten, bemühen sich die Eltern darum, einen Kompromiss zu finden, damit das entfallende Lernentwicklungsgespräch durch andere Formen einer regelmäßigen Rückmeldung ersetzt werden kann. Aus der Beschreibung des Prozesses der Kon-

fliktbearbeitung wird deutlich, dass das Abstimmungsergebnis in der Schulkonferenz gegen ein zweites Lernentwicklungsgespräch nicht als Niederlage empfunden wird, sondern als Ergebnis eines gemeinsamen Diskussions- und Entscheidungsprozesses von schulischen Akteuren und elterlichen Gremienvertretern.

Die Eltern loben die Darstellung des Förderkonzepts vor der Elternschaft durch den dafür zuständigen Koordinator. In ihrer Aussage kommt es ihnen vor allem auf die Qualität seines Vortrags an – die Tatsache, dass sie Informationen erhalten, erscheint selbstverständlich. Aus den Aussagen der Eltern wird deutlich, dass an der Schule ein dialogisches Miteinander praktiziert wird, das von einem Vertrauensverhältnis geprägt ist. Die Eltern gewinnen Sicherheit durch schulische Strukturen und signalisieren eine persönliche Wertschätzung der schulischen Akteure. Die Eltern dieser Schule betrachten das schulische Angebot darüber hinaus nicht als bloße Dienstleistung, sondern sehen sich als Partner in einer dialogischen Beziehung zur Schule, in der sie ihren eigenen Beitrag zu erfolgreichen Bildungsprozessen leisten.

4.4 Zusammenfassung der Ergebnisse aus der Analyse der Interviews

Die Unterschiede zwischen den drei Schulen werden besonders in den unterschiedlichen Kommunikationsstilen deutlich. An Schule A assoziieren die Eltern Schule mit dem Begriff Familie und betonen ein Gefühl von „Freiheit“ in der Kommunikation mit der Schule, also einer angstfreien Kommunikation mit den schulischen Akteuren. An Schule B wird das Gefühl eines Gegeneinanders von Schule und Eltern aus den Äußerungen erkennbar, das sich für die Eltern vor allem in der Kommunikation zwischen Schule und Eltern manifestiert. An Schule C gewinnt die Elternzufriedenheit über die Atmosphäre und den Dialog hinaus eine Qualität im Sinne einer Bildungspartnerschaft, die durch Möglichkeiten elterlicher Partizipation entstehen, die von den Eltern genutzt werden. Die im Modell von Hoover-Dempsey & Sandler (1995, 1997, 2005) beschriebene notwendige Passung zwischen elterlichem Engagement und schulischen Erwartungen und ihre Auswirkungen auf das elterliche Gefühl der Selbstwirksamkeit bezüglich der aktiven Unterstützung ihrer Kinder wird an Schule C deutlich.

Aus den Interviewbeispielen aller drei Schulen wird die Bedeutung eines informationsbasierten und interaktionsbasierten Vertrauens erkennbar, das sich nicht nur in der Transparenz von Lernzielen und Lernständen des Kindes ausdrückt, sondern sich auch in der Art und Weise der Kommunikationsgestaltung zwischen Schule und Familie manifestiert. Das Beispiel von Schule A bestätigt darüber hinaus die Erkenntnis von de Boer (2012, S. 202), dass es auch an Schulen mit eher bildungsfernen Elternhäusern gelingt, Eltern ins Boot zu holen, wenn der Fokus der Schule auf den Aufbau einer Anerkennungskultur gerichtet ist. Das Beispiel der Schule C demonstriert, dass dies in besonderer Weise erfolgversprechend ist, wenn

die Zusammenarbeit zwischen Schule und Familie als Dialog verstanden wird, nicht im Sinne einer anzuwendenden Technik, sondern als Haltung auf der Basis der „Bereitschaft gegenseitiger Wertschätzung“ (ebd., S. 203).

5 Fazit und Ausblick

Abschließend soll ein Abgleich der eingangs referierten Theorien zur Elternbeteiligung mit den Befunden aus der Schulinspektion erfolgen. Aus den Ergebnissen der Fragebogenerhebungen und der Interpretation der Interviews lassen sich Erkenntnisse zur Sicht der Eltern auf Hamburger Schulen gewinnen, die durch die Rückmeldung an die Schulen diesen als lernende Organisationen Entwicklungsmöglichkeiten eröffnen, darüber hinaus aber auch Hinweise für die Steuerung im Bildungswesen geben können.

Die Ergebnisse der Elternbefragungen im Rahmen der Hamburger Schulinspektion zeigen, dass Schulen bislang noch kaum Rückmeldungen von Eltern im Sinne einer systematischen internen Evaluation einholen, um Veränderungsprozesse in Gang zu setzen oder Schulentwicklungsprozesse fortlaufend zu analysieren. Ausgehend von den Überlegungen Matlands zur Nutzung der *Voice*-Strategie als Steuerungsinstrument im Bildungssystem (vgl. Matland 1995, S. 511) müssten Überlegungen angestellt werden, wie Schulen darin unterstützt werden können, Elternmeinungen systematisch zu erheben und zu nutzen.

Eltern-Feedback wird an Hamburger Schulen noch wenig für die Unterstützung der Lernprozesse genutzt. Vor dem Hintergrund der Erkenntnisse aus der Schulentwicklungsforschung zur positiven Auswirkung der Einbeziehung von Eltern auf die Leistungen der Schülerinnen und Schüler (vgl. Epstein 2001, Fullan 2001, Hoover-Dempsey & Sandler 2005) werden die Eltern somit nicht hinreichend als Partner für erfolgreiches Lernen gesehen. Die von Hoover-Dempsey und Sandler (1995, S. 322) für erfolgreiches Lernen notwendige Passung zwischen elterlichen Aktivitäten und den schulischen Erwartungen an elterliches Engagement ist vor allem an Schulen mit geringer Elternzufriedenheit gering, wie aus den analysierten Interviews deutlich wird. Wenn Schulen erfolgreich arbeiten wollen, dann müssen Schulleitungen sich nach Leithwood et al. (2004, S. 48) damit auseinandersetzen und verstehen, in welcher Weise Schule und Familie kooperieren und wie diese Kooperation für die Lernprozesse der Schülerinnen und Schüler produktiv werden kann.

Die im Rahmen der Hamburger Schulinspektion gewonnenen Daten stützen Theorien zum Vertrauen (vgl. Tiefel & Zeller 2012; Peters 2014, 2015a, 2015b). Aus den elterlichen Aussagen ergeben sich Hinweise für die Bestätigung der Generalisierbarkeit des Modells eines informationsbasierten und eines interaktionsbasierten Vertrauens (vgl. Peters 2015b), das nicht allein durch das Implementieren

von Kommunikationsstrukturen entsteht, sondern gelingender Interaktionsprozesse zwischen schulischen Akteuren und Eltern bedarf. Die Schulinspektion kann Schulen eine Rückmeldung darüber geben, wie sich die Beziehung von Eltern als Vertrauensgebern und Schulen als Vertrauensnehmern gestaltet. Die Inspektorinnen und Inspektoren können darüber hinaus vermitteln, dass die einzelne Schule durch eine dialogische, Transparenz herstellende Kommunikation mit den Eltern wichtige Impulse und Fokussierungen für die Schulentwicklung erhält.

Die Hamburger Schulinspektion hat im zweiten Zyklus Veränderungen des Verfahrens vorgenommen, um die Perspektive von Eltern stärker zu berücksichtigen. Die Modifikation der Befragung durch Abkehr von der digitalen Form, einer Kürzung des Fragebogens und der Ergänzung einzelner Fragen, hat zu einer Verdoppelung der Beteiligung von Eltern an der Befragung geführt und ermöglicht meist repräsentative Aussagen zur elterlichen Sicht auf das Angebot der Einzelschule. Die Schulen erhalten auf diese Weise eine Rückmeldung über die Außensicht von Eltern auf das System, über „blinde Flecke“ (vgl. Luhmann 2005, S. 17), die vom System selbst ansonsten nicht erkannt würden. Den schulischen Akteuren werden damit Möglichkeiten eröffnet, diese Erkenntnisse zur Selbstreflexion und zur Schulentwicklungsarbeit zu nutzen. Gleichzeitig reagiert die Schulinspektion auf die Erwartungen von Eltern nach mehr Information und Transparenz, indem sie im zweiten Zyklus dem Bedürfnis der Hamburger Elternschaft nachkommt, Informationen über die Schule zu erhalten. Die Veröffentlichung der schulspezifischen Ergebnisse in Form eines zusammenfassenden Berichts von Schulinspektionen im Internet ermöglicht Eltern bereits vor der Einschulung, sich einen Überblick über das schulische Angebot anhand von Qualitätskriterien zu machen. Sie unterstützt auf diese Weise den Umsteuerungsprozess im Bildungswesen und die Möglichkeit von Eltern, aus dem Bildungsangebot auszuwählen zu können und Entscheidungen zu treffen.

Für zukünftige Rückmeldungen an Schulen und hinsichtlich einer möglichen Modifikation des Verfahrens ergeben sich folgende Fragen:

- Wie können Schulen noch stärker motiviert werden, die gewonnenen Daten der Schulinspektion zur Zufriedenheit von Eltern für ihre Schulentwicklungsprozesse zu nutzen bzw. selbst Daten zu erheben, mit denen sie ein Feedback von Eltern erhalten?
- Wie können die Erkenntnisse zur bislang geringen Einbeziehung von Eltern in die Lernprozesse zur Steuerung im Bildungswesen genutzt werden?

Die Kommunikation zwischen Schule und Familie als Thema wahrzunehmen, das für gute Schule relevant und nicht nur klimatisch bedeutsam ist, sondern auch zu besseren Lernleistungen der Schülerinnen und Schüler führt (vgl. Leithwood 2011, S. 4), wird eine Aufgabe sein, der sich Schulen und Schulverwaltung in der

Zukunft stärker stellen müssen. Schulinspektion kann die schulinterne Evaluation der Elternsicht nicht ersetzen, sondern stellt nur *eine* Form der Beschreibung der Außensicht dar. In diesem Sinne wird auch zu überlegen sein, welchen Beitrag die Schulinspektion in Zukunft leisten kann, um selbstverantwortete Schulen darin zu unterstützen, Eltern als Stakeholder von Schule wahrzunehmen und die Kooperation mit ihnen dialogisch zu gestalten.

Literatur

- BSB (Hrsg.) (2006). *Orientierungsrahmen. Qualität an Hamburger Schulen*. Hamburg: Behörde für Schule und Berufsbildung. Verfügbar unter: <http://www.hibb.hamburg.de/index.php/file/download/887> [Zugriff am 9.3.2015]
- BSB & IfBQ (Hrsg.) (2013). *Jahresbericht der Schulinspektion. Schuljahr 2010/2011*. Hamburg: Behörde für Schule und Berufsbildung und Institut für Bildungsmonitoring und Qualitätsentwicklung. Verfügbar unter: <http://www.schulinspektion.hamburg.de/index.php/file/download/1782> [Zugriff am 23.07.2014]
- de Boer, H. (2012). Qualitätsentwicklung in der Grundschule. In J. Košinár (Hrsg.), *Vielfalt und Anerkennung. Internationale Perspektiven auf die Entwicklung von Grundschule und Kindergarten* (S. 198–208). Baltmannsweiler: Schneider-Verl. Hohengehren.
- Epstein, J.L. (2001). *School, Family and Community Partnerships. Preparing Educators and Improving Schools*. Boulder, Oxford: Westview.
- Fend, H. (2008). *Schule gestalten. Systemsteuerung, Schulentwicklung und Unterrichtsqualität*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Friedman, M. (1955). The Role of Government in Education. In R. A. Solo (Hrsg.), *Economics and the Public Interest* (S. 123–144). New Brunswick: Rutgers University Press.
- Fullan, M. (2001). *The New Meaning of Educational Change* (3. Aufl.). New York: Teachers College Press.
- Hirschman, A.O. (1970). *Exit, Voice and Loyalty: Responses to Decline in Firms, Organizations and States*. Cambridge: Harvard University Press.
- Hoover-Dempsey, K. & Sandler, H. (1995). Parental Involvement in Children's Education: Why Does It Make a Difference. *The Teachers College Record*, 97 (2), 310–331.
- Hoover-Dempsey, K. & Sandler, H. (1997). Why Do Parents Become Involved in Their Children's Education? *Review of Educational Research*, 61 (1), 3–42.
- Hoover-Dempsey, K. & Sandler, H. (2005). *The Social Context of Parental Involvement: A Path to Enhanced Achievement. Final Performance Report for the Office of Educational Research and Improvement (Grant No. R305T010673)*. Presented to Project Monitor, Institute of Education Sciences, U.S. Department of Education. March 22, 2005.
- Jäger-Flor, D. & Jäger, R. S. (2010). *Bildungsbarometer zur Kooperation Elternhaus–Schule 4/2009. Ergebnisse, Bewertungen und Perspektiven*. Verfügbar unter: http://vep-landau.de/Bildungsbarometer/Bildungsbarometer_2009_4.pdf [Zugriff am 25.07.2014]

- Killus, D. & Tillmann, K.-J. (Hrsg.) (2011). *Der Blick der Eltern auf das deutsche Schulsystem. 1. JAKO-O-Bildungsstudie*. Münster: Waxmann.
- Killus, D. & Tillmann, K.-J. (Hrsg.) (2012). *Eltern ziehen Bilanz. Ein Trendbericht zu Schule und Bildungspolitik in Deutschland. 2. JAKO-O-Bildungsstudie*. Münster: Waxmann.
- Leithwood, K. (2011): *Characteristics of High Performing School Systems in Ontario. Execute Summary*. Verfügbar unter: http://iel.immix.ca/storage/6/1314648943/IEL_Executive_Summary_Report_KLeithwood_Characteristics_of_HPerfroming_Bds_with_new_cover_page_August_2011.pdf [Zugriff am 18.8.2014]
- Leithwood, K., Louis, K. S., Anderson, S. & Wahlstrom, K. (2004). *How Leadership Influences Student Learning*. Center for Applied Research and Educational Improvement at the University of Minnesota and Ontario Institute for Studies in Education at the University of Toronto.
- Luhmann, N. (2000). *Vertrauen. Ein Mechanismus der Reduktion sozialer Komplexität* (4. Aufl.). Stuttgart: Lucius & Lucius.
- Luhmann, N. (2005). *Soziologische Aufklärung 5. Konstruktivistische Perspektiven* (3. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Matland, R. E. (1995). Exit, Voice, Loyalty, and Neglect in an Urban School System. *Social Science Quarterly*, 76 (3), 506–512.
- OECD: Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (2007). *PISA 2006 – Schulleistungen im internationalen Vergleich. Naturwissenschaftliche Kompetenzen für die Welt von morgen*. Paris: OECD. Verfügbar unter: <http://www.oecd.org/pisa/39728657.pdf> [Zugriff am 26.07.2014]
- Peters, S. (2014). Gemeinsam unterwegs zur inklusiven Schule – inklusive der Eltern? In S. Peters & U. Widmer-Rockstroh (Hrsg.), *Gemeinsam unterwegs zur inklusiven Schule* (S. 225–236). Frankfurt a. M.: Grundschulverband.
- Peters, S. (2015a). Eltern als Kooperationspartner von multiprofessionellen Teams – Impulsgeber für inklusive Schulentwicklung? In D. Blömer, M. Lichtblau, A.-K. Jüttner, K. Koch, M. Krüger & R. Werning (Hrsg.), *Gemeinsam anders lehren und lernen – Wege in die inklusive Bildung* (Jahrbuch Grundschulforschung, Bd. 18, S. 156–161). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Peters, S. (2015b). *Inklusive Bildung für Kinder mit komplexen Beeinträchtigungen? Elternperspektiven auf Schule*. Universität Koblenz-Landau: Unveröffentlichte Dissertation.
- Sacher, W. (2010). Elterngarbit ohne Qualitätsbewusstsein? Die fehlenden Standards. *SchulVerwaltung. Zeitschrift für Schulleitung und Schulaufsicht, Ausgabe Nordrhein-Westfalen*, 21 (6), 178–180.
- Tiefel, S. & Zeller, M. (Hrsg.) (2012). *Vertrauensprozesse in der Sozialen Arbeit*. Baltmannsweiler: Schneider-Verl. Hohengehren.
- Vodafone-Stiftung Deutschland (Hrsg.) (2013): *Qualitätsmerkmale schulischer Elternarbeit. Ein Kompass für die partnerschaftliche Zusammenarbeit von Schule und Elternhaus*. Verfügbar unter: <http://www.eltern-bildung.net/pages/publikationen/qua>

litaetsmerkmale_schulischer_elternarbeit/subpages/qualitaetsmerkmale_schulischer_elternarbeit/index.html [Zugriff am 26.07.2014]

Wagenblass, S. (2004). *Vertrauen in der Sozialen Arbeit. Theoretische und empirische Ergebnisse zur Relevanz von Vertrauen als eigenständiger Dimension*. Weinheim, München: Juventa Verlag.

Withey, M.J. & Cooper, W.H. (1989). Predicting Exit, Voice, Loyalty, and Neglect. *Administrative Science Quarterly*, 34 (4), 521–539.

Teil 4

Die Zukunft der Schulinspektion

Zukunft der externen Evaluation – Trends und Perspektiven in Deutschland

Sabine Müller

Zusammenfassung

In den letzten Jahren hat sich die Schulinspektion als fester Bestandteil der Qualitätssicherung in fast allen deutschen Bundesländern etabliert. Sie wurde mit unterschiedlichen Funktionen und Zielsetzungen eingeführt. Der Beitrag richtet nach einer kurzen Zwischenbilanz bezogen auf die verschiedenen Wirkungserwartungen den Blick auf Trends der Weiterentwicklung, die in einigen Bundesländern bereits umgesetzt werden, sowie auf mögliche langfristige Perspektiven für die künftige Rolle der Schulinspektion in Deutschland und die dazu notwendigen Voraussetzungen und Gelingensbedingungen im Kontext eines Gesamtsystems von Qualitätsentwicklung und -sicherung.

Seit gut zehn Jahren spielt externe Evaluation in den bildungspolitischen Steuerungsstrategien in Deutschland eine zentrale Rolle als verpflichtender Baustein der Qualitätssicherung. Dies erfolgte im Zuge eines Paradigmenwechsels in der Organisations- und Steuerungsphilosophie von einer Input- und Prozesssteuerung zur Outputsteuerung (vgl. Dederling & Müller 2008a, S. 241). Dabei steht die Schulinspektion als Verfahren externer Evaluation im Fokus, die im internationalen Kontext über eine teilweise jahrzehntelange Tradition verfügt.

Die einzelnen Bundesländer haben dabei individuelle Wege beschritten; bezogen auf das Verfahren und dessen Ablauf gibt es viele Parallelen. Inzwischen haben einige Länder einen ersten Zyklus durchlaufen und alle Schulen in ihrem Zuständigkeitsbereich bereits durch ein externes Evaluationsteam inspiziert und für den weiteren Einsatz des Verfahrens Kurskorrekturen und Modifizierungen vorgenommen. Bis auf eine Ausnahme (Schleswig-Holstein) geht der Trend in allen Ländern in Richtung einer dauerhaften Etablierung von externer Evaluation.

Die Bundesländer unterscheiden sich nicht nur in den entsprechenden Bezeichnungen (u. a. Schulinspektion, Schulvisitation, Qualitätsanalyse), sondern beispielsweise auch hinsichtlich der Anbindung der externen Evaluationen, der Zusammensetzung der Teams, die die Evaluation durchführen, hinsichtlich der Instrumente, der Dauer der Schulbesuche und der zugrunde liegenden Qualitätsaspekte, die in Referenz- bzw. Orientierungsrahmen oder Qualitätstableaus festgeschrieben sind.

In diesem Beitrag soll auf der Grundlage der zu beobachtenden Entwicklungen in den Ländern versucht werden, einen Blick in die Zukunft zu werfen und vorsichtige Prognosen bzw. Thesen für die künftige Rolle und Entwicklung externer Evaluations- bzw. Inspektionsverfahren zu wagen, die naturgemäß ein Stück weit Spekulation sind. Der Fokus richtet sich dabei ausschließlich auf die nationale Perspektive in Deutschland. Aufgrund der föderalen Unterschiede werden dabei länderübergreifend nur generelle Trends und Richtungsentscheidungen in den Blick genommen.

Zunächst sollen die Funktionen, Ziele und (Wirkungs-)Erwartungen skizziert werden, mit denen die Schulinspektionen in Deutschland eingeführt wurden. Anschließend werden schlaglichtartig bisherige Erfahrungen und Erkenntnisse bilanziert – eine umfassende Bestandsaufnahme bezogen auf die einzelnen Bundesländer würde einen eigenen Beitrag erfordern. Daraus werden Trends und Perspektiven abgeleitet, die sich teilweise in den Umsteuerungen der Länder abzeichnen. Zum Abschluss werden noch Desiderata formuliert, die als mögliche Gelingensfaktoren für eine erfolgreiche Einbindung von externer Evaluation in eine Gesamtstrategie von Qualitätsentwicklung eine Rolle spielen können.

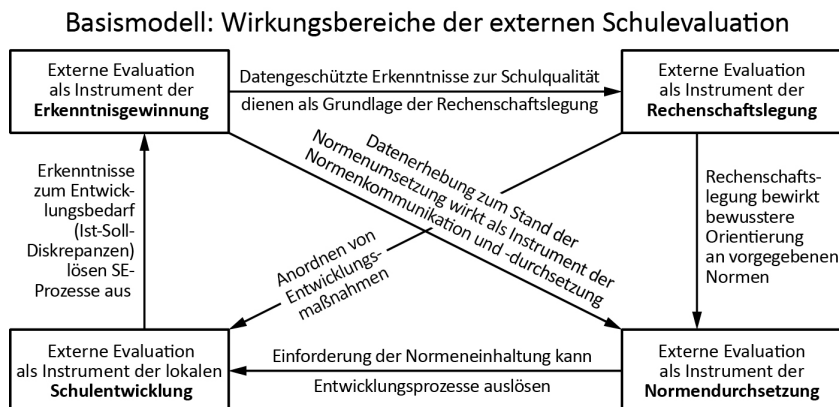
1. Externe Evaluation: Erwartungen und Grenzen

Externe Evaluation wird mit unterschiedlichen Funktionen und Zielsetzungen verbunden. Diese stehen auch in Zusammenhang mit Wirkungsmodellen, die in den Strategien der Bundesländer bei der flächendeckenden Einführung der Inspektionsverfahren mit unterschiedlichen Akzentsetzungen eine Rolle gespielt haben.

1.1 Funktionen und Wirkungsmodelle

Ein Wirkungsmodell aus dem Forschungskontext ist das von Quesel, Husfeldt, Landwehr und Steiner (2011), das vier Wirkungsbereiche externer Evaluation nennt: die Erkenntnisgewinnung, die Rechenschaftslegung, die Normendurchsetzung und die Schulentwicklung (vgl. Abbildung 1).

Abbildung 1: Wirkungsbereiche externer Schulevaluation



Quelle: Quesel et al. 2011

Diese Wirkungsbereiche gelten für die Ebene der Einzelschulen sowie auch für das Gesamtsystem und sind miteinander verbunden. Erkenntnisgewinnung erfolgt für die einzelnen Schulen durch die Rückmeldung ihrer Ergebnisse. Für die Systemebene – z. B. Landesebene – liefern landesweite Auswertungen Daten, die als eine Grundlage für ein Systemmonitoring bzw. für die Generierung administrativen Steuerungswissens genutzt werden können. Die Rechenschaftslegung der Schulen richtet sich sowohl auf den schulinternen als auch auf den schulexternen Kontext. Schulexterne Evaluation soll sowohl Impulse für die Schulentwicklung der inspizierten Schulen geben als auch der Steuerungsebene bei der Bereitstellung bedarfsgerechter und geeigneter Unterstützungsmaßnahmen helfen. Durch eine Festlegung der Gegenstandsbereiche der Evaluation in Orientierungsrahmen bzw. Qualitätstableaus soll eine normierende Wirkung erreicht werden, indem der Blick der Schulen auf die dort festgelegten Aspekte von Schulqualität gelenkt wird (andererseits wird dies auch als Implementationsfunktion bezeichnet).

In den unterschiedlichen Bundesländern in Deutschland kann von einem Funktionenmix bei der Einführung der externen Evaluation mit verschiedenen Akzenten und Schwerpunktsetzungen gesprochen werden. Maritzen (2006) schreibt nach Sichtung der Ansätze in den Bundesländern dem Verfahren zusammenfassend eine *Spiegel- oder Feedbackfunktion* aus externer Sicht und eine *Unterstützungsfunktion* insbesondere für Schulleitungen und Lehrkräfte zu. Ferner führt er die *Impulsfunktion für die Schul- und Unterrichtsentwicklung* sowie die *Erkenntnisfunktion* in Bezug auf die Wirkungen schulischer Arbeit an (vgl. Maritzen 2006).

Die Gewinnung von Steuerungswissen auf administrativer Ebene, die Rechenschaftslegung und die Unterstützung der Einzelschulen wurden weitgehend in den

Ländern als Funktion festgelegt. Strittig wurden Zielsetzungen und Wirkungserwartungen im Zusammenhang mit Schülerleistungen diskutiert, die teilweise gefordert, von den Ländern jedoch nicht als primäres bzw. explizites Ziel formuliert wurden.

1.2 Grenzen standardisierter Inspektionsverfahren

Was Inspektionsverfahren konkret zu leisten vermögen, wird in unterschiedlichen Kontexten kontrovers diskutiert. Im Kern geht es um Rückmeldungen – in der Regel verbunden mit Bewertungen – von geschulten schulexternen Expertinnen und Experten zu unterschiedlichen Aspekten von Schulqualität mit weitgehend standardisierten Verfahren und Instrumenten. Bei aller Standardisierung handelt es sich letztlich um ein Expertenrating, in dem die subjektive Einschätzung der als Evaluatorinnen und Evaluatoren eingesetzten Personen bei der Bewertung eine nicht unbedeutende Rolle spielt. Insofern kann – trotz standardisierter Instrumente – nicht von objektiven „Messungen“ der Qualität von Einzelschulen ausgegangen werden. Und bei allen Wirkungserwartungen darf nicht vergessen werden, dass bereits durch die Anlage der Verfahren standardisierter Inspektion Grenzen gesetzt sind, die insbesondere die Erwartungen an die Impulsfunktion für Schulentwicklung einschränken:

- Schulinspektion leistet in der Regel keine fachliche Rückmeldung zum Unterricht.
- Die Erkenntnisse der Inspektionsteams münden in Bewertungen, die den Schulen in einem Bericht zur Verfügung gestellt werden. In der Regel werden keine Empfehlungen, Tipps oder Hilfestellungen gegeben, da die Rolle der Inspektion klar von einer beratenden Funktion abgegrenzt wird.
- Die Dokumentation der Ergebnisse erfolgt auf Schulebene. Die Inspektion gibt keine Rückmeldungen, geschweige denn konkrete Hinweise zur Weiterentwicklung zum Unterricht einzelner Lehrkräfte – Lehrer erhalten keine unmittelbaren Anregungen und Hilfen.
- Es handelt sich um eine administrativ „verordnete“ Rückmeldung mit extern festgelegten Themen.
- Auch wenn die Personen, die die Inspektionen durchführen, gegenüber den Schulen nicht unmittelbar weisungsbefugt sind, handelt es sich in der Regel um ein hierarchisches Verhältnis, das insbesondere durch den Aspekt der Bewertung geprägt ist.
- Sie ist ein eher singuläres Ereignis, das die einzelnen Schulen je nach Zyklus im Abstand von mehreren Jahren betrifft.

- Letztlich sind die Konsequenzen im positiven wie im negativen Sinne begrenzt – es erfolgen weder umfangreiche Unterstützungen in personeller oder materieller Hinsicht noch „Sanktionen“.

Vor diesem Hintergrund scheint es schwierig, ein solches Verfahren (primär) als Unterstützungsinstrument für Schulen zu begreifen. Die Funktion als Impulsgeber für Schulentwicklung bleibt fraglich, wenn nicht andere Akteure unterstützend hinzukommen und eine Einbettung in ein Gesamtsystem der Qualitätsentwicklung und -sicherung gewährleistet ist.

Die dargestellten Einschränkungen beziehen sich auf die Verfahrensmerkmale einer standardisierten Inspektion, die in den einzelnen Ländern nicht in allen Elementen so zutrifft. Sie dienen hier vielmehr dazu, plakativ und pointiert auf Grenzen hinzuweisen. Was die Inspektionsverfahren bislang bezogen auf ihre Ziele und (erhofften) Wirkungen geleistet haben, ist in den Ländern auf unterschiedlichen Ebenen erforscht und ausgewertet worden.

2. Erkenntnisse zu Wirkungen externer Evaluation

Konkrete Wirkungen oder Effekte externer Evaluationen bzw. Schulinspektionen empirisch eindeutig nachzuweisen, ist schwierig. In einem komplexen Kontext des Bildungssystems sind monokausale Wirkzusammenhänge kaum seriös anzunehmen. So sind Wirkungen hinsichtlich der beschriebenen Kernfunktionen, die mit externer Evaluation verbunden sind (Generierung von Steuerungswissen, Rechenschaftslegung, Implementation und Impulsgeber für Schulentwicklung) in unterschiedlichem Maße erforscht und empirisch belegt.

Inzwischen liegen Forschungsergebnisse aus verschiedenen wissenschaftlichen Studien vor, die von den Ländern und von Universitäten durchgeführt wurden (vgl. u. a. Böttcher et al. 2010; Müller 2010; Müller et al. 2011; Dederling & Müller 2008b und 2011; Dederling et al. 2012; Böhm-Kaspar & Selders 2013).

Ferner gibt es auch zahlreiche Veröffentlichungen der Länder mit (landesweiten) Auswertungen der Daten aus den externen Evaluationen. Die Inspektionen selbst haben inzwischen in großem Umfang differenzierte Daten zu Schul- und Unterrichtsqualität produziert, die in der Regel in den Ländern ausgewertet und dokumentiert werden (vgl. u. a. Institut für Bildungsmonitoring und Qualitätsentwicklung Hamburg 2013).

2.1 Externe Evaluation als Steuerungsinstrument?

Die Länder haben (Jahres-)Berichte mit den bisherigen Erkenntnissen veröffentlicht, in denen landesweite Ergebnisse dokumentiert sind, die als Grundlage für

Steuerungswissen dienen sollen. Welche (unmittelbaren) bildungspolitischen Steuerungserkenntnisse daraus gezogen wurden und werden und wie diese in konkrete Maßnahmen umgesetzt werden, bleibt diffus. Inwieweit die erheblichen Datenmengen aus den Schulbesuchen zur Funktion des Steuerungswissens beigetragen haben und dem Gesamtsystem wie den Einzelschulen als Instrument der Rechenschaftslegung dienen, lässt sich vor allem mit Blick auf daraus abgeleitete Handlungen kaum bestimmen.

Zunächst ist festzustellen, dass durch die Schulinspektion in einem großen Umfang differenzierte Daten zu unterschiedlichen Aspekten schulischer Qualität vorliegen. Landesweite Auswertungen der Ergebnisse von Einzelschulen liefern erst einmal Informationen, die von den beteiligten Akteuren in ein Handlungswissen transformiert werden müssen, um entsprechende Steuerungen daraus ableiten zu können. Hier spielen auch bildungspolitische Faktoren eine Rolle sowie finanzielle, materielle und personelle Ressourcen, die für diesen Bereich zur Verfügung stehen bzw. eingesetzt werden können. Bildungspolitische Steuerung steht immer in einem Gesamtkontext, in dem Inspektionsergebnisse eine Datenquelle liefern, deren Bedeutsamkeit schwer eindeutig zu bestimmen ist. Im Rahmen des BMBF Forschungsprogramms „Steuerung im Bildungssystem“ (SteBis) wurde mit einer qualitativen Untersuchung die Rolle des Schulträgers näher beleuchtet (vgl. Preuß 2013). Dabei wird unter anderem festgestellt, „dass der Schulträger im Kontext von Schulinspektion gerne ein echter ‚Mitspieler‘ wäre und für ihn die Frage nach seinem Steuerungseinfluss eine Grundsatzfrage ist“ (ebd., S. 167).

2.2 Externe Evaluation als Impulsgeber für Schulentwicklung?

Am ehesten sind Erkenntnisse zur Unterstützungsfunktion aus Schulsicht zu berichten. Aus verschiedenen Untersuchungen (der Länder sowie von unabhängiger Forschung) lässt sich ableiten, inwieweit Schulen die Schulbesuche und die mündlichen und schriftlichen Rückmeldungen und Bewertungen als Impuls für ihre weiteren Entwicklungsprozesse sehen.

Befragungen von Schulleitungen und Lehrkräften zeigen, dass die anfänglichen Befürchtungen der Schulen nicht eintreten und die Schulbesuche weniger unangenehm ausfallen als erwartet. Ferner scheinen die Ergebnisse und Rückmeldungen aus Sicht von betroffenen Schulen insgesamt geeignet, Impulse für Schulentwicklung zu geben. Allerdings gibt es hierzu vorwiegend Studien, die auf Einschätzungen und Wirksamkeitserwartungen der Akteure basieren. Befragungen zeigen in der Tendenz, dass Schulen (insbesondere Schulleitungen) im Nachgang zu den Schulbesuchen als grundsätzlich hilfreich ansehen (vgl. u. a. Müller 2010; Sommer 2011; Böhm-Kaspar & Selders 2013), dass sie Entwicklungen befruchten; ein

„harter“ empirischer Nachweis für einen Einfluss auf Schulentwicklung kann nicht erbracht werden, da solche monokausalen Wirkzusammenhänge und eindeutige Rückführungen auf die externe Evaluation eher schwierig sind.

Husfeldt (2011) fasst die Erkenntnisse verschiedener Studien im deutschsprachigen Raum zur Wirksamkeit externer Evaluation zusammen, die auf Befragungen der Beteiligten beruhen. „Insgesamt liegt der Schwerpunkt der Studien eher auf der Erfassung einer zielunabhängigen Wirkung bzw. der Wirksamkeitseinschätzung durch die Akteure als auf einer an konkret formulierten Zielen abgeglichenen Wirksamkeit.“ (ebd., S. 278) Ferner formuliert Husfeldt als Forschungsdesiderat ein Wirkungsmodell, das die Verarbeitungsprozesse in der Schule und die administrative Seite in den Blick nimmt. „Die Maßnahmen werden als Wirkung der externen Schulevaluation angesehen, ohne dass überprüft wird, ob sie hinsichtlich der von der Schule aufgrund des Evaluationsberichts formulierten Zielvereinbarungen auch wirksam sind bzw. sein können.“ (ebd.)

Ferner gibt es auch Untersuchungen zur konkreten (quantitativen) Erfassung von Maßnahmen, die Schulen in Folge externer Schulevaluationen initiiert haben (vgl. u. a. Wurster & Gärtner 2013). Sie ordnen diese Maßnahmen unterschiedlichen Evaluationstypen zu und weisen nach, dass eine Entwicklungsfunktion nicht für alle Schulen gleichermaßen angenommen werden kann. Diese gilt insbesondere für „aktive“ und „aktiv unzufriedene“ Schulen (ebd., S. 441). In den anderen Schulen zeigen sich weniger Entwicklungsaktivitäten. Die Studie von Wurster und Gärtner mit Schulen aus Berlin und Brandenburg zeigt ferner, dass ein Großteil von Maßnahmen und Aktivitäten der Schulen im Vorfeld der externen Evaluation stattfinden (62 % der Schulen berichten von Maßnahmen im Vorfeld). Die Autoren weisen in diesem Zusammenhang auf die Funktion der Normdurchsetzung bzw. Implementation hin, da die Schulen sich im Vorfeld mit den Gegenstandsbereichen der Evaluation auseinandersetzen.

Man könnte aber auch weitergehend interpretieren, dass Schulen sich – wie ausdrücklich *nicht* intendiert – auf die externe Evaluation vorbereiten und dann ein „*Window Dressing*“ betreiben. Ob die dokumentierten Aktionen nach erfolgtem Schulbesuch und Erhalt des Berichts fortgeführt werden, ist eine andere Frage. Ebenso bleibt fraglich, inwieweit eine nachhaltige Implementationsfunktion in der Fläche aller Schulen gegeben ist, die über den Kontext der Vorbereitung auf eine anstehende Inspektion in den Einzelschulen hinausgeht.

2.3 Wirkungen externer Evaluation auf Schülerleistung?

Insbesondere die Diskussion, ob Auswirkungen von Schulinspektionen auf Schülerleistungen überhaupt zu erwarten und faktisch zu verzeichnen sind, ist problematisch und wird kontrovers geführt. Ein Ergebnis von Husfeldt nach Betrachtung

verschiedener empirischer Studien deutet einen eher gegenteiligen Trend an. „Dabei zeigt sich meist, dass die Leistung im Zusammenhang mit externer Evaluation kurz- bis mittelfristig eher leidet.“ (Husfeldt 2011, S. 277) Pietsch, Janke und Mohr (2014) kommen zu einem anderen Ergebnis und resümieren nach einer Studie mit Hamburger Schulen vorsichtig, dass sich unter dem Paradigma, durch Schulinspektion Schulentwicklung stimulieren zu wollen, empirische Hinweise verdichten, „dass Schulinspektion mit dem Blick auf Schülerleistungen, zumindest unter diesem Paradigma, nicht schadet und ggf. sogar positive Effekte nach sich ziehen kann [...] Die Befunde zeigen, dass bei Einsatz maßgeschneiderter kausalanalytischer Verfahren sowohl Effekte auf Lernzuwächse als auch Leistungstrends von Schülerinnen und Schülern in Hamburg nachgewiesen werden können.“ (Pietsch et al. 2014, S. 466) Die Effekte variieren dabei zwischen den beteiligten Fächern. Die Ergebnisse basieren auf zwei Messzeitpunkten, sodass noch keine Aussagen zur Nachhaltigkeit der Effekte getroffen werden können.

Insgesamt bleiben wissenschaftliche Nachweise hinsichtlich einer Wirksamkeit von Schulinspektion auch deshalb problematisch, da unterschiedliche Kontextfaktoren eine Rolle spielen. Dazu gehört die unterschiedliche Akteurskonstellation in den Ländern, die einen länderübergreifenden oder -vergleichenden Ansatz erschweren. Auch Längsschnittuntersuchungen innerhalb der Länder sind in ihrer Aussagekraft begrenzt, da in den meisten Ländern Umsteuerungen im Verfahren und in den Instrumenten vorgenommen wurden. Und angesichts der Komplexität der Zielsetzungen und Erwartungen ist eine konkrete Evaluation der Wirksamkeit schwierig. Es müssten zunächst im Land konkrete Ziele und Erwartungen sowie entsprechende Erfolgsindikatoren formuliert werden, die dann wissenschaftlich überprüft werden können.

3. Trends und Perspektiven

Seit der Einführung der externen Evaluation haben die Bundesländer unterschiedliche Entwicklungen durchgeführt und teilweise Umsteuerungen auch hinsichtlich der externen Evaluation vorgenommen. Fast alle Länder haben dabei bislang entschieden, die Inspektion als verpflichtendes Verfahren der externen Evaluation beizubehalten, wenn auch mit zum Teil erheblichen Modifizierungen. Die Entwicklungsstände in den Ländern sind – teilweise auf aufgrund der großen Varianz der Anzahl von Schulen – unterschiedlich. In den meisten Bundesländern ist mindestens ein erster Zyklus von Inspektionen in allen Schulen abgeschlossen und eine zweite Phase hat begonnen. Die Modifizierungen beziehen sich sowohl auf die Inspektion im Kern als auch auf andere Maßnahmen zur Qualitätssicherung, die die Rolle der Inspektion im Gesamtsystem anders ordnet. Wie eingangs beschrieben, sollen hier nicht die Entwicklungen in den einzelnen Ländern

nachgezeichnet werden, sondern generelle Trends, die Grundlage für mittel- und längerfristige Prognosen zur Entwicklung von Inspektionen in Deutschland sein können.

Veränderungen hinsichtlich der Inspektionen beziehen sich u. a. auf

- Überarbeitungen der Orientierungs- bzw. Referenzrahmen, die Grundlage für die Rückmeldungen an die Schulen sind, und teilweise Fokussierungen bzw. Schwerpunktsetzungen mit der Option für die Schulen, zusätzlich eigene Schwerpunkte zu wählen,
- Überarbeitungen der Erhebungsinstrumente.

Daneben sind in zahlreichen Ländern Entwicklungen erkennbar, die sich auf eine Neujustierung unterschiedlicher Verfahren der Qualitätsentwicklung und -sicherung beziehen und die externe Evaluation als Baustein in ein Gesamtkonzept einbetten.

Neben der externen Evaluation als zentralem Steuerungsinstrument ist in stärkerem Maß die interne Evaluation (wieder) in den Fokus gerückt.

3.1 Verzahnung von interner und externer Evaluation

Betrachtet man die steuerungspolitischen Konzepte zur Qualitätsentwicklung und -sicherung in den Bundesländern, so sind in regelmäßigen Abständen „Wellenbewegungen“ erkennbar, in denen der jeweilige Fokus wechselt. In einigen Phasen steht dabei Stärkung der Einzelschule im Mittelpunkt mit ihrer Verantwortung für ein kontinuierliches Qualitätsmanagement. Hierzu gehören auch Maßnahmen wie die Einrichtung von Steuergruppen und die Verpflichtung zu Schulprogrammarbeit sowie zu Fortbildungen und interner Evaluation (Inputsteuerung). In anderen Phasen stehen Verfahren des Systemmonitorings und der externen Kontrolle und Rechenschaft sowie schulstrukturelle Maßnahmen im Vordergrund (Outputsteuerung). Möglicherweise steht eine Phase bevor, in der beide Paradigmen in Einklang gebracht werden und sich sinnvoll ergänzen. Voraussetzung dafür ist eine starke und professionelle interne Evaluationskultur in den Schulen.

Im Zuge der Weiterentwicklung der externen Evaluationsverfahren ist die Zielperspektive der Impulsfunktion für interne Schulentwicklungsprozesse in einigen Bundesländern noch deutlicher als in der Implementationsphase in den Fokus gerückt. Eine externe Evaluation kommt jedoch punktuell und viel zu selten – nämlich im Abstand von mehreren Jahren –, um einen kontinuierlichen Schulentwicklungsprozess zu gewährleisten. Die Sicherung der Nachhaltigkeit solcher Prozesse muss in der Einzelschule erfolgen. Dies erfordert ein systematisches internes Qualitätsmanagement, zu dem auch eine regelmäßige Selbstvergewisserung durch schulin-

terne Evaluationen gehört. Idealerweise beziehen sich dabei interne und externe Evaluation so aufeinander, dass die Schulen davon profitieren können. Folglich erscheint eine enge Verzahnung von interner und externer Evaluation als ein vielversprechender Ansatz, um Schulentwicklung zu fördern. Ähnliche Überlegungen haben dazu geführt, dass in verschiedenen Bundesländern beide Verfahren stärker aufeinander bezogen werden, wobei die interne Evaluation ein deutlich höheres Gewicht erhält. Dabei sind die Schulen aufgefordert, unter anderem auch die Instrumente der externen Evaluation für ihre internen Evaluationen zu verwenden, um die internen Einschätzungen mit der Perspektive der externen Experten abgleichen zu können. Um die Schulen bei der internen Evaluation zu unterstützen, haben die Länder zum Teil Portale entwickelt, die die Schulen in der Regel online-basiert für die interne Evaluation nutzen können. Hier gibt es auch Angebote für fokussierte interne Evaluationen, die die konkreten Fragestellungen der Einzelschulen in den Blick nehmen.

Insgesamt könnte der Trend einer langfristigen Perspektive dahin gehen, dass externe Evaluationen sich auf die Prozesse und jeweiligen (Weiter-)Entwicklungen der Schulen beziehen und insbesondere auf die Qualität der intern erhobenen Daten. Je stärker die Evaluationskultur einer Schule ausgebaut ist, desto eher kann eine Balance zwischen interner und externer Evaluation erreicht werden. Und je besser eine Schule intern evaluiert, desto fokussierter kann eine externe Evaluation ausfallen. Gleichzeitig sollten bestimmte Qualitätsaspekte der externen Evaluation als verbindliche Gegenstandsbereiche der externen Evaluation erhalten bleiben, um über alle Schulen hinweg einen landesweiten Datenpool für eine kontinuierliche Gewinnung von Steuerungswissen zu erhalten.

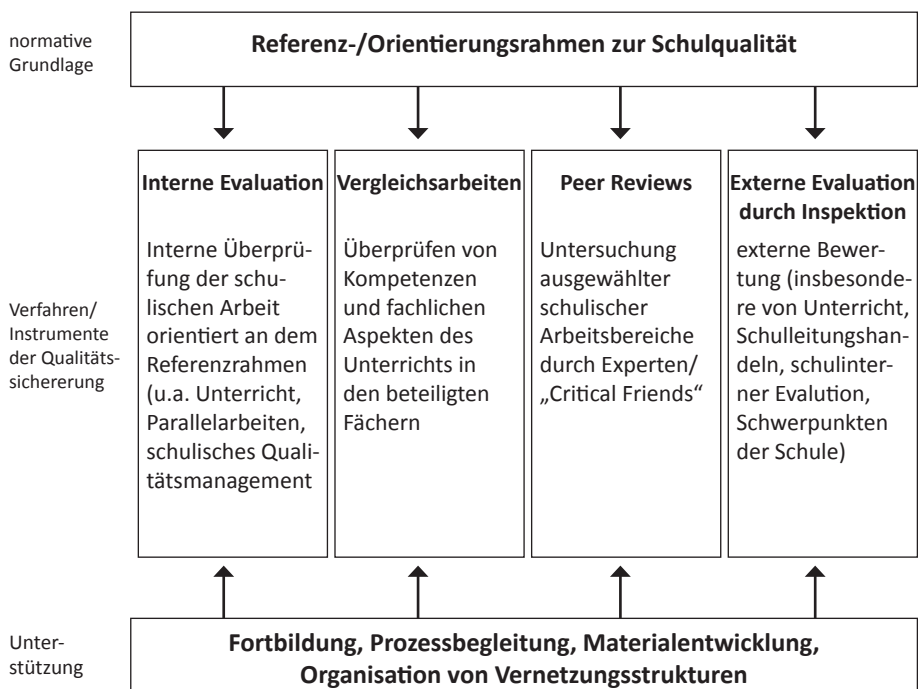
3.2 Externe Evaluation als Baustein im Gesamtsystem von Qualitätsentwicklung und -sicherung

Wie oben beschrieben, ist von externer Evaluation als Einzelmaßnahme für eine kontinuierliche Schulentwicklung keine nachhaltige Wirksamkeit zu erwarten, die über eine (singuläre) Impulsfunktion hinausgeht. Dazu wird ein stimmiges Gesamtkonzept benötigt, das verschiedene Maßnahmen und Akteure in einen gemeinsamen Handlungskontext bringt. Wichtig dabei ist eine Passung der Rollen aller Akteure, die in den Kreislauf der Schulentwicklung eingebunden sind (Schulaufsicht, Schulinspektion, Unterstützungssystem, Einzelschule). Zielvereinbarungen zwischen Schule und Schulaufsicht können helfen, Entwicklungsprozesse verbindlich zu steuern und Kontinuität mit orientierenden „Meilensteinen“ zu sichern. Die Schulinspektion liefert in bestimmten Abständen einen Blick von außen auf den Stand und Entwicklungsprozess der Einzelschule, der für die Weiterentwicklung genutzt wird. Dabei sollten Schulen auf ein Unterstützungssystem zugreifen können, das

eine professionelle Begleitung in Fragen von Qualitätsmanagement, Schul- und Unterrichtsentwicklung ermöglicht.

In verschiedenen Ländern sind Orientierungs- oder Referenzrahmen entwickelt worden, die ein gemeinsames Verständnis der beteiligten Akteure hinsichtlich der Zielrichtungen von Schulqualität festlegen und einen gleichsinnigen Handlungsrahmen definieren. Sie dienen der Orientierung und bilden gleichsam ein normatives „Dach“ für Qualitätsentwicklung und -sicherung und sind auch Bezugsrahmen für interne und externe Evaluation. Schulen können mithilfe interner Evaluationen überprüfen, wo sie hinsichtlich unterschiedlicher Qualitätskriterien „stehen“, und eine externe Evaluation gibt Rückmeldungen auf der Grundlage solcher Rahmen.

Abbildung 2: Qualitätsentwicklung und Qualitätssicherung



Die im Referenzrahmen festgelegten Qualitätsaspekte und -kriterien bieten Orientierung für alle Beteiligten und geben der Schule systematische Anhaltspunkte für eine Standortbestimmung als Grundlage für einen kontinuierlichen, datenbasierten Schulentwicklungsprozess. Dabei spielen auch Daten aus den landesweiten Vergleichsarbeiten und Lernstandserhebungen eine Rolle, die eine Datenbasis für in-

terne Evaluationen liefern. Ebenso können Ergebnisse von Parallelarbeiten in eine systematische Unterrichtsentwicklung einbezogen werden. Die Dokumentation der Entwicklung kann im Schulprogramm erfolgen, um Transparenz für die Schulöffentlichkeit (die intern und extern Beteiligten) sowie auch Orientierungen für die Schulinspektion zu gewährleisten.

Ein weiteres Verfahren der Evaluation ist das *Peer Review*, das unter anderem in den Deutschen Schulen im Ausland durchgeführt wird (vgl. Gieske-Roland et al. 2014). Dabei handelt es sich um Rückmeldungen durch externe Experten, die die Schule selbst mit auswählt und bestimmt. Peers können beispielsweise Kolleginnen und Kollegen aus anderen Schulen sein, aber auch Experten, die die Schule als „*Critical Friend*“ einlädt und die Rückmeldungen auch mit Beratung verbinden können. Ferner können hier fachliche Aspekte in den Blick genommen werden, die in den gängigen Inspektionsverfahren nicht im Vordergrund stehen.

Wichtig ist insgesamt ein Unterstützungssystem, das auf dieser Grundlage adressatenbezogene Angebote vorhält und die Schulen bei ihrer Entwicklung unterstützt. Neben einem vom Land administrierten Unterstützungssystem mit Fortbildungsangeboten und Materialien spielen angesichts knapper Ressourcen schulische Netzwerke eine immer wichtigere Rolle. Hier können Schulen voneinander lernen und sich gegenseitig sowohl mit Feedback als auch mit fachlichen Expertisen unterstützen.

3.3 Perspektiven und Gelingensbedingungen

In den vorangegangenen Kapiteln wurden Entwicklungen skizziert, die als Trends in verschiedenen Bundesländern bereits erkennbar sind. Abschließend sollen nun noch einige Perspektiven formuliert werden, die als Ansprüche und Gelingensfaktoren für eine sinnvolle Nutzung von externer Evaluation bedeutsam sind.

Wenn man den Blick auf Ansprüche an externe Evaluation richtet, so haben viele Faktoren Einfluss auf die Qualität und die Akzeptanz sowie auch auf Wirksamkeitserwartungen des Verfahrens, u. a.:

- Eine eindeutige Zielsetzung und Rollenbeschreibung der Inspektion, die für alle Beteiligten transparent ist (und die im Kontext einer klaren Rollendefinition anderer Akteure wie Schulaufsicht, Fortbildungs- und Unterstützungssysteme steht). Um Effekte und Wirkungen von Inspektion überhaupt evaluieren zu können, müssen erst einmal konkrete Zielstellungen und Erfolgsindikatoren definiert und festgelegt werden.
- Keine „Überfrachtung“ der Rolle mit (überhöhten?) Wirkungserwartungen, die sich auf unterschiedliche Ebenen beziehen. Die Anlage und Gestaltung von externer Evaluation kann je nach Zielsetzung variieren. Eine Inspektion, die in

erster Linie der Normdurchsetzung und Rechenschaftslegung dient, kann gut mit standardisierten Grundlagen und Verfahren arbeiten; geht es aber vorrangig um Impulse für Schulentwicklung, so erfassen standardisierte Rückmeldungen in Form von Bewertungen die Frage- und Problemstellungen der Schulen nicht immer adäquat. Es ist fraglich, ob alle weiter oben beschriebenen Wirkungsebenen gleichermaßen „bedient“ werden können. Hier wäre ein Ansatz für weitere wissenschaftliche Studien.

- Eine gute Qualifizierung und Fortbildung der Inspektorinnen und Inspektoren, die ein gleichsinniges Rollenverständnis und Handeln ermöglicht, sowie kontinuierliche Fortbildungen.
- Handhabbare und transparente Erhebungsmethoden, die dem Verfahren entsprechen, sowie Rückmeldeverfahren, die Impulse für Schulentwicklung ermöglichen.
- Die Schule mit ihrem Entwicklungsprozess in den Mittelpunkt stellen und die Aufgaben der anderen Akteure so „konzertieren“, dass sie die Schulen unterstützen. Dazu gehört, dass die Schulen selbst Themen in die Inspektion einbringen können, ihre internen Evaluationsergebnisse eine Rolle spielen sowie auch die Entwicklungsprozesse und -ergebnisse, die zwischen den Inspektionszeitpunkten stattfinden.

Hinzu kommen weitere (Gelingens-)Faktoren, die mit der externen Evaluation zusammenspielen:

- Ausbau von Netzwerken, um die Potenziale zu nutzen, dass Schulen sich gegenseitig systematisch unterstützen und anregen können.
- Passende und zeitnahe Unterstützungsangebote von verschiedenen Anbietern, vor allem regional, ggf. für mehrere Schulen einer Region gemeinsam.
- Vermittlung von Evaluationskompetenz in der Lehrerbildung.
- Stärkungen der Selbstevaluationskompetenz in den Schulen.
- Systematische Kooperationsstrukturen von Administration und Wissenschaft. Einschlägige Wirkungs- und Evaluationsstudien könnten zielgerichteter erfolgen, wenn für die Inspektion konkrete Ziele und Erfolgsindikatoren formuliert würden, die über vage Funktionsbeschreibungen hinausgehen. Auch die wissenschaftliche Auswertung der Inspektionsdaten könnte dann noch stärker zur Generierung von Steuerungswissen beitragen.

Daten aus externen Rückmeldungen allein führen nicht automatisch zu nachhaltigen Entwicklungen. Wichtig ist es, die Kompetenzen von Schulen hinsichtlich interner Evaluation und für den Aufbau eines systematischen schulinternen Qualitätsmanagements zu unterstützen und deutlich auszubauen. Ob externe Evaluation im Sinne der Schulinspektion erfolgreich sein kann, hängt also sehr stark mit ihrer

Einbindung in ein Gesamtsystem von Qualitätsentwicklung und Qualitätssicherung zusammen. Ferner sollte über weitere, „weichere“ Verfahren als Ergänzung nachgedacht werden, wie beispielsweise Peer Reviews, die einer Impuls- und Unterstützungsfunktion noch anders gerecht werden können, da sie sich an den Einzelschulen orientieren, und bei dem die externen Experten den Schulen konkrete Anregungen geben können. Entwicklungsprozesse sollten von den eigenverantwortlichen Schulen aus gedacht werden und nicht bei einer externen Evaluation beginnen. „Idealerweise sollten extern praktizierte Qualitätsmanagementverfahren erst eingeführt werden, wenn die Lehrerschaft und die Schulleitung mit den Verfahren der Selbstevaluation so vertraut sind, dass die externe Evaluation nicht als Bedrohung und Kontrolle erlebt wird.“ (Kuhn 2015, S. 104) Die landesweiten Auswertungen der Inspektionsergebnisse zeigen, dass hier noch Handlungsbedarf besteht.

Als langfristiger Trend wäre eine Proportionalität von internen und externen Verfahren zu denken. Je professioneller die eigenverantwortliche Schule sich selbst evaluiert, desto fokussierter kann eine externe Evaluation stattfinden (wie beispielsweise in den Niederlanden). Die Verantwortung für Qualitätsentwicklung und deren Sicherung würde dann bei den Schulen liegen, denen dazu ein adäquates Unterstützungs- und Beratungssystem zur Seite stehen muss. Die externe Evaluation kann dann zum einen eine gezielte Rechenschaftsfunktion erfüllen und zum anderen zu einem konkreten Systemmonitoring beitragen. So könnte ein Kreislauf professioneller und kontinuierlicher Qualitätsentwicklung und -sicherung umgesetzt werden.

Literatur

- Böhm-Kaspar, O. & Selders, O. (2013). „Schulinspektionen sollten regelmäßig durchgeführt werden.“ In I. van Ackeren, M. Heinrich, F. Thiel (Hrsg.), *Evidenzbasierte Steuerung im Bildungssystem?* (Die Deutsche Schule: Beiheft 12, S. 121–153). Münster: Waxmann.
- Böttcher, W., Keune, M. & Neiwert, P. (2010). *Evaluationsbericht zum Projekt „Schulinspektion in Hessen – Wirkungen auf die Qualitätsentwicklung von Schulen und die Arbeit der Schulaufsicht“*. Münster: WWU.
- Bonsen, M. (2014). Qualität in der Schule. Zur Diskussion um Qualitätsmanagement in der Schule. *Lernende Schule*, 17 (66), 4–7.
- Dederig, K., Fritsch, K. & Weyer, Ch. (2012). Die Ankündigung von Schulinspektionen und deren innerschulische Effekte – hektisches Treiben oder genügsame Gelassenheit? In S. Hornberg & M. Parreira do Amaral (Hrsg.), *Deregulierung im Bildungswesen* (S. 205–222). Münster: Waxmann.
- Dederig, K. & Müller, S. (2008a). Schulinspektion in Deutschland – Forschungsbereiche und -desiderata. In W. Böttcher, W. Bos, H. Döbert & H.G. Holtappels (Hrsg.), *Bil-*

- dingsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive* (S. 241–252). Münster u. a.: Waxmann.
- Dederling, K. & Müller, S. (2008b). Schulinspektion als Gegenstand wissenschaftlicher Forschung. In S. Müller, K. Dederling & W. Bos (Hrsg.), *Schulische Qualitätsanalyse in Nordrhein-Westfalen* (S. 337–347). Köln: LinkLuchterhand.
- Dederling, K. & Müller, S. (2011). Wirkungen von externer Schulevaluation. *Journal Schulentwicklung*, 15 (1), 17–23.
- Gieske-Roland, M., Buhren, C.G. & Rolff, H.-G. (2014). *Peer Review an Schulen*. Weinheim und Basel: Beltz.
- Husfeldt, V. (2011). Wirkungen und Wirksamkeit der externen Schulevaluation. *Zeitschrift für Erziehungswissenschaft*, 14 (2), 259–282.
- Institut für Bildungsmonitoring und Qualitätsentwicklung Hamburg (2013). *Jahresbericht der Schulinspektion. Schuljahr 2010/2011*. Hamburg: Institut für Bildungsmonitoring und Qualitätsentwicklung.
- Kuhn, H. J. (2015). Schulinspektion in Deutschland – Erfolgsmodell oder neue Illusion? In C. G. Buhren (Hrsg.), *Handbuch Feedback in der Schule* (S. 79–107). Weinheim: Beltz.
- Maritzen, N. (2006). Schulinspektion in Deutschland. In H. Buchen, L. Horster & H.-G. Rolff (Hrsg.), *Schulinspektion und Schulleitung* (S. 7–26). Stuttgart: Raabe.
- Müller, S. (2010). Erste Effekte von Schulinspektion – eine Zwischenbilanz. In N. Berkemeyer, W. Bos, H. G. Holtappels, N. McElvany & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung*, Bd. 16 (S. 289–308). Weinheim und München: Juventa.
- Müller, S., Pietsch, M. & Bos, W. (2011). *Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht*. Münster u. a.: Waxmann.
- Pietsch, M., Janke, N. & Mohr, I. (2014). Führt Schulinspektion zu besseren Schülerleistungen? *Zeitschrift für Pädagogik* 60 (3), 446–470.
- Preuß, B. (2013). Akteurkonstellation zwischen Schulträger und Schule. Empirische Analysen zur Governance von Schulinspektion. In I. van Ackeren, M. Heinrich, F. Thiel (Hrsg.), *Evidenzbasierte Steuerung im Bildungssystem?* (Die Deutsche Schule: Beiheft 12, S. 154–171). Münster: Waxmann.
- Quesel, C., Husfeldt, V., Landwehr, N. & Steiner, P. (Hrsg.) (2011). *Wirkungen und Wirksamkeit der externen Schulevaluation*. Bern: hep-Verlag.
- Sommer, N. (2011). Wie beurteilen schulische Gruppen die erlebte Schulinspektion? In S. Müller, M. Pietsch & W. Bos (Hrsg.), *Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht* (S. 137–164). Münster u. a.: Waxmann.
- Wurster, S. & Gärtner, H. (2013). Schulen im Umgang mit Schulinspektion und deren Ergebnissen. *Zeitschrift für Pädagogik*, 59 (3), 425–445.

Rechtliche, strukturelle und politische Rahmenbedingungen für externe Evaluation in Deutschland

Hans-Jürgen Kuhn

Zusammenfassung

Vor mehr als zehn Jahren haben fast alle Länder damit begonnen, ein System einer standardisierten und systematischen externen Evaluation ihrer Schulen aufzubauen. Orientiert an landesspezifischen Orientierungsrahmen für Schulqualität werden damit sehr unterschiedliche Ziele verfolgt. Die Umsetzung dieser Konzeption ist in den Ländern sehr stark von den schulrechtlichen, strukturellen und bildungspolitischen Rahmenbedingungen beeinflusst. Trotz großer Gemeinsamkeiten im Vorgehen, in den eingesetzten Methoden und Instrumenten unterscheiden sich die Länder inzwischen zum Teil erheblich, wenn man die Rahmenbedingungen betrachtet, in die die Strategie der externen Evaluation eingebettet ist. Der Beitrag analysiert diese Rahmenbedingungen und zeigt an Beispielen auch die Konsequenzen, die getroffene Grundsatzentscheidungen für Wirkungen und Wirksamkeit dieses neuen Instruments haben. Dabei wird deutlich, dass vor allem die personelle Ausstattung, die fachlichen Qualifikationen des eingesetzten Personals und das Verhältnis der Evaluatoren zur operativen Schulaufsicht eine Schlüsselrolle bei der Weiterentwicklung des Gesamtsystems einnehmen. Es wird auch deutlich, dass die Kultusministerkonferenz im Rahmen ihrer Überlegungen zur Weiterentwicklung der Gesamtstrategie des Bildungsmonitorings vor der Frage steht, wie sie dieses relativ neue Instrument der Schulevaluation in eine produktive Beziehung zu den bundesweit festgeschriebenen Verfahren der standardisierten Leistungsmessungen bringt.

Die im Jahr 2015 existierenden Verfahren externer Evaluation von Schulen in den deutschen Ländern sind das Ergebnis eines seit zehn Jahren andauernden Entwicklungs- und Gestaltungsprozesses. Ausgelöst durch politische Diskussionen über die Qualität von Schule und Unterricht, die im Nachgang zur ersten PISA-Studie 2000 (vor allem nach der Veröffentlichung des PISA Ländervergleichs im Juni 2002) eine deutliche Dynamik entwickelten, haben die politisch Verantwortlichen für Bildungspolitik in den Ländern mit dem Aufbau einer systematischen, standardisierten externen Evaluation begonnen. Vor dem Hintergrund der jeweiligen landesspezifischen Strategie zur Qualitätssicherung und -entwicklung wurden bildungspolitisch durchaus unterschiedlich akzentuierte Rahmenbedingungen festgelegt, in denen sich die Akteure entwickeln konnten. Dies betraf wesentlich die schulrechtliche

Einordnung und Absicherung dieses Instruments der neuen Steuerung, die institutionelle Abgrenzung zu anderen bereits tätigen Akteuren (z. B. der Schulaufsicht) und die personelle Ausstattung. Vor allem auf diesem Gebiet wurden in den letzten zehn Jahren sehr unterschiedliche Rahmenbedingungen gesetzt, die nicht ohne Konsequenzen für die Güte, Akzeptanz und Entwicklungsperspektiven der externen Evaluation geblieben sind.

Während die einmal geschaffenen strukturellen und rechtlichen Rahmenbedingungen ein gewisses Beharrungsvermögen besitzen und über die Jahre eine Eigendynamik entfalten, können die alle vier bis fünf Jahre wechselnden politischen Konstellationen nach Landtagswahlen Veränderungen bewirken, die die Existenzberechtigung des neuen Instruments, seine Zielsetzung, Auftragslage und personelle Ausstattung erheblich erschüttern und verändern. Im Folgenden wird daher nach einer Darstellung der gegenwärtigen Vielfalt der Ausgestaltung der externen Evaluation der Frage nach den absehbaren Veränderungsprozessen nachgegangen. Dabei wird auch die Rolle der KMK behandelt, die sich, anders als auf dem Gebiet des Bildungsmonitorings durch empirisch fundierte Leistungsvergleiche auf unterschiedlichen Ebenen, bisher eher randständig mit dem Instrument der externen Evaluation beschäftigt hat.

1. Definition, Zielsetzung und erreichter Sachstand externer Evaluation von Schulen in Deutschland

Es gibt vielfältige Versuche, die mit der externen Evaluation verfolgten Ziele zu beschreiben und zu bündeln (vgl. z. B. Brägger et al. 2005; Burkard 2005; Maritzen 2007; Rürup 2008). Dabei gelingt es vor allem Landwehr (2011), die verschiedenen Hauptfunktionen systematisch voneinander abzugrenzen:

- *Katalysefunktion:* Schulevaluation erhebt den Anspruch, die Schulentwicklung zu befördern, indem sie durch die Rückmeldung eines umfassenden, datengestützten Qualitätsprofils Handlungsfelder identifiziert und nächste Entwicklungsschritte aufzeigt. Durch den Evaluationsbericht kann und soll der inner-schulische Diskussions- und Entwicklungsprozess stimuliert werden.
- *Rechenschaftsfunktion:* Schulevaluation leistet einen Beitrag zur staatlichen Gewährleistung, indem schulische Qualität gegenüber den beteiligten Akteuren, der Schulöffentlichkeit und den verantwortlich Steuernden sichtbar gemacht und so (Mindest-)Standards gesichert werden.
- *Normendurchsetzungsfunktion:* Die in den Referenz- oder im Orientierungsrahmen kodierten normativen Erwartungen an schulische Qualität werden in den Schulen vor allem mit Blick auf eine anstehende Evaluation intensiv verarbeitet und so aktiv aufgenommen.

- *Erkenntnisfunktion*: Schulevaluation leistet einen Beitrag zum Bildungsmonitoring, indem sie die einzelschulischen Befunde zu Aussagen über die Qualität des Gesamtsystems verdichtet und Steuerungserfordernisse offenlegt. Diese können dann Grundlage für Entscheidungen zur Weiterentwicklung des Schulsystems sein.

Externe Evaluationen zielen mit den drei ersten Zielen vorrangig auf die Ebene der Einzelschule, die vierte Zielsetzung bedient den Informationsbedarf auf der Ebene des Gesamtsystems. Ausgehend von den allgemeinen Ansätzen zur Schulentwicklung soll die externe Evaluation einen ergänzenden Beitrag dazu leisten, die Handlungskoordination der Akteure zu stimulieren und zu verbessern; das Wissen über die eigene Organisation soll erweitert werden um Ebenen, die der Schule selbst kaum zugänglich sind (blinde Flecken).

Die Umsetzung dieser Ziele ist in unterschiedlichen institutionellen Formen möglich:

- Im *Netzwerkmodell* kooperieren fallbezogen Evaluatoren, die im Hauptamt andere Funktionen im System haben (z. B. Schulleitungen, Schulaufsicht, Schulberater). Dieses Modell wurde in Schleswig-Holstein bis 2009 praktiziert und dann eingestellt.
- Im *Expertenmodell* kommt es ebenfalls zu einer fallbezogenen Kooperation von systemexternen Evaluatoren auf Werkvertragsbasis. Teilweise wird dies auch mit einer Auftragsvergabe des gesamten Prozesses an ein wissenschaftliches Institut kombiniert. Dieser Weg wurde bisher lediglich in Bremen erprobt und wird im zweiten Durchgang nicht mehr praktiziert.
- Beim *Inspektorenmodell* werden die externen Evaluationen durch einen festen Stamm von meist hauptamtlichen Personen in einer eigenen staatlichen Einrichtung oder Teilorganisation durchgeführt. Die Evaluatoren sind Landesbeschäftigte mit speziellen Kenntnissen, die sie speziell für diese Aufgabe zusätzlich erworben haben.

Das Inspektorenmodell bildet in fast allen Ländern bisher die Grundform der externen Evaluation und lässt sich auch klar abgrenzen zu anderen Formen der externen Evaluation, wie sie mit den Begriffen *Peer Review* oder *Critical Friends* beschrieben werden (vgl. Buhren & Düring 2008).

Der Begriff des „kritischen Freundes“ spiegelt nach Gieske-Roland (2015, S. 109) die Balance zwischen partnerschaftlicher und gleichzeitig im positiven Sinne kritischer Haltung wider, die die Peers gegenüber der fallgebenden Schule vertreten sollen, und impliziert, dass das Verhältnis zwischen Peers und besuchter Schule auf Freiwilligkeit, nicht auf einem formalen Kontrollverhältnis beruht. Die Peers sollen in der Regel aus einem ähnlichen Umfeld stammen, z. B. Lehrkräfte der

gleichen Schulform oder von benachbarten Schulen. Es wird keine Rechenschaft nach außen abgelegt, sondern die Ergebnisse sollen im Sinne des „Impulsegebens“ innerschulische Prozesse und Entwicklungen unterstützen.

Dagegen ist externe Evaluation im oben definierten Verständnis nach Stryck (2014) ein unbequemer Akteur

„der eigenen Professionalität wegen – distanziert und analytisch, sie versucht die Daten und Informationen, die sie sammelt und herstellt, nach einem eigenen Indikatorensystem zu ordnen. Lehrende Angehörige der Schule sind dagegen in ihre eigene Geschichte eingebunden (also nicht distanziert), sind ein Teil des Geschehens, greifen insofern auch immer bevorzugt auf (individuelle) Erfahrungen (und eben nicht auf Daten) zurück, wenn sie ihren Unterricht reflektieren sollen – sie konstituieren systematisch eine andere Wirklichkeit des Unterrichtsgeschehens.“ (ebd., S. 8)

Ab 2004 haben die Länder begonnen, flächendeckend die externe Evaluation einzuführen. Vierzehn Länder haben inzwischen einen ersten Besuch in fast allen öffentlichen Schulen abgeschlossen und ausgewertet. Dabei sind bis Ende 2014 ca. 21.000 Schulen zum ersten Mal extern evaluiert worden¹ und durch einen umfassenden schriftlichen Bericht über die wahrgenommenen Stärken und Schwächen in ausgewählten Dimensionen von Schul- und Unterrichtsqualität informiert worden. Im Rahmen einer zweiten Schulevaluation sind bis Ende November 2014 ca. 4.500 Schulen erneut evaluiert worden. Vorzeitig abgebrochen wurde das Verfahren 2009 in Schleswig-Holstein (EVIT) und NRW hat aufgrund von jahrelangen personellen Engpässen erst ca. 65 % seiner Schulen ein erstes Mal evaluiert. Schulen in freier Trägerschaft werden bisher durch die Schulevaluatoren der Länder nur in Ausnahmefällen auf eigenen Wunsch besucht. Insbesondere die evangelische und katholische Kirche haben für die von ihnen getragenen Schulen in den letzten Jahren eigene Qualitätssysteme mit externen Schulbesuchen entwickelt.

Noch nie wurde so viel Unterricht besucht und analysiert wie im Rahmen der externen Evaluationen. So haben die Berliner Schulinspektorinnen und Schulinspektoren im Rahmen der Besuche von ca. 700 Schulen im ersten Durchgang den Unterricht von über 30.000 Stunden bewertet. Eine Quelle von Erkenntnissen, wie sie mit dem vorhandenen schulaufsichtlichen Personal nie hätte erlangt werden können.

In zwölf Ländern findet inzwischen der zweite Evaluationsdurchlauf statt, mindestens vier Länder (Berlin, Brandenburg, Hessen, Rheinland-Pfalz) entwickeln bereits die konzeptionellen Eckpunkte für einen dritten Zyklus ab Sommer 2016.

Fast alle Länder haben im ersten Durchgang nach dem Modell einer *standardisierten Vollinspektion* gearbeitet. Dabei werden weitgehend schulformunabhängig für alle Schulen die gleichen Qualitätsmerkmale untersucht. Im zweiten Durchgang

1 Die Zahl ergibt sich aus den in 15 Ländern durch den Autor abgefragten Fallzahlen.

haben einige Länder dieses Vorgehen aufgrund der Erfahrungen fachlich weiterentwickelt, das Grundmodell aber (auch zur Sicherung der Anschlussfähigkeit beider Durchgänge) als *modifizierte Vollinspektion* beibehalten. In einigen Ländern ist man auf der Basis des Feedbacks der Schulen dazu übergegangen, eine sogenannte Fokus-Evaluation durchzuführen, bei der den Schulen auch die Möglichkeit eingeräumt wird, aus einem Angebot unterschiedlicher Module eigene Themen auszuwählen.

2. Gemeinsamkeiten und Differenzen der Länder zu ausgewählten Rahmenbedingungen

2.1 Schulrechtliche Einordnung in den Ländern

Eine Analyse der Schulgesetze der Länder (Stand November 2014) zeigt, dass in den letzten Jahren in allen Ländern (bis auf Schleswig-Holstein) neue Paragraphen oder Ergänzungen zu bestehenden Paragraphen in die Schulgesetze aufgenommen wurden, die den Bereich der Qualitätsentwicklung und -sicherung auch gesetzlich normieren. Dabei ist in fast allen Ländern im Schulgesetz eine Verpflichtung der Schulen zur Durchführung von Maßnahmen der internen Evaluation (auch Selbstevaluation genannt) sowie die Teilnahme an externen Evaluationen/Fremdevaluationen festgelegt.

Der Grad der Festlegungen bezüglich der Verpflichtung zur internen Evaluation ist dabei unterschiedlich konkret definiert. Schleswig-Holstein und Berlin markieren dabei die Pole der Festlegungen. So ist im Schleswig-Holsteinischen Schulgesetz vom 24. Januar 2007 im §3, Abs. 1 *Selbstverwaltung der Schule* lediglich die Überprüfung eines selbstverfassten Schulprogramms festgelegt:

(1) Die Schulen sind im Rahmen der geltenden Rechts- und Verwaltungsvorschriften selbständig in der Durchführung des Auftrages der Schule und in der Verwaltung ihrer eigenen Angelegenheiten. Die einzelne Schule gibt sich zur Ausgestaltung ihrer pädagogischen Arbeit und des Schullebens ein Schulprogramm, das sie der Schulaufsichtsbehörde vorlegt. Vor der Beschlussfassung ist der Schulträger zu hören. Das Schulprogramm ist von der Schulkonferenz in regelmäßigen Abständen zu überprüfen.

In Schulgesetz von Berlin vom 26. Januar 2004 heißt es dagegen im §9, Abs. 2 *Qualitätssicherung und Evaluation* sehr konkret:

(2) 1. Die interne Evaluation obliegt der einzelnen Schule und wird von Personen vorgenommen, die der Schule angehören. 2. Bei der Konzeption, Durchführung und Auswertung kann sich die Schule Dritter bedienen. 3. Für die Bereiche und Gegenstände der

internen Evaluation sind von der Schule Evaluationskriterien und Qualitätsmerkmale zu entwickeln und anzuwenden. 4. Die Schulkonferenz beschließt auf Vorschlag der Gesamtkonferenz der Lehrkräfte ein Evaluationsprogramm für die Schule. 5. Die Verantwortung für die interne Evaluation hat die Schulleiterin oder der Schulleiter. 6. Die Schule legt der Schulkonferenz und der Schulaufsichtsbehörde einen schriftlichen Evaluationsbericht vor.

Ähnlich differenziert wie bei der schulgesetzlichen Absicherung der internen Evaluation verhält es sich bei der externen Evaluation. Betrachtet man die Vielfalt der in den Schulgesetzen der Länder gewählten Bezeichnungen für die externe Evaluation, fällt auf, dass nur in vier Ländern (Hamburg, Niedersachsen, Hessen und Sachsen-Anhalt) der Begriff „Schulinspektion“ benutzt wird. In acht Ländern wird das Verfahren als „Externe Evaluation“ bzw. in Baden-Württemberg als „Fremdevaluation“ bezeichnet. Brandenburg hat sich für den Begriff „Schulvisitation“ entschieden, um eine klare Abgrenzung zum in der DDR bestehenden flächendeckenden System einer fachlichen und politischen Kontrolle herzustellen, die durch ca. 800 hauptberufliche Bezirks- und Kreisschulinspektoren ausgeübt wurde (vgl. Geißler 2008). Auch die anderen ostdeutschen Länder haben durch die Verwendung des Begriffs „Externe Evaluation“ eine abgrenzende Bezeichnung zum Inspektionsbegriff gewählt. In Nordrhein-Westfalen wird der Begriff „Qualitätsanalyse“ verwendet und das Modell „Evaluation im Team/EVIT“ aus Schleswig-Holstein ist 2009 ersatzlos eingestellt worden.

Mit den gesetzlichen Festlegungen hat die externe Evaluation als neues Instrument der Qualitätssicherung in 15 Ländern eine gewisse Stabilität erhalten, da die gesetzliche Verankerung durch Parlamente im Rahmen der Schulgesetzgebung auch als Signal für Dauerhaftigkeit interpretiert werden kann. Dies schließt allerdings, wie am Beispiel Schleswig-Holsteins sichtbar wurde, nicht aus, dass es nach einer Landtagswahl auch zur Abschaffung der externen Evaluation kommen kann, wenn neue Regierungsparteien die Notwendigkeit dieses Instrument grundsätzlich infrage stellen.

Die Analyse der einschlägigen Paragraphen der Schulgesetze, in denen die externe Evaluation beschrieben wird, zeigt eine erhebliche Varianz zwischen den Ländern bezüglich der Regelungsbereiche im Gesetz. Während in einzelnen Ländern die Ziele der externen Evaluation genauer beschrieben werden – z. B.: „Die Qualität schulischer Arbeit zu sichern und zu verbessern“ (Bayern); „Die Schulvisitation unterstützt die Qualitätsentwicklung der Schule durch regelmäßige systematische Schulbesuche“ (Brandenburg); „Die externe Evaluation dient dazu, die Standards, die für die Schulen gelten, zu sichern, die Entwicklung und Fortschreibung der Schulprogramme zu unterstützen, Erkenntnisse über den Stand und die Qualität von Unterricht und Erziehung, Schulorganisation und Schulleben zu liefern sowie die Gleichwertigkeit, Durchgängigkeit und Durchlässigkeit des schulischen Bil-

dungsangebots zu gewährleisten“ (Berlin) –, wird in Gesetzen anderer Länder eher die Unabhängigkeit der eingesetzten Evaluatoren betont – z. B.: „Schulinspektoren und Schulinspektoren sind in der Bewertung der Qualität einzelner Schulen an Weisungen nicht gebunden“ (Hamburg) oder: „Sie sind hinsichtlich ihrer Feststellungen bei der Durchführung der Qualitätsanalyse und deren Beurteilung an Weisungen nicht gebunden“ (Nordrhein-Westfalen).

Weitere gesetzliche Normierungen, die einzelne Länder vorgenommen haben, beziehen sich auf die Festlegung, dass Schulen und Schulaufsicht auf der Grundlage der Ergebnisse der externen Evaluation Zielvereinbarungen treffen (z. B. Bayern, Thüringen), regeln die Möglichkeit des Systemmonitorings mit Evaluationsdaten und verpflichten sich zu Bildungsberichten, in denen auch die Ergebnisse der externen Evaluation landesweit ausgewertet werden, oder legen fest, wer Empfänger der Ergebnisberichte der Einzelschule sein darf.

Wo im Gesetz eher knappe Ausführungen zum Gegenstand gemacht werden, gibt es oft die Ermächtigung für Rechtsverordnungen oder ergänzende Verwaltungsvorschriften, in denen weitere Präzisierungen vorgenommen werden. Es bleibt offen, ob das, was gesetzlich geregelt sein muss oder untergesetzlich entschieden werden darf, dauerhaft in der aktuellen Beliebigkeit erhalten bleibt oder ob die einsetzende Tendenz, Ergebnisberichte der Einzelschulen zu veröffentlichen, zu Konflikten führt, die eine präzisere gesetzliche Normierung erfordern.

Insgesamt lassen viele Schulgesetze den Eindruck entstehen, dass hinter der in den letzten zehn Jahren erfolgten additiven Aufnahme neuer Instrumente der Qualitätsentwicklung, wie die Verpflichtung der Schulen zum Erstellen von Schulprogrammen, Verpflichtungen zur Selbst- und Fremdevaluation und die oft gleichzeitig angestrebte erweiterte Selbstständigkeit der Einzelschule, kein stimmiges Gesamtkonzept steht. Insofern ist Avenarius (2006) zuzustimmen, wenn er das ergänzte Niedersächsische Schulgesetz vom Sommer 2006 kritisch kommentiert und feststellt:

„Wie die Niedersächsische Schulinspektion ihrerseits externe Evaluationen handhabt, nach welchen Kriterien sie die Qualität pädagogischer Arbeit letztlich bewertet: Das geht aus dem Schulgesetz nicht hervor. Das Gesetz lässt der Schulinspektion in ihren Aktivitäten freien Lauf. Sie kann die Interna der Schulen nach einem ihnen aufgezwungenen Raster durchleuchten und bewerten. [...] Das Niedersächsische Schulgesetz in der durch das Gesetz zur Einführung der Eigenverantwortlichen Schule geänderten Fassung enthält kaum vereinbare Ziele: Es strebt die Stärkung der Eigenständigkeit der Schule an, unterwirft die Schule aber zugleich einer ins Belieben der Schulinspektion gestellten unbegrenzten Qualitätskontrolle. Es soll die schulische Qualität verbessern, lässt aber Kriterien vermissen, an denen sich Schulen, Schulaufsicht und Schulinspektion im Bemühen um Qualitätsentwicklung und Qualitätssicherung orientieren könnten.“ (ebd., o. S.)

Es wäre ein Gewinn für die weitere Diskussion, wenn es den Schulrechtsexperten der Länder gelänge, einen Konsens über die Sachverhalte zu erreichen, die entsprechend der Wesentlichkeitstheorie einer gesetzlichen Regelung bedürfen.

2.2 Umgang mit den Ergebnissen

2.2.1 *Wer erhält die Schulberichte?*

Obwohl bereits durch die Ankündigung und Vorbereitung einer externen Evaluation in Schulen sinnvolle Aktivitäten angestoßen werden, soll das Abschlussprodukt, der Schulbericht, den entscheidenden Entwicklungsimpuls auslösen. Inhalt, Umfang, Struktur und Detaillierungsgrad sind daher in allen Ländern immer wieder Gegenstand von Überarbeitung und Weiterentwicklung.

Ebenso relevant ist allerdings auch, wer den Bericht und damit die Chance erhält, gestützt auf den Inhalt und die Befunde Konsequenzen abzuleiten oder einzufordern. Zwar ist die Schule immer der wichtigste Adressat des Berichts, allerdings ist damit noch nicht geklärt, ob nur die Schulleitung den vollständigen Bericht erhält oder ob auch Eltern, Lehrkräfte und Schüler (alle oder nur die gewählten Gremienvertreter?) Gelegenheit erhalten, den gesamten Bericht einzusehen.

Externe Evaluationen zielen als vom Land extern gesetzter Steuerungsimpuls auf die einzelne Schule als Handlungseinheit. In der Regel erhält aber auch die für die jeweilige Schule zuständige Schulaufsichtsbehörde den Bericht direkt oder kann ihn von der Schule anfordern. Schließlich soll sie dafür Sorge tragen, dass durch eine intensive Auswertung des Berichts oder durch den Abschluss von Zielvereinbarungen mit der Schulaufsicht Konsequenzen erfolgen und mögliche Schwächen der Schule bearbeitet werden.

Angesichts der verfassungsrechtlich festgelegten Aufgabenteilung, wonach in Deutschland in der Regel der kommunale Schulträger für die Errichtung, Unterhaltung und Verwaltung der Schulen verantwortlich ist und die Sachkosten trägt und das Land die Personalkosten für Lehrkräfte an öffentlichen Schulen, ist nicht zwingend davon auszugehen, dass der Schulträger an der externen Evaluation beteiligt wird oder am Ende auch den vollständigen Bericht erhält. Dies ist tatsächlich auch eher die Ausnahme. Häufig wird er lediglich im Rahmen von Vorabbefragungen oder bei Schulrundgängen einbezogen und erhält nur Auszüge des Berichts, die sich mit den räumlichen Gegebenheiten oder anderen Fragen beschäftigen, bei denen der Schulträger eine unmittelbare Zuständigkeit hat. Angesichts der Tatsache, dass Schulen Teil der örtlichen Gemeinschaft sind und als weicher Standortfaktor von vielen Schulträgern geschätzt und unterstützt werden, läge es nahe, auch dem Schulträger die Ergebnisse der externen Evaluation „seiner“ Schulen zur Verfügung zu stellen. Dies bedarf jedoch einer konkreten normativen Regelung

im Schulgesetz oder einer Rechtsverordnung bei entsprechender Ermächtigung. Tatsächlich haben etliche Länder die vollständige Weitergabe der Schulberichte an die Schulträger schulgesetzlich ausgeschlossen.

Obwohl inzwischen in vielen Ländern Versuche unternommen werden, im Rahmen von „Bildungslandschaften“ die verschiedenen Institutionen in vertikaler und horizontaler Ebene zusammenzubringen und zu einem gemeinsam koordinierten Vorgehen anzuhalten, wird im Umgang mit Evaluationsberichten von Schulen innerhalb der zuständigen Bildungsministerien noch häufig von einer starren Trennung in innere und äußere Schulangelegenheiten ausgegangen und das Thema Schulqualität klar zur Seite der inneren Schulangelegenheiten gerechnet.

Dies entspricht jedoch keinesfalls mehr den Vorstellungen der kommunalen Schulträger, wie zuletzt in einer länderübergreifenden empirischen Studie im Rahmen des BMBF-Projekts „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene“ in 52 qualitativen Experteninterviews erhoben wurde (vgl. Brüsemeister et al. 2009). Als ein Ergebnis dieser Studie kann nach Preuß (2013) festgestellt werden, „dass der Schulträger im Kontext von Schulinspektion gerne ein echter ‚Mitspieler‘ wäre und für ihn die Frage nach dem Steuerungseinfluss eine Grundsatzfrage ist. Faktisch sieht er sich jedoch nicht angeschlossen und in einer Rolle als Außenbeobachter am Rande des Geschehens.“ (ebd., S. 167)

In ähnlicher Weise hat der Deutsche Städtetag dies im Rahmen seiner „Münchener Erklärung“ formuliert: „Die Länder sollten dem kommunalen Engagement durch erweiterte schulgesetzlich geregelte Kompetenzzuweisungen Rechnung tragen. Hierzu gehört insbesondere die Mitgestaltung bei der inneren Schulentwicklung, [...]“ (Deutscher Städtetag 2012, S. 3).

Für eine künftig bessere Nutzung der Potenziale der externen Evaluation wäre es sehr sinnvoll, den Schulträger in seinem „Steuerungseinfluss“ aufzuwerten, indem er grundsätzlich den vollständigen Evaluationsbericht erhält. Zu zahlreich sind inzwischen die Anforderungen an Schulen und die daraus resultierenden Verflechtungen mit anderen kommunalen Institutionen und Partnern, die u. a. durch den forcierten Ausbau von Ganztagsangeboten ständig weiter zunehmen. Unterschiedliche Zuständigkeiten für das ergänzende pädagogische und nicht-pädagogische Personal, Kooperationen mit Jugendhilfeträgern und erweiterte Eigenverantwortlichkeit der Schulen machen es erforderlich, bei der Ausgestaltung der regionalen Bildungs- und Verantwortungsgemeinschaften die Ergebnisse der externen Evaluation breiter zu kommunizieren und so die Chancen für evaluationsgestützte Schulentwicklung unter Einbeziehung der Schulträger zu erhöhen.

2.2.2 *Entwicklungsappell oder Sanktionsstrategie?*

Vergleicht man die vorliegenden Erfahrungen im deutschsprachigen Raum zum Umgang mit externen Evaluationsergebnissen bzw. zur Akzeptanz der Verfahren mit denen aus angelsächsischen Studien (vgl. Böhm-Kasper & Selders 2013) fällt auf, dass die Akzeptanz der Verfahren unabhängig von landesspezifischen Ausprägungen in Deutschland insgesamt relativ hoch ausfällt, während in angelsächsischen Studien eher die Belastungen und der Kontrollaspekt thematisiert werden. Tatsächlich versteht sich die externe Evaluation in Deutschland vor allem als Dienstleistung gegenüber den Schulen, um mit transparenten und standardisierten Methoden Informationen zu Stärken und Schwächen zu liefern. Sie setzt damit auf Schulentwicklung durch Einsicht als Handlungsoption und vermeidet alles, was als negative Sanktion verstanden werden könnte. Dem dienen auch Regelungen, wonach die Evaluatoren ausdrücklich im inneren Dienstbetrieb einer Verschwiegenheitspflicht unterliegen. Auch wird oft sichergestellt, dass die Mitglieder der externen Evaluation nicht gleichzeitig oder kurz zuvor noch schulaufsichtliche Aufgaben gegenüber den untersuchten Schulen wahrnehmen. Lediglich die in einigen Ländern praktizierte Form von „Nachinspektionen“ bei Schulen, die definierte Standards nicht erreicht haben und als Schulen mit erheblichem Entwicklungsbedarf gelten, kann in Ansätzen als Sanktion verstanden werden. Schließlich müssen in diesen Fällen Schule und Schulaufsicht verbindliche Zielvereinbarungen abschließen, in denen konkrete Entwicklungsarbeiten verabredet werden. Aber auch in Ländern, die diese Vorgehensweise praktizieren, besteht eine grundsätzlich positive Wahrnehmung der externen Evaluation.

Demgegenüber setzen angelsächsische Inspektionssysteme stärker auf Transparenz durch die Veröffentlichung von Inspektions- und Leistungsdaten und verstehen ihren Auftrag stärker kontrollierend. In der Folge schlechter Schulergebnisse kann dies zu massiven Eingriffen in die Autonomie der Schulträger und Schulleitung führen, die vom Austausch des Leitungspersonals bis zur Schließung der Schule reichen kann. Während deutsche Schulen nach schwachen externen Evaluationsergebnissen (z. T. vergeblich) auf Unterstützung hoffen, besteht in anderen Ländern eher die Sorge vor Sanktionen.

Während für den ersten Durchgang einer Inspektion in allen Ländern Konsens darüber bestand, dass der Schulbericht nicht öffentlich publiziert wird, werden im zweiten Durchgang in Berlin, Brandenburg und Hamburg die Schulberichte ganz oder nur geringfügig gekürzt auf öffentlich zugänglichen Internetplattformen („Schulporträt“) nach einer Frist von zwei bis sechs Monaten veröffentlicht. Rankinglisten werden allerdings weiterhin ausgeschlossen und könnten nur durch aufwendige Analysen der veröffentlichten Evaluationsberichte ggf. durch Medienvertreter selbst erzeugt werden.

Im Vorfeld der bildungspolitischen Entscheidung einzelner Länder, die Veröffentlichung der Berichte vorzunehmen, hat es vonseiten der Schüler- und Elternvertreter verstärkten Druck gegeben, durch eine Veröffentlichung vollständige Transparenz für alle an Schule Interessierten herzustellen. Als Argument wurde dabei auch die Hoffnung geäußert, nur so würden Schulleitungen vor allem bei kritischen Schulrückmeldungen gezwungen sein, auch zu reagieren und Probleme tatsächlich zu bearbeiten. Das Argument, der Erfolg der externen Evaluation hänge davon ab, dass die Schulen darauf vertrauen können, dass sich aus dem Untersuchungsergebnis für sie keine Nachteile ergeben, wird von Elternvertretern durchaus kritisch gesehen. Vor allem aufseiten der Schulen hat es dagegen Befürchtungen gegeben, mit der Öffentlichkeit der Ergebnisse würde das Instrument seine Funktion erheblich verändern und es wäre mit unerwünschten Nebenfolgen zu rechnen (z. B. „*Window Dressing*“).

Anders als befürchtet, gibt es bisher keinerlei Anzeichen, dass die Veröffentlichung mit unerwünschten Nebenfolgen verbunden ist. Ob das tatsächlich so ist und auch dauerhaft so bleibt, ist bisher wissenschaftlich nicht untersucht und methodisch auch nur schwer zu erfassen. Das mediale Interesse blieb letztlich gering, aufgrund der Ungleichzeitigkeit des Erscheinens der Berichte wurden bisher auch keine Rankings erstellt und veröffentlicht. Ob die interessierte Öffentlichkeit die nun gegebene Möglichkeit, sich umfassender über wesentliche Qualitätsparameter einer Schule zu informieren, auch tatsächlich nutzt, bleibt allerdings ebenfalls mangels geeigneter Erhebungsmöglichkeiten unklar. Da aufgrund des zunehmenden Drucks auf die Bildungsadministration durch gesetzliche Vorgaben (Informationsfreiheitsgesetzgebung) ohnehin immer mehr einzelschulbezogene Daten veröffentlicht werden müssen, werden weitere Länder diesem Beispiel vermutlich folgen. Ein Schritt in diesem Sinne muss allerdings gut kommuniziert werden und die Bedeutung der dahinterstehenden Transparenzerwartungen der Öffentlichkeit positiv aufnehmen und sich damit gegen ein fehlgeleitetes Marktverständnis von Schulqualität abgrenzen.

2.2.3 Externe Evaluation und Datenschutz

Das in den letzten Jahren stark gewachsene Interesse der Öffentlichkeit an Informationen zur Qualität von Einzelschulen ist nicht nur vor dem Hintergrund der Eltern verständlich, die bei der Wahl der Schule im Rahmen der Einschulung oder der Übergangsentscheidungen nach Klasse 4 bzw. 6 Informationen suchen, um für ihr Kind die „richtige“ Schule zu finden. Dieses Informationsinteresse muss allerdings in Einklang gebracht werden mit dem Recht der betroffenen Schülerinnen und Schüler, Lehrkräfte, Schulleiterinnen und Schulleiter auf informationelle Selbstbestimmung. Ergebnisse von Jahrgangstests (z. B. VERA 3 oder 8), Abschlussprü-

fungen am Ende der Jahrgangsstufe 10, aber auch Berichte der externen Evaluation lassen nicht nur Rück- und Fehlschlüsse auf den Unterricht einer Schule zu, sondern u. U. auch auf die Arbeit einzelner Lehrkräfte oder die Arbeit des Schulleiters oder der Schulleiterin.

Damit ergeben sich etliche datenschutzrelevante Fragestellungen:

- Ist die Erhebung von Werturteilen über Lehrkräfte und Schulleiterinnen und Schulleiter überhaupt rechtmäßig?
- Kann ausgeschlossen werden, dass die zur Evaluation erhobenen Daten nicht für eine dienstliche Beurteilung zweckentfremdet werden?
- Wie ist die Verwendung der im Rahmen der Evaluation gewonnenen, auch personenbezogenen Daten geregelt?
- Haben schulische Gremien oder die gesamte Schulöffentlichkeit ein Recht darauf, die Berichte zu lesen?
- Dürfen externe Evaluationsberichte ganz oder teilweise veröffentlicht werden?

In den Schulgesetzen der Länder sind diese Fragen teilweise rechtlich geklärt, teilweise greifen für Einzelaspekte aber auch die sich in Deutschland ausbreitenden Informationsfreiheitsgesetze. Einige bereits vorliegende Entscheidungen von Verwaltungsgerichten (vgl. Dix 2011) stützen das Vorgehen der Länder, dass z. B. die Erörterung eines Schulinspektionsberichtes, in dem eine Schulleiterin kritisch bewertet wird, in der Schulkonferenz von dieser hingenommen werden müsse.

Für die Veröffentlichung der Berichte sind die Länder unterschiedliche Wege gegangen. In Bayern und Brandenburg ist geregelt, dass die personenbezogenen Informationen zur Schulleitung aus dem Inspektionsbericht vor der Veröffentlichung zu entfernen sind (Bayern) oder nur mit Zustimmung der Schulleiterin oder des Schulleiters veröffentlicht werden dürfen (Brandenburg). Allerdings ist Dix zuzustimmen, der bedauernd feststellt, dass dies aber zu einem Transparenzverlust in einem wesentlichen Bereich führt (vgl. Dix 2011, S. 39), und ergänzend anmerkt: „Wo Schulen Qualitätsprobleme haben, lassen diese sich nicht durch überzogene Geheimhaltung im Sinne eines ‚Schonraums Schule‘ lösen.“ ebd., S. 41)

2.3 Schulevaluator als Beruf oder Durchgangsstadium?

Eine externe Evaluation ist nur so gut

- wie die Daten, die ihr zweckgebunden zur Verfügung stehen,
- wie deren strukturierte, zeitgerechte und kriterienorientierte Aufbereitung,

- wie die Verknüpfung, die ihr zwischen qualitätsrelevanten Datenbeständen gelingt,
- wie die Standards empirischer Prüfung, die sie auf die eigenen Verfahren anwendet, und
- wie die Professionalität der Personen, die die Evaluation durchführen.

Bezüglich der Beurteilung der Professionalität der mit der Durchführung von externen Evaluationen beauftragten Personen ist es wichtig, eine Trennung zwischen Selbst- und Fremdwahrnehmung vorzunehmen. Während die Akzeptanz der Evaluatoren bei Schulleitungsmitgliedern und Lehrkräften der besuchten Schulen vor allem davon abhängt, ob diese über hinreichend lange Berufserfahrungen als Pädagogen verfügen (nach Möglichkeit auch in der Schulform der besuchten Schule) und welchen Status in Bezug auf Nähe zur Schulaufsicht bzw. Dienstbehörde sie besitzen, spielen für die Evaluatoren zur Einschätzung der eigenen Kompetenz noch andere Aspekte wie die Beherrschung der Anwendung methodischer Kompetenzen oder „Erfahrungen im Evaluationsgeschäft“ eine große Rolle. Diese wiederum sind zum Teil in Abhängigkeit von Rahmenbedingungen für den Arbeitsplatz „Schulevaluator/-in“ zu sehen, die wesentlich durch bildungspolitische Entscheidungen der Länder festgelegt werden.

Obwohl es seit fast zehn Jahren mehrere hundert Personen in Deutschland gibt, die sich überwiegend oder ausschließlich mit der externen Evaluation von Schulen beschäftigen, handelt es sich um eine Personengruppe, über die wenig systematisch erhobene Informationen vorliegen. Anders als für Lehrkräfte und Schulleitungsmitglieder sowie teilweise auch für Schulrätinnen und Schulräte, gibt es in fast keinem Land einheitliche, schriftlich fixierte und durch entsprechende Regelungen im Beamten-, Besoldungs- oder Laufbahnrecht kodifizierte Festlegungen darüber, wie man Schulevaluator/-in wird, unter welchen Rahmenbedingungen man diese Aufgabe bearbeitet, wie man dafür bezahlt wird und welche weiteren Verwendungsmöglichkeiten es beim Ausscheiden aus dieser Funktion gibt.

Mit dem Beginn der flächendeckenden Einführung der externen Evaluation ab 2005 war dies auch kaum zu leisten; mangels Erfahrungen behalf man sich mit unterschiedlichen, oft befristeten Funktionsbeschreibungen, für die sehr unterschiedliche Personen rekrutiert wurden. Nachdem nun aber in fast allen Ländern bereits der zweite Durchgang stattfindet und ein dritter Durchgang in mehreren Ländern bereits geplant wird, liegen zahlreiche Erfahrungen mit unterschiedlichen Modellen vor, die es möglich und notwendig machen, die Fragen nach der Qualität der Arbeit, dem Maß an angestrebter Professionalisierung systematisch und möglichst länderübergreifend weiterzuentwickeln.

Anders als beim Schulleitungspersonal hat sich bisher auch die Forschung wenig mit dieser neuen Akteursgruppe, ihrer Rolle, ihrer Qualifikation und ihrem Selbstkonzept beschäftigt. Dieses Schicksal teilt sie mit den Mitarbeiterinnen und

Mitarbeitern der operativen Schulaufsicht in den Ländern, die ebenfalls in den letzten 15 Jahren eher aus dem Blickfeld der forschenden Erziehungswissenschaften gefallen sind.

Eine der wenigen aktuell vorliegenden Arbeiten zur Untersuchung der Kompetenzen von Schulevaluatoren ist eine 2013 abgeschlossene Studie, die an der Universität Vechta durchgeführt wurde (vgl. Dederling & Sowada 2013) und im Rahmen von teilstandardisierten Leitfadeninterviews mit niedersächsischen Schulinspektoren der Frage nachgegangen ist, über welche Kompetenzen Schulinspektoren aus eigener Sicht verfügen müssen und inwiefern sich diese von Kompetenzen anderer schulischer Akteure unterscheiden. Ein Ergebnis dieser Untersuchung verweist auf die große Bedeutung, die die Schulinspektoren den Kompetenzen beimessen, die dem interpersonalen Bereich und dem Bereich der Selbstreflexion zugeordnet werden können. Außerdem werden kaum Unterschiede zwischen Inspektoren und anderen schulischen Akteuren (z. B. Schulleitungen) gesehen, was auf eine geringe Differenz der beschriebenen Kompetenzprofile hinweist. Erfahrungen als Schulleitung werden als wichtige Voraussetzung für die Arbeit als Schulinspektor betrachtet und umgekehrt.

In den Schulgesetzen der Länder werden selten Angaben zur Qualifikation des Evaluationspersonals gemacht. Lediglich im Schulgesetz Thüringen wird für die Expertenteams bei der externen Evaluation festgelegt: „Sie bestehen in der Regel aus dafür besonders geschulten Lehrkräften, insbesondere Schulleitungsmitgliedern und Mitarbeitern aus Schulämtern außerhalb des für die Schule zuständigen Schulamtsbereichs.“ (§ 40b, Abs. 4 ThürSchulG). Knappe Aussagen finden sich auch in Bayern: „Die Schulaufsichtsbehörden setzen Evaluationsgruppen ein, die speziell für diese Aufgabe qualifiziert werden.“ (Art. 113c, Abs. 2 BayEUG), sowie in Brandenburg: „[...] systematische Schulbesuche, die von fachlich geeigneten Personen durchgeführt werden.“ (§ 129, Abs. 3 BbgSchulG)

Bei der konkreten Rekrutierung des erforderlichen Personals mussten die Länder zahlreiche Strukturentscheidungen treffen, die für die Akzeptanz, Qualifikation und Perspektive der Evaluatoren nach innen und außen von erheblicher Bedeutung sind:

- Welche formalen oder tatsächlichen Qualifikationsanforderungen werden im Auswahlverfahren vorausgesetzt?
- Gibt es ein fachlich definiertes Ausbildungscurriculum vor der Aufnahme der Tätigkeit als Evaluator/-in?
- Wird die Aufgabe in Vollzeit oder in Teilzeit ausgeübt?
- Wird sie befristet oder dauerhaft übertragen?
- Ist das ausgewählte Personal status- und laufbahnrechtlich dem Schulaufsichtsdienst gleichgestellt oder wird in bewusster Abgrenzung eine Trennung vorgenommen?
- Wonach richtet sich die Bezahlung/Besoldung der beauftragten Personen?

Mit der Verstetigung des Systems der externen Evaluation war in den meisten Ländern auch eine deutliche Steigerung der Ansprüche an das Personal verbunden. Die Rekrutierung der künftigen Evaluatoren erfolgt meist über Ausschreibungen mit hohen Qualifikationsanforderungen, die in einigen Ländern durch aufwendige Verfahren (Self-Assessment-Instrumente, Interviews, mehrtägige fachliche Assessments) überprüft werden. Nach der Zulassung zur Ausbildung als Evaluator/-in finden umfangreiche Qualifizierungen zu einem breiten Themenspektrum statt, die in Blockwochen, Wochenendseminaren und verpflichtenden Weiterbildungsseminaren angeboten werden und in einigen Ländern auch mit einer Prüfungsevaluation enden (vgl. für Bayern: Mühlbauer 2010). Lehrkräfte, Schulentwicklungsmoderatoren, Schulleitungsmitglieder, Seminarleiter der Lehrerbildung, Schulrätinnen und Schulräte, teilweise aber auch Sozial- oder Erziehungswissenschaftler bilden den Bewerberpool, aus dem die künftigen Evaluatorinnen und Evaluatoren ausgewählt werden.

Vor allem in großen Flächenländern werden Grundbestandteile der Ausbildung landeseinheitlich und zentral angeboten, vertiefende und berufsbegleitende Praxismodule oft in regionaler Zuständigkeit organisiert.

In mehr als der Hälfte der Länder arbeiten die ausgebildeten Evaluator/-innen auf festen Planstellen in Vollzeit ohne zeitliche Befristung. Allerdings werden die Evaluationsteams, die je nach Schulgröße in der Regel zwei bis vier Personen umfassen, meist ergänzt durch Evaluatorinnen und Evaluatoren, die als abgeordnete Mitarbeiter/-innen und meist auf befristeten Teilzeit- oder Vollzeitstellen tätig werden. Während die Dauerstellen in Vollzeit oft aus dem Stellenplan der Schulaufsicht durch Verlagerung stammen oder als zusätzliche Planstellen eingerichtet wurden, nutzen einige Länder ausschließlich Stellen aus den Lehrkräftekapiteln und ordnen geeignete Personen befristet mit der besonderen Funktionsübertragung an die zuständige Stelle im Land ab. In einigen Ländern (z. B. Berlin, Hamburg, Rheinland-Pfalz) werden die hauptamtlichen Evaluatoren durch aktive Schulleitungsmitglieder zum Teil gegen gesondertes Honorar als schulfachliche Experten mehrmals im Jahr oder mit halber Stelle einbezogen. Damit wird die Absicht verbunden, sowohl Praxiserfahrungen als auch Schulformkompetenz in den jeweiligen Teams sicherzustellen.

Neben schulischen Evaluatoren werden in einigen Ländern (z. B. Bayern, Hamburg, Berlin) auch nicht-schulische Personen systematisch in die Evaluationsarbeit eingebunden. Eltern, Vertreter von Unternehmen und Verbänden oder andere Personen erhalten dazu eine vorbereitende Qualifizierung und können dann bei Erstattung der anfallenden Kosten an den Evaluationen auch verantwortlich mitwirken. Während in Stadtstaaten diese Einbeziehungen eher organisiert werden können, wurden derartige Versuche in Flächenländern (z. B. Brandenburg oder Niedersachsen) erfolglos abgebrochen.

In Ländern mit auf Dauer angelegten Vollzeitstellen für die externe Evaluation sind diese meist laufbahnrechtlich dem Schulaufsichtsdienst zugeordnet, wobei darauf geachtet wird, sie durch eine besondere organisationsrechtliche Einbindung aus der unmittelbar operativ arbeitenden Schulaufsicht herauszunehmen. Damit verbunden ist eine Besoldung, die sich nach dem Eingangsamt der jeweiligen Schulaufsicht richtet und mit A 14 bis A 16 für leitende Funktionen im System der externen Evaluation verbunden sein kann.

Dies gilt nicht für Evaluatoren/-innen, die auf der Basis von Abordnungen in Teil- oder Vollzeit weiterhin so bezahlt werden wie zuvor. Für die statusrechtlich eher unsichere Absicherung der Evaluatoren/-innen auf Abordnungsbasis im Teilzeitmodell gibt es gute fachliche Begründungen: Erstens soll die tägliche Schulwirklichkeit präsent bleiben und Einfluss auf die Weiterentwicklung der Instrumente und Verfahren haben und zweitens soll die Evaluationstätigkeit direkt in die eigene schulische Arbeit der abgeordneten Mitarbeiterinnen und Mitarbeiter einfließen. Ein weiterer Grund liegt in der von vielen Ländern beabsichtigten „Drehtürfunktion“ der externen Evaluation. „Mehr als 60 ehemalige Inspektorinnen und Inspektoren haben im Anschluss an ihre Inspektorentätigkeit ein höherwertiges Amt angetreten und sind in schulischen Leitungsfunktionen, der Lehrerbildung, der Schulaufsicht oder in Forschungseinrichtungen tätig. Damit wirkt sich die Mitarbeit in der Schulinspektion nachhaltig positiv auf ein gemeinsames Verständnis über Qualitätsentwicklungsprozesse in der Bildungslandschaft Berlins aus.“ (Abgeordnetenhaus Berlin 2012, S. 3) Diese für Berlin dargestellte Erfahrung wird auch in anderen Ländern (z. B. Brandenburg, Hessen, Niedersachsen) angestrebt und erreicht.

Als unerwünschte Nebenfolge bei abgeordneten Evaluatoren kommt es allerdings in den Evaluationsteams zu einer großen Spreizung der Bezahlung bei teilweise identischer Arbeitsleistung. In Brandenburg wird versucht, dieses Problem teilweise durch eine Zulagenregelung aufzufangen, die abgeordneten Lehrkräften zusteht, wenn sie mit mehr als 50% ihrer Arbeitszeit in der externen Evaluation tätig sind. Mit dem Abordnungsstatus ist auch eine hohe zeitliche Unsicherheit und Instabilität im System der Evaluation verbunden, weil Abordnungen relativ leicht in Schuljahresrhythmen von beiden Seiten aufgehoben werden können und damit nur eine geringe berufliche Planbarkeit ermöglichen.

Es bleibt offen, welche Effekte mittelfristig in den Ländern zu beobachten sind, in denen der Wechsel in eine unbefristete Beamtenstelle der externen Evaluation zu einem auf Dauer angelegten Rollenwechsel führt. Als Nebenfolge könnten Evaluatoren dieses Typs nach mehr als zehn Jahren fehlender Schulpraxis dem Vorwurf ausgesetzt sein, sie wären nur bedingt in der Lage, die aktuelle Schulpraxis angemessen beurteilen zu können.

Die Menge des in den Ländern zur Verfügung stehenden bzw. notwendigen Personals richtet sich nach politischen und finanziellen Vorgaben: Zentral ist dabei vor

allein die Evaluationsfrequenz, die derzeit in fast allen Ländern so festgelegt wurde, dass innerhalb von vier bis fünf Jahren alle öffentlichen Schulen einmal evaluiert sein sollen. Die Begründung für diesen Zeitraum ist eher der Absicht geschuldet, den erforderlichen Personalbedarf klein zu halten, als einer plausiblen Begründung oder gar empirischen Evidenz, in welchem Abstand Schulen sinnvollerweise evaluiert werden sollten. Tatsächlich hat der erste Durchgang in vielen Ländern viel länger gedauert, da es entweder nicht gelang, das erforderliche Personal zu gewinnen und zu qualifizieren, oder die Evaluatoren schneller als angenommen in andere berufliche Positionen gewechselt sind.

3. Institutionelle Einordnung der externen Evaluation in die Bildungsadministration der Länder

Mit der Absicht, externe Evaluationen als neues Instrument der Qualitätssicherung einzuführen, war zwangsläufig auch die Frage verbunden, ob diese neue Akteursgruppe der Schulevaluatoren als Ersatz oder als Ergänzung für eine bestehende Schulaufsicht verstanden werden soll, ob die externe Evaluation als Teil der operativen Schulaufsicht oder personell und organisatorisch getrennt davon praktiziert werden soll. Tatsächlich hat die klärende Diskussion um die künftige Rolle und Funktion der operativen Schulaufsicht in vielen Ländern nicht stattgefunden. Die Analyse zahlreicher Schulgesetze der Länder offenbart ein Ausklammern dieser Debatte, wie sich am Beispiel Hamburgs gut zeigen lässt. Dort wurde der § 85 mit der Überschrift *Schulaufsicht, Schulberatung und Schulinspektion* neu gefasst, wobei die klassische Definition der Schulaufsicht als Fach-, Dienst- und Rechtsaufsicht (vgl. § 85, Abs. 1) beibehalten wurde und lediglich ein neuer Absatz 3 hinzugefügt wurde, mit dem der Auftrag der Schulinspektion zur Untersuchung der Qualität der Bildungs- und Erziehungsprozesse beschrieben wurde. Gleichzeitig wird im Absatz 2 daran festgehalten, dass die Schulaufsicht die Schulen in der Wahrnehmung ihrer Aufgaben auch im Bereich der erweiterten Selbstverantwortung berät und unterstützt. Damit wird es in die Hand der Akteure gegeben, wie sie dieses Regelungsdefizit persönlich auflösen.

Das Handlungsrepertoire der Schulaufsicht in allen Ländern ist dabei durchaus vielfältig: Durch Schulbesuche und Hospitationen im Unterricht, Teilnahme an Konferenzen, Gespräche mit Schulleitungen und Lehrkräften, Empfehlungen, Hinweise, durch Weisungen zur Durchsetzung allgemeiner Vorgaben bestehen Möglichkeiten zur Einflussnahme auf die Qualität der Arbeit der Schulen. Zur Fachaufsicht gehört auch, darauf zu achten, dass die Schulen ihren Auftrag zweckmäßig und auf der Grundlage der gesetzlichen Vorgaben wahrnehmen. Das Feedback, das die Schulaufsicht an Schulen gab, ist allerdings eher individuell (bezogen auf die

Schulleitung, einzelne Lehrkräfte), selten schriftlich dokumentiert, spontan und in der Regel unsystematisch.

Stryck ist zuzustimmen, wenn er feststellt:

„Die funktionale Diffusität hatte eine geradezu paradoxe Konsequenz: Qualitätskontrolle – implizite Aufgabe einer Aufsicht über das Schulwesen – ist nachgerade tabuisiert worden! Qualitätskontrolle war nie wirklich eine zentrale und – wo sie denn stattfand – folgenreiche Signatur des Systems. Es gab weder eine politisch offensiv vertretene Explikation eines Aufsichtsprogramms noch einen wirksamen Praxisauftrag an die Schulaufsicht, die Qualitätsdifferenzen auf allen Systemebenen systematisch zu untersuchen und folgenreich zu thematisieren.“ (Stryck 2000, S. 116)

Ähnlich beschreibt es Avenarius in einer Bilanz nach dem ersten Ländervergleich PISA-E 2002 mit der Feststellung:

„Bislang war und ist Qualitätskontrolle nicht wirklich ein prägendes Merkmal der Schulaufsicht. Es gab und gibt keine politisch offensiv vertretenen, keine verbindlichen Leitlinien, an denen sich die Schulaufsicht orientieren könnte. Auch gab und gibt es, jedenfalls in den meisten Ländern, keinen verbindlichen Auftrag an die Schulbehörden, die Qualitätsdifferenzen im Schulbereich systematisch zu untersuchen und auf Ursachen und Konsequenzen hin zu thematisieren.“ (Avenarius 2002, S. 9)

Die in allen Ländern faktisch auch als Auftrag der externen Evaluation bestehende Kontrollfunktion, festzustellen, ob und inwieweit die Schulen ihrem Auftrag nachkommen, bestehende Regelungen, Vorschriften oder inhaltliche Anforderungen aus dem jeweiligen Qualitätsrahmen zu beachten bzw. zu erfüllen, beschreibt im Grunde genommen auch einen von der bisherigen Schulaufsicht einzulösenden Auftrag, den diese augenscheinlich aber nicht systematisch wahrgenommen hat.

Ein möglicher Grund dafür wird in der Mentalitätsallianz zwischen der Schulaufsicht und dem pädagogischen Leitungspersonal gesehen. Dem praktizierenden Schulaufsichtsbeamten, der seinen beruflichen Werdegang oft vom Lehrer über den Schulleiter zur Aufsicht über Schulleiter und Lehrer genommen hat, wird fehlende professionelle Distanz zu seiner Klientel unterstellt. Im Ausbalancieren zwischen Macht- und Einflusskonstellationen kommt es zu oft zur Verbrüderung.

„Dem Dilemma ist auch nicht dadurch zu entkommen, dass Schulaufsicht künftig – in der Figur einer rhetorischen Adaptation – Beratung leisten soll. Bleibt es gleichzeitig beim Kontrollauftrag, kommt Schulaufsicht in das Zirkeldilemma, die Qualität der von ihr beratenen Schulen kontrollieren zu sollen. Aber was wird dann eigentlich kontrolliert? Sicher doch auch die beabsichtigten oder unbeabsichtigten Induktionswirkungen der Beratung.“ (Stryck 2000, S. 117)

Der Versuch, das Personal der mit der externen Evaluation beauftragten neuen Teile der Bildungsverwaltung von den befürchteten oder tatsächlichen negativen Kontaminationen der operativen Schulaufsicht fernzuhalten, hat innerhalb der Länder zu sehr unterschiedlichen Vorgehensweisen geführt. Grundsätzlich gibt es dafür vier Organisationsmodelle:

- Modell 1: Durchführung der externen Evaluationen durch eine rechtlich selbstständige, neugegründete Einrichtung
- Modell 2: Vergabe der externen Evaluationen an ein externes Institut
- Modell 3: Erweiterung der Aufgaben der bisherigen pädagogischen Landesinstitute durch neue Abteilungen oder Referate oder Neugründungen und Zusammenlegung mit anderen Teilen der bisherigen Bildungsverwaltung
- Modell 4: Beauftragung der staatlichen Schulaufsichtsbehörden mit der Durchführung der externen Evaluation, wobei darauf geachtet wird, diese organisatorisch oder räumlich voneinander getrennt zu organisieren.

Das Modell 1, ähnlich wie in verschiedenen europäischen Ländern vollkommen eigenständige Landesorganisationen aufzubauen, wurde in keinem Land dauerhaft verfolgt. Die 2006 gegründete eigenständige Landesschulinspektion in Niedersachsen mit Sitz in Bad Iburg wurde 2012 im Zuge von Verwaltungsmodernisierungen zu einer Abteilung im neu gegründeten Niedersächsischen Landesinstituts für schulische Qualitätsentwicklung (NLQ). Und das in Bremen in einem ersten Durchgang praktizierte Modell 2, die Evaluation einem externen Institut zu übertragen, wurde für den zweiten Durchgang aufgegeben.

Heute dominiert als Organisationsmodell die Integration der externen Evaluation als mehr oder weniger eigenständige Abteilung in nachgeordneten Einrichtungen. Diese verknüpfen häufig als Landesbehörde allgemeine schulfachliche Aufgaben, üben die Funktion von Fortbildungsinstituten aus oder nehmen in den anderen Abteilungen andere Funktionen aus dem jeweiligen Aufgabenspektrum der Qualitätssicherung wahr (Bildungsberichterstattung, zentrale Prüfungsplanung und -organisation, Lernstandsmessungen, Rahmenplanentwicklung usw.).

Das Modell 4 wird in wenigen Ländern (z. B. Berlin, Nordrhein-Westfalen, Thüringen) praktiziert. Dabei wird der Einsatz der externen Evaluation direkt durch Referate des zuständigen Schulministeriums bzw. in NRW durch das Dezernat Qualitätsanalyse in der Schulabteilung der Bezirksregierung gesteuert. Durch besondere Festlegungen (z. B. Weisungsungebundenheit bei der Qualitätsfeststellung, Verzicht auf Weisungsbefugnisse gegenüber den Schulen, Verschwiegenheitsverpflichtungen) wird dabei versucht, eine Vertrauensbeziehung zu den Schulen zu schaffen, um so die Akzeptanz und die Wirkungen der Qualitätsfeststellungen zu erhöhen.

Inwieweit die mit derartigen organisationsrechtlichen Entscheidungen erhofften Wirkungen („Unabhängigkeitsparadigma“) tatsächlich eintreten und die subjektive Wahrnehmung der Tätigkeit der Schulevaluatorinnen und -evaluatoreen beeinflussen, ist bisher nicht untersucht. Es gibt jedoch Hinweise darauf, dass die rechtlich-strukturelle Einbindung weniger relevant ist als die Frage, wie stark Schulrätinnen und Schulräte der operativen Schulaufsicht das „Gesicht“ der externen Evaluation prägen.

Trotz aller Bekundungen in den Ländern, den neu geschaffenen Organisationen oder Behördenteilen für die externe Evaluation in Fragen der Vorgehensweisen weitgehende Unabhängigkeit einzuräumen, besteht natürlich die Möglichkeit, dass die jeweils gewählte Ministeriumsspitze bildungspolitisch geprägte Entscheidungen zum Rahmenprogramm, den Indikatoren, dem Umgang mit Ergebnissen und der Rekrutierung des Personals trifft. So bleibt es nicht aus, dass fertiggestellte Jahresberichte mit Ergebnissen der externen Evaluation auf Landesebene (zuletzt in Sachsen und Niedersachsen) auch mal unveröffentlicht bleiben, bis die gerade stattfindenden Landtagswahlen vorbei sind.

Die Entscheidung der meisten Länder, die externe Evaluation personell, institutionell und oft auch räumlich von der „alten“ Schulaufsicht abzukoppeln, macht es umso dringlicher, das arbeitsteilig gedachte Modell der Qualitätssicherung durch Rollen- und Verfahrensklarheit und ein präzises Schnittstellenmanagement zu konturieren.

Die Einbindung der operativen Schulaufsicht in den Prozessverlauf der externen Evaluation verläuft dabei in den Ländern sehr unterschiedlich und reicht von der persönlichen Anwesenheit bei der Berichtspräsentation in der Schule über die zeitgleiche Zustellung des Endberichts bis zur Verpflichtung zu Bilanzgesprächen bzw. dem Abschluss von Zielvereinbarungen.

Neu zu definieren ist vor diesem Hintergrund allerdings, wer die Ergebnisverantwortung hat – ausschließlich die Einzelschule oder auch die zuständige Schulaufsicht? Da die Evaluatoren weder Weisungsbefugnisse gegenüber der Schule haben noch als Schulberater qualifiziert sind bzw. sich qua Rollenfestlegung nicht als solche verstehen, kommt es vor allem bei Schulen mit erheblichem Entwicklungsbedarf häufig zu Enttäuschungen. Während vonseiten der Schulbehörden gelegentlich beklagt wird, dass es mangels harter Sanktionsmöglichkeiten gegenüber entwicklungsresistenten Schulen kaum Einflussmöglichkeiten gibt, beklagen Schulleitungen eher den fehlenden „Rechtsanspruch“ auf die erforderliche Unterstützung durch die in den Ländern oft nicht hinreichend ausgestatteten Unterstützungssysteme. Ein systemisch und strukturell verankertes Zusammenspiel der Akteure externe Evaluation, Schulaufsicht und Unterstützungseinrichtungen ist oft erst in Ansätzen erkennbar.

In diesem Zusammenhang wird derzeit in einigen Ländern die Frage neu gestellt, ob die bisherige strikte Abgrenzung von Qualitätsfeststellungen durch die externe Evaluation einerseits und der Aufgabe der Beratung und Unterstützung der

Schulen durch die Schulaufsicht andererseits sinnvoll ist. Eine Antwort zu dieser Frage wird nicht isoliert gegeben werden können, sondern nur, wenn alle zum Teil nebeneinander und unverbunden implementierten Bausteine eines behaupteten Qualitätsmanagements überdacht und verknüpft werden.

4. Zukünftige Herausforderungen und Perspektiven

4.1 Externe Evaluation als Profession?

Die in den letzten Jahren zu beobachtende Ausweitung der Handlungsfelder, in denen sich Evaluation im Bildungsbereich etabliert hat, hatte nicht immer eine Entsprechung in den vorhandenen personellen Ressourcen. Dies ist auch darauf zurückzuführen, dass (Schul-)Evaluation bisher kein ausgewiesenes Berufsfeld mit entsprechenden Anforderungen an die Aus- und Weiterbildung darstellt und das dafür benötigte Personal nicht aus dem Bestand rekrutiert werden kann. Der Charakter der Evaluation als Querschnittswissenschaft, die sich auf Anforderungen aus unterschiedlichen Fachdisziplinen bezieht, macht eine entsprechende Qualifizierung auch nicht einfach.

Während die Anfangsphase des Aufbaus der externen Evaluation eher durch zurückhaltende Festlegungen zur personellen Ausstattung und Nutzung von Personen aus unterschiedlichen beruflichen Kontexten gekennzeichnet war, stellt sich mit der dauerhaften Institutionalisierung der externen Evaluation die Frage, ob die Länder beginnen werden, die bisher sehr unterschiedlichen Vorstellungen von den erforderlichen Kompetenzen und dem daraus resultierenden (Nach-)Qualifizierungsbedarf für externe Schulevaluatoren abzustimmen und sich künftig auf gemeinsame Standards für die Qualifizierung des Personals einigen wollen bzw. können. Eine Frage, die sich natürlich auch auf andere Aspekte der Evaluation beziehen lässt.

Eine gute Diskussionsgrundlage für die Entwicklung solcher Standards sind die von der DeGEval – Gesellschaft für Evaluation (2008) entwickelten Empfehlungen. Sie können als grundlegende Anforderungen und Kompetenzen verstanden werden, die für die angemessene Durchführung von externen Evaluationen unerlässlich sind, und lassen sich in vier Kompetenzfeldern zusammenfassen, wobei die Vermittlung der Evaluationskompetenzen an die Praxis der Evaluation angebunden sein muss:

- Theorie und Geschichte der Evaluation
- Methodenkompetenzen
- Organisations- und Feldkenntnisse
- Sozial- und Selbstkompetenzen
- Praxis der Evaluation

Eine ähnliche Strukturierung der erforderlichen Kompetenzen haben u.a. auch amerikanische Forscher entwickelt. In Anlehnung an Stevahn et al. (2005) lassen sich danach sechs unterschiedliche Kompetenzbereiche definieren, die von professionellen Evaluatoren abgedeckt werden sollten:

1. *Professional Practice* (bezieht sich auf wesentliche Werte und Normen, die der Evaluationspraxis zugrunde liegen)
2. *Systematic Inquiry* (umfasst eher technische Kompetenzen der Datenerhebung und -auswertung, der Analyse, Interpretation und Berichterstattung)
3. *Situational Analysis* (bezieht sich auf Kompetenzen zur Analyse und Beachtung des Einzelfalls und seiner spezifischen Interessen, Probleme und Kontextbedingungen)
4. *Project Management* (Kompetenzen der operativen Durchführung von Evaluationen wie Kostenplanung, Koordination und Überwachung)
5. *Reflective Practice* (umfasst die Selbstwahrnehmung hinsichtlich professioneller Stärken und Entwicklungsbedarfe)
6. *Interpersonal Competence* (bezieht sich auf Kompetenzen, die für die zwischenmenschliche Kommunikation relevant sind)

Trotz ähnlicher Begrifflichkeiten und Dimensionen werden vor allem im Bereich *Reflective Practice* und *Interpersonal Competence* Vertiefungen und Konkretisierungen vorgenommen, die für die externe Evaluation besonders relevant sind. Angesichts der deutschen Ausgangslage käme es allerdings zuerst darauf an, die Diskussion länderübergreifend zu führen und Festlegungen mit dem Ziel zu treffen, ein grundlegendes Anforderungsprofil zu formulieren und so aufseiten der Auftraggeber wie auch bei den Evaluatorinnen und Evaluatoren Klarheit und Sicherheit über die erwarteten Kompetenzen zu erlangen.

Es spricht viel dafür, die Tätigkeit der Schulevaluatorinnen und -evaluatorennen nicht beamtenrechtlich als Amt auf Lebenszeit auszugestalten, sondern den in einigen Ländern beschrittenen Weg einer hohen Durchlässigkeit zwischen verschiedenen Ebenen im Schulsystem systematisch weiterzuentwickeln. Lehrkraft, Schulleitung, Schulaufsicht, Seminarleiter, Fortbildner/Berater – in allen und aus allen Tätigkeitsfeldern sollte es möglich sein, die Qualifikation als externer Evaluator auf der Basis bundesweit abgestimmter professioneller Standards zu erwerben und nach mehreren Jahren der Ausübung dieser Funktion sich auf andere Funktionsstellen bewerben zu können. Dazu wären allerdings dienst- und laufbahnrechtliche Veränderungen notwendig, die innerhalb der Bildungsadministration der Länder derzeit nicht erkennbar sind. Eine deutlich höhere Flexibilität und Durchlässigkeit innerhalb und zwischen den Systemebenen wäre ein Gewinn für alle Beteiligten.

„In der Evaluation jedoch reicht es in der Regel nicht aus, alleine hervorragend methodisch, theoretisch, organisations- und feldspezifisch oder im Hinblick auf Sozial- und Selbst-Kompetenzen ausgebildet zu sein. Auch ein ausschließlich praxisorientierter Zugang greift zu kurz. Viel mehr ist eine Synthese zwischen diesen unterschiedlichen Fertigkeiten herzustellen, um in einem häufig von der Gleichzeitigkeit unterschiedlicher politischer und wissenschaftlicher Intentionen geprägten Feld bestehen zu können.“ (DeGEval 2008, S. 31)

Angesichts der sehr unterschiedlichen biografischen Vorkenntnisse bei den bisher tätigen Personen sollte durch Definition von Modulen die Möglichkeit der Anrechnung von bereits erworbenen Qualifikationen z.B. aus der Schulleitungsfortbildung geschaffen werden. Ohnehin sollte aus fachlichen Überlegungen geprüft werden, Teilqualifikationen für Schulleitungen, Schulaufsicht und Schulleitungen gemeinsam als Fortbildungsbausteine anzubieten, müssen diese drei Akteure doch bei der Planung, Durchführung und Auswertung der externen Evaluation künftig intensiver zusammenarbeiten als bisher. Klar ist allerdings, dass eine derartige Professionalisierung unter den aktuell in mehreren Ländern bestehenden Bedingungen der Unsicherheiten in Bezug auf Rolle und Status von Teilzeitevaluatoren auf Abordnungsbasis nicht zu leisten ist. Die bildungspolitisch Verantwortlichen müssen sich daher entscheiden, ob der vor zehn Jahren begonnene Prozess einer flächendeckenden Einführung von externer Evaluation mit hoher Qualität fortgesetzt werden soll oder unter dem Druck von Haushaltssanierung eine halbherzige Reformidee bleiben wird, weil man die erforderlichen personellen und finanziellen Rahmenbedingungen nicht schaffen will, die für einen Erfolg dieses Instruments der Qualitätssicherung notwendig sind. Betrachtet man die für eine Schulevaluation erforderlichen Kosten und setzt diese in Beziehung zu den im Gesamtsystem aufgewendeten Ausgaben, wird deutlich, dass vergleichsweise wenig Geld für dieses Instrument der Qualitätsentwicklung erforderlich ist. In Brandenburg werden jährlich ca. 1,3 Millionen Euro für das in der externen Evaluation tätige Personal aufgewendet, hinzu kommen ca. 140.000 Euro für Sachausgaben.² Legt man diesen Betrag auf die 730 öffentlichen Schulen um, fallen pro Jahr ca. 2.000 Euro Kosten pro Schule an. Die externe Evaluation einer Schule (im Fünf-Jahres-Zyklus je einmal) kostet pro Schule danach durchschnittlich 10.000 Euro. Bezieht man den Gesamtbetrag auf die ca. 245.000 Schülerinnen und Schüler an öffentlichen Schulen, entspricht das einem Betrag von ca. 6 Euro pro Jahr pro Person. Gleichzeitig sind pro Jahr in Brandenburg ca. 900 Millionen Euro für die Personalkosten aller Lehrkräfte erforderlich. Entgegen der gelegentlich aus Gewerkschaftskreisen vorgebrachten These liefert der Verzicht auf externe Evaluationen keinen nennenswer-

2 Siehe Haushaltsplan MBJS für 2014, Kapitel 05 140, zusätzlich 10 Lehrerstellen aus dem allgemeinen Schulkapitel.

ten finanziellen Beitrag der für Verbesserungen im Schulsystem an anderer Stelle eingesetzt werden könnte.

4.2 Externe Evaluation im Kontext der politischen Auseinandersetzungen zwischen Regierungsparteien

Bildungspolitik ist im Rahmen des Föderalismus in Deutschland einer der wichtigen Politikbereiche, der fast ausschließlich in den Händen der Länder liegt. Die bestehende Konstruktion des Art. 91 Grundgesetz begrenzt die Kooperationsmöglichkeiten des Bundes mit den Ländern erheblich, ein Umstand, der seit vielen Jahren auf immer stärkere Kritik stößt. Umso bedeutsamer ist es daher, welche programmatischen Vorstellungen die jeweils regierenden Parteien in ihre Regierungsbündnisse auf Landesebene einbringen. Eine gute Grundlage zur Analyse der jeweils in einem Regierungsbündnis für bedeutsam gehaltenen Überzeugungen und Absichtserklärungen sind die inhaltlichen Aussagen zu Bildung in den jeweiligen Koalitionsverträgen.

Anders als bei Wahlprogrammen, in denen Parteien sich aufgrund der Wettbewerbssituation eher voneinander abgrenzen wollen und „ideologische Grundmuster“ das Fundament für konkrete Forderungen und Konzeptionen bilden (vgl. Nikolai & Rothe 2013), sind Koalitionsverträge das Ergebnis komplizierter Prozesse der Konsenssuche. Dabei werden tatsächliche Meinungsverschiedenheiten zwischen den Parteien oft in unverbindliche Prüfaufträge gekleidet, als ungelöste Streitfragen durch Nichtbefassung entschieden oder einem Koalitionspartner als (notwendiger) „Erfolg“ zugestanden. Bei der Interpretation dieser Texte muss außerdem in Rechnung gestellt werden, dass sie oft nur in Kenntnis der bildungspolitischen Konfliktlagen verständlich sind, die in der Legislaturperiode vor dem Abfassen dieser Texte durch die neue (alte) Regierung bestanden haben. Andererseits zeigt die Analyse der schriftlich formulierten Positionen aber auch eine für alle Länder und über Parteiendifferenzen hinweg geltende Prioritätensetzung, die sich in den letzten 20 Jahren deutlich erkennbar verändert hat. Insbesondere die nach PISA 2000 einsetzende, bundesweite Diskussion über vermeintliche oder tatsächliche Defizite im deutschen Schulsystem hat auch in den Koalitionsverträgen zu neuen Themensetzungen und spezifischen Schwerpunktverschiebungen geführt. Eine Analyse der 16 Koalitionsverträge aller Länder (seit 2008) sowie die Hinzunahme der 5 nach Landtagswahlen in 2013 und 2014 neu abgeschlossenen Koalitionsverträge mit dem Fokus auf Aussagen zur Schul- und Unterrichtsqualität bzw. zu interner und externer Evaluation zeigt beträchtliche Unterschiede, aber auch eine zunehmende Themenkonvergenz.

Alle untersuchten Koalitionsverträge behandeln im Bildungsteil mehr oder weniger konkret u. a. die Aspekte Bildungsgerechtigkeit, Durchlässigkeit, individuelle

Förderung, Umsetzung der Inklusion, Eigenverantwortlichkeit der Einzelschule, Ausbau von Ganztagsangeboten, Ressourcenplanungen in Bezug auf demografische Renditen, Einstellungskorridore, Schulsozialarbeit, Lehrerbildung und Finanzierung freier Schulen. Fragen der Schul- und Unterrichtsqualität werden eher randständig behandelt und in vielen Texten auf die Verabredung von Maßnahmen zur Vermeidung von Unterrichtsausfall reduziert.

Aussagen zur Einführung, Fortführung oder Weiterentwicklung von externer Evaluation findet sich in den Koalitionsverträgen von acht Ländern (Stand November 2014: BB, BY, HE, HH, NI, NRW, SL, ST), wobei in vier dieser Länder (BY, HE, NRW, SL) auch die Verknüpfung von interner und externer Evaluation als Ziel beschrieben wird.

„Wir werden die Qualität der bayrischen Bildung durch interne und externe Evaluation sowie länderübergreifende Vergleiche sichern.“ (Koalitionsvertrag Bayern, CSU/FDP, 2008–2013)

„Die selbstständig arbeitenden Schulen müssen die Qualität ihrer Arbeit überprüfen lassen. Deswegen werden wir interne und externe Systeme zur Qualitätssicherung weiterentwickeln.“ (Koalitionsvertrag Hessen, CDU/FDP, 2009–2014)

„SPD und GRÜNE werden die Qualitätsanalyse neu ausbalancieren. Sie soll zielgerichteter werden und den für die Schulen zu leistenden Aufwand reduzieren. Wir werden die Selbstevaluation von Schulen unter Einbeziehung von Eltern, Lehrkräften und Schülerschaft in Balance zur Qualitätsanalyse stärken.“ (Koalitionsvertrag NRW, SPD/Grüne, 2012–2017)

„Die jetzige Struktur der systematischen externen Evaluation soll gefestigt, mit Zielvereinbarungen insbesondere bezüglich Fortbildungskonzepten verknüpft und durch eine interne Evaluation der Schulen systematisch ergänzt werden.“ (Koalitionsvertrag Saarland, CDU/FDP/Grüne, 2009–2014)

Im Koalitionsvertrag von Bremen (2011–2015, SPD/Grüne) wird eine Verabredung getroffen, wonach „[...] jede Schule ein Qualitätsmanagement einführen [soll]. Die Schülerinnen und Schüler sollen regelmäßig die Qualität des Unterrichts und das Schulklima bewerten“. Tatsächlich hat das Land Bremen in den letzten Jahren auch andere Wege beschritten, die Qualität der Schul- und Unterrichtsentwicklung zu beurteilen, als durch eine systematische, standardisierte externe Evaluation, wie sie von den übrigen Ländern durchgeführt wird. Schwerpunkt wird auf die verbindliche Möglichkeit zum Schüler-Lehrer-Feedback gelegt.

Obwohl die Einführung von systematischen externen Evaluationen erhebliche Auswirkungen auf die Rolle und Funktion der bestehenden Schulaufsicht hat, wird

dieser Zusammenhang nur in zwei Ländern auch Gegenstand von koalitionspolitischen Verabredungen:

„Vor dem Hintergrund der zunehmenden inhaltlichen und organisatorischen Eigenverantwortung der Schulen auf der einen Seite und neu entwickelter Instrumente zur Wahrnehmung der staatlichen Aufsichtspflicht durch externe Evaluation auf der anderen Seite müssen Aufgaben und Organisation der Schulaufsicht neu bestimmt werden.“ (Koalitionsvertrag Sachsen-Anhalt, CDU/SPD, 2011–2016)

„Die staatlichen Schulämter werden vermehrt zu Einrichtungen der Schulaufsicht und zu Servicestellen für Verwaltungsaufgaben umgebaut. Ihre Eingriffsmöglichkeiten auf die Schulen werden vermindert; darüber hinaus werden sie – ebenso wie die Schulen – einer regelmäßigen Evaluation unterzogen.“ (Koalitionsvertrag Hessen, CDU/FDP, 2009–2014)

Die explizite Übertragung des Instruments der externen Evaluation auf die staatlichen Schulämter findet sich in keinem anderen Koalitionsvertrag und stellt auch fachlich eine Herausforderung dar, die sich auch bisher kaum eine Landesregierung vorgenommen hat.

In allen übrigen sieben Ländern gibt es in den Koalitionsverträgen keine Aussagen zum Thema der Verbesserung der Unterrichtsqualität oder dem Stellenwert interner oder externer Evaluation. Dies überrascht insofern, da in fast allen Ländern (bis auf Schleswig-Holstein) tatsächlich seit Jahren zum Teil mit großem Aufwand betriebene Systeme externer Evaluation implementiert worden sind und weiterentwickelt werden. Die Nicht-Befassung mit der bildungspolitischen Strategie, Schul- und Unterrichtsentwicklung durch externe (und interne) Evaluation zu befördern, lässt sich auch nicht durch die jeweils vertragschließenden Parteien erklären. Alle fünf untersuchten Parteien (CDU/CSU, SPD, Grüne, Linke, FDP) sind in Regierungsbündnissen oder Alleinregierungen beteiligt, die im Koalitionsvertrag oder Regierungsprogramm keine expliziten Verabredungen zur externen Evaluation treffen, gleichwohl jedoch bestehende Maßnahmen auf diesem Gebiet unterstützen und fortführen.

Wird die externe Evaluation in einem Land im Konflikt zwischen Regierung und Opposition eingeführt oder Gegenstand öffentlicher Kontroversen, erlangt sie gelegentlich auch eine politische Bedeutung, die sich in den Koalitionsvereinbarungen mehrerer Legislaturperioden niederschlägt. So vereinbarten CDU und SPD in Schleswig-Holstein für die 16. Legislaturperiode (2005–2010), dass „[...] der Schul-TÜV (EVIT) unter Einbeziehung der Leistungsvergleiche fortgeführt und weiterentwickelt [wird]“. Für die nachfolgende Regierungszeit von CDU und FDP wurde 2009 im Koalitionsvertrag auf Drängen der FDP festgelegt: „[...] deshalb werden wir den bürokratischen Schul-TÜV (‘EVIT’) genauso abschaffen wie Re-

gelingen, die in der Vergangenheit gute Pädagogik durch Papierberge ersetzt haben.“ Im Koalitionsvertrag von SPD, Grünen und dem SSW für die aktuelle 18. Legislaturperiode (2012–2017) finden sich keine Aussagen zur externen Evaluation. Das durch vorzeitige Neuwahlen 2012 entstandene Regierungsbündnis hat sich zu dieser Frage nicht festgelegt, gleichwohl wird das Thema die Koalitionsfraktionen weiter beschäftigen.

Die Analyse der Koalitionsverträge aller Länder für die vergangenen zehn Jahre zeigt, dass die Nennung bzw. Nicht-Behandlung der externen Evaluation nur bedingt Rückschlüsse auf das Vorhandensein derartiger Verfahren oder ihre dauerhafte Stabilität über mehrere Legislaturperioden zulässt. Da von der interessierten Öffentlichkeit in Wahlkämpfen eher Themen wie „Vermeidung von Unterrichtsausfall“, „Sicherung des Lehrkräftebedarfs“ oder „Ausbau von Ganztagsangeboten“ nachgefragt werden, bleibt die Nennung anderer Themen, die für die Qualität eines Schulsystems von Bedeutung sind, eher subjektiven Vorlieben und Schwerpunktsetzungen der konkreten Verhandlungspartner geschuldet als rational nachvollziehbaren Argumenten.

4.3 Stellenwert der externen Evaluation in der Gesamtstrategie der Länder zum Bildungsmonitoring

Nachdem sich Deutschland nach langen Jahren der Abstinenz 1995 zum ersten Mal wieder an der TIMSS-Studie beteiligte, führte die danach einsetzende bildungspolitische Diskussion dazu, dass die KMK sich unter dem öffentlichen Druck entschloss, ihre Zurückhaltung hinsichtlich einer systematischen und vergleichenden empirischen Überprüfung von Schülerleistungen aufzugeben. Schon acht Monate später fasste die Kultusministerkonferenz am 24. Oktober 1997 den sogenannten Konstanzer Beschluss (KMK 1997) und leitete damit eine Vielzahl von Maßnahmen ein, deren nachhaltigste Wirkung darin bestand, künftig auch an weiteren internationalen Leistungsvergleichen der OECD und IEA teilzunehmen:

„Die Kultusministerkonferenz sieht im Hinblick auf die Gleichwertigkeit der schulischen Ausbildung, die Vergleichbarkeit der Schulabschlüsse sowie die Durchlässigkeit des Bildungssystems innerhalb der Bundesrepublik Deutschland in der Entwicklung von Maßnahmen zur Sicherung der Qualität schulischer Bildung eine wichtige Aufgabe. Im Hinblick auf diese Zielsetzung und zur Qualitätssicherung in Schulen halten es die Mitglieder der Kultusministerkonferenz für erforderlich, in den Ländern Instrumente zur Evaluation zu entwickeln und zu erproben und über die gewonnenen Ergebnisse in einen breiten Erfahrungsaustausch einzutreten.“ (Auszug aus dem Konstanzer Beschluss der Kultusministerkonferenz vom 24. Oktober 1997)

Die im Beschluss der Kultusministerkonferenz von Konstanz 1997 postulierte Absicht, „in den Ländern Instrumente zur Evaluation zu entwickeln und zu erproben und über die gewonnenen Ergebnisse in einen breiten Erfahrungsaustausch“ zu treten, wurde zu einem späteren Zeitpunkt, nachdem die nationalen Bildungsstandards beschlossen, das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) gegründet und die Nationalen Bildungsberichte in Auftrag gegeben worden waren, in einem umfassenden Text noch einmal präzisiert:

„[...] ist es erforderlich, Prozesse der Qualitätsentwicklung und Standardsicherung auf allen Ebenen [...] systematisch umzusetzen und miteinander zu verbinden. Insbesondere muss sichergestellt werden, dass Informationen über die Qualität des Bildungssystems so weit wie möglich auch für die Entwicklung jeder einzelnen Schule genutzt werden können.“ (KMK 2006)

Trotz der Intention, qualitätsrelevante Informationen auch zur Entwicklung der Einzelschule bereitzustellen, enthält die Gesamtstrategie der KMK keinerlei Hinweise auf die von vielen Ländern zu diesem Zeitpunkt bereits eingeführten Verfahren der externen Evaluation, sondern beschränkt ihre Zielsetzung bezüglich der Maßnahmen auf Schulebene auf die Einführung nationaler Vergleichsarbeiten. Auch die bestehenden Gremien wie die Amtschefkommission Qualitätssicherung oder der Schulausschuss der KMK haben sich bisher mit dem Aufbau, den Ergebnissen oder Problemen der externen Evaluation in den Ländern nicht befasst. Während die Steuerungsgruppe Vergleichsarbeiten als Gremium der Absprache zwischen den Ländern seit vielen Jahren etabliert ist und die von ihr gewählte Sprechergruppe regelmäßig den Schulausschuss der KMK über ihre Arbeit informiert bzw. auch vom Schulausschuss Aufträge erhält, ist eine derartige Verzahnung der inzwischen bestehenden Konferenz der Leitungen der Schulinspektorate der Länder mit anderen KMK-Gremien noch nicht vorgesehen.

In der Vergangenheit haben die Länder ihre Konzeptionen der externen Evaluation eher nebeneinander entwickelt; gelegentliche Absprachen oder Ländercluster mit gemeinsamen Entwicklungslinien haben daran nicht wirklich etwas geändert. Inzwischen besteht eine gewachsene Kontinuität im gemeinsamen Arbeits- und Diskussionsprozess, in den alle Länder einbezogen sind. Dazu haben u. a. auch mehrere Ländertreffen in den letzten Jahren beigetragen sowie eine durch kritische Ergebnisse der Bildungsforschung ausgelöste Infragestellung der Nützlichkeit der Verfahren der externen Evaluation. Ziel dieser Kooperation ist es, durch Vereinbarungen zwischen den Ländern bundesweite Absprachen über professionelle Standards für Evaluаторinnen und Evaluatoren, gemeinsame Arbeit an der Verbesserung der Evaluationsinstrumente und Erarbeitung eines gemeinsamen Qualitätsleitbildes vorzunehmen.

Die im Jahr 2013 erklärte Absicht der KMK, fünfzehn Jahre nach dem Konstanzer Beschluss ihre Gesamtstrategie zu überdenken und auf bestehende Probleme und Kritik zu reagieren, könnte eine Chance sein, eine kritische Würdigung des bisherigen Prozesses der sogenannten empirischen Wende in der Bildungspolitik vorzunehmen. Nicht erst in dem von der KMK organisierten ersten Fachgespräch über Bilanz und Perspektive der Gesamtstrategie am 4. Dezember 2013 wurde deutlich, dass es zahlreiche Probleme gibt, auf die Bildungspolitik und Bildungsforschung reagieren müssen, um die bestehenden Akzeptanzprobleme und Transferdefizite zu bearbeiten. Neben den unbestreitbaren Leistungen der Gesamtstrategie sind konzeptionelle Brüche und Grenzen des Programms deutlich geworden, die weitere Finanzierbarkeit bereitet Probleme, die Nutzung der Ergebnisse des Programms auf den unterschiedlichen Ebenen ist widersprüchlich, konfliktträchtig und zum Teil unterkomplex und die strukturellen Grenzen der Verständigungsbereitschaft und Koordinierungsfähigkeit der Länder wurden immer deutlicher. Der Spagat zwischen der Aufgabe der KMK, gesamtgesellschaftliche Verantwortung im Bildungswesen wahrzunehmen, und den zum Teil parteipolitisch-ideologisch motivierten Länderinteressen, die von häufig wechselnden Ministerinnen und Minister repräsentiert werden, gelingt heute schlechter als zu den Zeiten des PISA-Schocks. Die in den Monaten danach durch KMK-Arbeitsgruppen erarbeiteten Vorschläge zur Weiterentwicklung sind nach einem zweiten Fachgespräch am 10. Dezember 2014 weitgehend Konsens und sollen im Juni 2015 endgültig beschlossen werden.

Während die bisherige Gesamtstrategie als Instrument der Qualitätssicherung auf Ebene der Schule nur die Vergleichsarbeiten kannte, gibt es nunmehr die Absicht, in Ergänzung der Strategie auch die interne und externe Evaluation ausdrücklich als ein Element einzufügen. Während Vergleichsarbeiten an den ländergemeinsamen Bildungsstandards ausgerichtet sind und zu deren Implementation und Unterstützung einer kompetenzorientierten Unterrichtsentwicklung beitragen sollen, orientieren sich die externen Evaluationen an kriterialen Referenzen für Prozessstandards, die von den Ländern definiert werden.

Eine systematische Aufnahme der externen Evaluation in Korrespondenz zu den bisherigen Vergleichsarbeiten würde auf der Ebene der Einzelschule eine neue Balance zwischen Produkt- und Prozessqualität ermöglichen. Die KMK würde damit faktisch auch eine inhaltliche Verantwortung für die Weiterentwicklung der externen Evaluation übernehmen und die Möglichkeit für bundesweite Verabredungen schaffen, wie sie sie Anfang 2012 auch zur Handhabung des VERA-Instruments beschlossen hat. Die Schwierigkeiten, zwischen den Ländern zu einem Minimalkonsens zu kommen, sind absehbar, der Versuch ist jedoch lohnend und würde das Instrument der externen Evaluation und die in den Ländern daran arbeitenden Akteure eher stärken.

5. Schlussbemerkung

Die externe Evaluation als ein neues Element der Qualitätsentwicklung kann nach zehn Jahren auf einen insgesamt positiven Entwicklungsstand blicken. Die gelegentlich vorgetragene Fundamentalkritik der ersten Jahre, wonach die externe Evaluation in Deutschland ein unwissenschaftliches Verfahren zur Qualitätsbewertung von Schulen darstelle, das sich mit unzureichend ausgebildetem Personal, fragwürdigen Indikatorensets und mit unseriösen Methoden Urteile über Schule und Unterricht anmaßt, wird heute kaum noch vertreten. Im Gegensatz zu anderen Elementen der Reformpolitik der letzten zehn Jahre, wie nationalen und internationalen Schulleistungsstudien oder den bundesweiten Vergleichsarbeiten in den Jahrgangsstufen 3 und 8, hat sich die externe Evaluation eine deutlich bessere Akzeptanz bei den Akteuren auf Schulebene und in der interessierten Öffentlichkeit erworben.

Allerdings haben die Entwicklungen in den Ländern aufgrund der sehr unterschiedlichen rechtlichen, strukturellen und politischen Rahmenbedingungen, unter denen dieses Instrument eingeführt wurde, einen so unterschiedlichen Verlauf genommen, dass durchaus auch Grund zur Sorge bezüglich der weiteren Entwicklung in einzelnen Ländern besteht.

Ähnlich dem Vorwurf fehlender Steuerungsrelevanz für praktische Schulreformen, dem sich nationale oder internationale Schulleistungsstudien ausgesetzt sehen, wird vonseiten der Politik, ausgehend von überzogenen Hoffnungen auf unmittelbare Effekte solcher Verfahren auf die Qualitätssteigerung der Einzelschule, nach empirischen Belegen für die Nützlichkeit bzw. Wirksamkeit dieser Maßnahmen gefragt. Da die vorliegenden Forschungsbefunde zu Wirkungen der externen Evaluation darauf weder umfassend noch einheitlich zu Antworten kommen, könnte die Bildungspolitik ungeduldig werden und den Ausbau oder die Fortsetzung des eingeschlagenen Weges stoppen.

Auch wenn die Gefahr einer Abschaffung der externen Evaluation zum gegenwärtigen Zeitpunkt nicht gegeben ist, ist nicht auszuschließen, dass vor allem der personelle Aufwand, der für eine professionelle externe Evaluation betrieben werden muss, einigen Ländern zu teuer wird und sie anfällig werden für „billige“ Varianten, bei denen das anspruchsvolle Gesamtkonzept der externen Evaluation auf anlassbezogene Schulbesuche mit geringem Tiefgang zurückgefahren wird.

Ungeduld ist auf dem Gebiet der Bildungspolitik zwar nachvollziehbar, allerdings hat sich das Denken in Zyklen von vier- oder fünfjährigen Wahlperioden an keiner Stelle in der Qualitäts- und Reformdiskussion in Deutschland als sinnvoll erwiesen, da die angestrebten Veränderungen in der Systemarchitektur nicht unmittelbar kurzfristig auf die Qualitätsverbesserung in der Einzelschule durchschlagen. Es gibt eben keinen Hebel im Schulsystem, den man nur ergreifen und umlegen muss.

Lehrkräfte und Schulleitungen sind die zentralen Akteure in Schulen, die Qualität ihres professionellen Handelns ist entscheidend für den Organisationserfolg. Zugleich besteht in jeder Schule ein Netz aus Überzeugungen, Praktiken und Normen als konstitutives Element der jeweiligen Schulkultur, an der die Mitglieder der Organisation ihr individuelles und soziales Handeln orientieren. Auch die Wirkung externer Evaluationsformen wird entscheidend von der bereits vorhandenen Evaluationskultur beeinflusst, weniger von den konkreten Ergebnissen einer erfolgten externen Evaluation. Eine Schlüsselrolle kommt dabei der Bereitschaft der Lehrkräfte und Schulleitungen zu, evidenzfundierte Informationen und evidenzbasierte Steuerungsinstrumente aufzugreifen und handlungsleitend wirksam werden zu lassen.

Der Nutzen der Evaluationsberichte in Verbindung mit den zahlreichen anderen Datenquellen, die der Schulaufsicht zur Einschätzung der Qualität einer Schule zur Verfügung stehen, lässt sich noch erheblich steigern. Wenig genutzt wird gegenwärtig die Möglichkeit, die erfolgreich evaluierte Praxis von Schulen für andere Schulen bekannt und nutzbar zu machen. Die Chance, von erfolgreichen Schulen zu lernen, indem man zeitlich befristet regionale Netzwerke und Schulverbünde animiert, ist in einzelnen Ländern mit sehr unterschiedlichen Erfolgen erprobt worden. Allerdings benötigen die Mitarbeiterinnen und Mitarbeiter der Schulaufsicht dazu häufig eine zusätzliche Qualifizierung sowie Arbeitsstrukturen innerhalb ihrer Dienststellen, die eine gemeinsame und systematische Reflexion der vielfältigen vorliegenden Daten ermöglichen. Die Sorgfalt bei der (auch methodisch-empirischen) Weiterentwicklung und Verbesserung der Instrumente und Verfahren sollte in den nächsten Jahren nicht aufgegeben werden. Noch stärker sollte man sich aber darauf konzentrieren, die Übergabepunkte der Daten und Berichte von der externen Evaluation zur Schulaufsicht strukturell zu verankern. Dies wird nur auf der Basis institutioneller Zielvereinbarungen zwischen Schulaufsicht und den Dienststellen oder zuständigen Institutionen für externe Evaluation gelingen.

Insbesondere die mit hoher Stabilität identifizierten Schwächen von Schulen bedürfen, will man ihre Qualität ernsthaft verbessern, einer längerfristigen und umfassenden Unterstützung. Die gelegentlich feststellbare Praxis in Ländern, die externe Evaluation personell zulasten der operativen Schulaufsicht zu verstärken, wird nicht funktionieren. Mit der Implementation einer flächendeckenden Evaluation von Schulen nehmen die Aufgaben der Schulaufsicht eher zu und nicht ab. Das Verhältnis von externer Evaluation zur Schulaufsicht bedarf dringend einer Rollenklärung und sinnvollen Arbeitsteilung. Dazu gehört auch, dass durch gemeinsame Weiterbildungen ermöglicht werden muss, dass externe Evaluation und Schulaufsicht die gleichen Annahmen über Indikatoren für gute Schule teilen, wie sie in den meisten Ländern in sogenannten Handlungs- oder Orientierungsrahmen festgelegt sind.

Die Datenfülle der in den Schulinspektionen erhobenen Qualitätsmerkmale sollte dafür künftig noch zielgerichteter ausgewertet werden, um rechtzeitig zu

erkennen, wo Schwerpunkte im bildungspolitischen Steuerungshandeln gegenüber einer Schule erforderlich sind. Aus Sicht der Schulen mangelt es nach wie vor an adressatenspezifischen und nachfrageorientierten Unterstützungsangeboten, die langfristig angelegt sind.

Ein weiterer Schwerpunkt der kommenden Jahre sollte darin liegen, mit den Schulen intensiv daran zu arbeiten, dass ihre Selbstevaluationsfähigkeiten ausgebaut und die dazu erforderlichen Haltungen bei den Lehrkräften erworben werden. Erst damit würden die Grundlagen gelegt, eines Tages auch eine stärkere Verzahnung von interner und externer Evaluation zu erreichen und die Voraussetzung für Verfahren einer proportionalen Evaluation zu schaffen.

Bei allen offenen Problemen und unsicheren Perspektiven gibt es aber auch ermutigende Zeichen: Selbst in Schleswig-Holstein wächst nach einem politischen Wechsel an der Spitze des Bildungsressorts im Herbst 2014 die Einsicht, dass externe Evaluation ein sinnvolles Instrument zur Verbesserung der Qualität von Schule und Unterricht sein kann. So erklärte die neue Bildungsministerin Britta Ernst (SPD) nach ihren ersten hundert Tagen im Amt: „Was in Schleswig-Holstein jedoch fehlt, ist eine externe Evaluation. Derzeit ist Schleswig-Holstein das einzige Bundesland, in dem keine Evaluation der Schulen mehr stattfindet, nachdem das alte Verfahren EVIT 2009 abgeschafft worden ist. Doch es gibt eindeutig den Bedarf, denn externe Evaluation ist ein wertvolles Unterstützungsinstrument für die schulische Arbeit. Deshalb werden wir den Schulen noch in dieser Wahlperiode wieder ein Angebot machen zur externen Evaluierung auf freiwilliger Basis.“ (Pressemitteilung des Ministeriums für Schule und Berufsbildung vom 15. Dezember 2014).

Die externe Evaluation stellt ein neues und wichtiges Instrument zur Qualitätsentwicklung dar, das erheblich mehr Potenzial enthält, als bisher sichtbar geworden ist. Dieses lässt sich aber nur darstellen, wenn die erforderlichen Verknüpfungen mit anderen Elementen der Qualitätsentwicklung und -sicherung planvoll vorgenommen werden. Eine Gesamtstrategie in diesem Sinne gibt es bisher nur in Ansätzen. Sie zu entwickeln, bleibt auch in den nächsten Jahren eine Herausforderung an die Länder und die Kultusministerkonferenz als koordinierendes Gremium.

Literatur

- Abgeordnetenhaus Berlin (2012). Drucksache 17/10078 der 17. Wahlperiode. *Antwort des Senats zur Kleinen Anfrage vom 3. Januar 2012 über: Stand und Perspektiven der Schulinspektion*. Verfügbar unter: <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/17/KlAnfr/ka17-10078.pdf> [Zugriff am 19.02.2015]
- Avenarius, H. (2002). Welche Rechte und Pflichten haben Lehrkräfte, Schulleitungen und Schulaufsicht bei der Qualitätsentwicklung der Schulen und bei der Sicherung gleicher Lebenschancen? *Zeitschrift für Bildungsverwaltung (ZBV)*, 17 (2), 5–17.

- Avenarius, H. (2006). Die rechtlichen Regelungen der Länder zur erweiterten Selbständigkeit der Schulen. Vortrag vor der Landesschulrätekonzferenz der GEW Niedersachsen am 31. Oktober 2006 in Visselhövede-Jeddingen.
- Böhm-Kaspar, O. & Selders, O.(2013). „Schulinspektionen sollten regelmäßig durchgeführt werden?“ In I. van Ackeren, M. Heinrich & F. Thiel (Hrsg.), *Evidenzbasierte Steuerung im Bildungssystem?* (Die Deutsche Schule: Beiheft 12, S. 212–153). Münster: Waxmann.
- Brägger, G., Bucher, B. & Landwehr, N. (2005). Voraussetzungen für eine gute Praxis der externen Schulevaluation. In G. Brägger, B. Bucher & N. Landwehr (Hrsg.), *Schlüsselfragen zur externen Schulevaluation* (S. 21–50). Bern: hep-Verlag.
- Brüsemeister, T., Rürup, M., Böhm-Kasper, O., Heinrich, M. & Wissinger, J. (2009). *Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene*. Forschungsantrag an das BMBF im Schwerpunktprogramm „Steuerung im Bildungssystem“. Unveröffentlichtes Manuskript.
- Buhren, C.-G. & Düring, K. (2008). *Critical Friends – Schulentwicklung mit kritischen Freunden: Verfahren, Methoden und Wirkungen von Peer Reviews*. München: Oldenbourg.
- Burkard, C. (2005). Ergebnisorientierte Systemsteuerung: Konsequenzen für die externe Evaluation. In G. Brägger, B. Bucher & N. Landwehr (Hrsg.), *Schlüsselfragen zur externen Schulevaluation* (S. 79–109). Bern: hep-Verlag.
- Dedering, K. & Sowada, M. (2013). *Die Sicht der Externen. Abschlussbericht zum Projekt „Erfahrungen niedersächsischer Schulinspektor/innen“*. Universität Vechta: unveröffentlichtes Manuskript.
- Dedering, K., Sowada, M., Fritsch, N., Kierchhoff, A. & Weyer, C. (2014). Kompetent inspizieren? Was dazu nötig ist. *Schulverwaltung Niedersachsen*, 25 (4), 132–135.
- DeGEval – Gesellschaft für Evaluation e. V. (Hrsg.) (2008). *Empfehlungen für die Aus- und Weiterbildung in der Evaluation*. Verfügbar unter: http://www.degeval.de/fileadmin/Publikationen/Publikationen_Homepage/DeGEval_-_Empfehlungen_Aus-_und_Weiterbildung.pdf [Zugriff am 19.02.2015]
- Deutscher Städtetag (2012). *Bildung gemeinsam verantworten. Münchner Erklärung des Deutschen Städtetages anlässlich des Kongresses „Bildung gemeinsam verantworten“ am 8./9. November 2012*. Verfügbar unter: http://www.staedtetag.de/imperia/md/content/dst/muenchner_erklaerung_2012_final.pdf [Zugriff am 19.02.2015]
- Dix, A. (2011). Schulinspektion, Schulvisitation und Vergleichsuntersuchungen zwischen Datenschutz und Informationsfreiheit. *Zeitschrift für Bildungsverwaltung* 2/2011, 33–41.
- Döbert, H., Rürup, M. & Dedering, K. (2008). Externe Evaluation von Schulen in Deutschland – die Konzepte der Bundesländer, ihre Gemeinsamkeiten und Unterschiede. In H. Döbert & K. Dedering (Hrsg.), *Externe Evaluation von Schulen* (S. 63–152). Münster: Waxmann.
- Geißler, G. (2008). Geschichtliches zur „Vorwegnahme periodischer Untersuchungen“ der Schulen. In H. Döbert & K. Dedering (Hrsg.), *Externe Evaluation von Schulen* (S. 23–62). Münster: Waxmann.

- Gieske-Roland, M. (2015). Das Peer Review als Verfahren der Schulentwicklung. In C. G. Buhren (Hrsg.), *Handbuch Feedback in der Schule* (S. 108–125). Weinheim und Basel: Beltz.
- KMK (1997). *Grundsätzliche Überlegungen zu Leistungsvergleichen innerhalb der Bundesrepublik Deutschland – Konstanzer Beschluss* (Beschluss der Kultusministerkonferenz vom 24.10.1997). Verfügbar unter: www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1997/1997_10_24-Konstanzer-Beschluss.pdf [Zugriff am 10.7.2014]
- KMK (2001). *Presseerklärung vom 6.12.2001*. Verfügbar unter: www.kmk.org/presse-und-aktuelles/pm2001/296plenarsitzung.html [Zugriff am 16.6.2014]
- KMK (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. Neuwied: LinkLuchterhand.
- Landwehr, N. (2011). Thesen zur Wirkung und Wirksamkeit der externen Schulevaluation. In C. Quesel, V. Husfeldt, N. Landwehr & P. Steiner (Hrsg.), *Wirkungen und Wirksamkeit der externen Schulevaluation* (S. 35–70). Bern: hep-Verlag.
- Maritzen, N. (2007). Schulinspektion – ein neues Element der Systemsteuerung. *Journal für Schulentwicklung*, 11 (3), 6–14.
- Ministerium für Schule und Berufsbildung Schleswig-Holstein (2014). *Pressemitteilung vom 15.12.2014*. Verfügbar unter: http://www.schleswig-holstein.de/MSB/DE/Service/Presse/PI/2014/Dezember_2014/III_100Tage.html [Zugriff am 19.02.2015]
- Mühlbauer, K. (2010). Bedeutsamkeit der Auswahl, Ausbildung und Betreuung der Evaluatoren durch die Qualitätsagentur. In B. Schaal & F. Huber (Hrsg.), *Qualitätssicherung im Bildungswesen* (S. 185–196). Münster: Waxmann.
- Nikolai, R. & Rothe, K. (2013). Konvergenz in der Schulstruktur? Programmatik von CDU und SPD im Vergleich. *Zeitschrift für Politikwissenschaft*, 23 (4), 545–572.
- Preuß, B. (2013). Akteurskonstellation zwischen Schulträger und Schule. In I. van Ackeren, M. Heinrich & F. Thiel (Hrsg.), *Evidenzbasierte Steuerung im Bildungssystem?* (Die Deutsche Schule: Beiheft 12, S. 154–171). Münster: Waxmann.
- Rürup, M. (2008). Typen der Schulinspektion in den deutschen Bundesländern. *Die Deutsche Schule*, 100 (4), 467–477.
- Stevahn, L., King, J. A., Ghore, G. & Minnema, J. (2005). Establishing Essential Competencies for Program Evaluators. *American Journal of Evaluation*, 26 (1), 43–59.
- Stryck, T. (2000). Qualitätssicherung in der Geisterbahn. Was hat Schulaufsicht mit Schulqualität zu tun? In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich; Schule, Sozialpädagogik, Hochschule*. (Zeitschrift für Pädagogik: Beiheft 41, S. 111–125). Weinheim: Beltz.
- Stryck, T. (2014). Schulinspektion – ein notwendiger, weil unbequemer Steuerungsakteur. In Senatsverwaltung für Bildung, Jugend und Sport (Hrsg.), *7 Jahre Schulinspektion in Berlin* (S. 7–9). Berlin: Eigendruck der Pressestelle.

Der zweite Zyklus der Schulinspektion Hamburg: ein Ausblick

Martina Diedrich

Zusammenfassung

Der Beitrag stellt das Verfahren der Hamburger Schulinspektion in ihrem zweiten Zyklus dar. Dazu werden zunächst die Bewertungsgrundlagen auf Basis des neuen Orientierungsrahmens Schulqualität erläutert. Davon ausgehend werden die getroffene Schwerpunktbildung und die notwendig gewordene Reduktion der Bewertungskriterien erklärt. Neben der ausführlichen Darstellung der Instrumente und des Verfahrens geht der Beitrag vertiefend auf das gegenüber dem ersten Zyklus veränderte Berichtsformat ein. Neu daran ist insbesondere die Veröffentlichung eines Berichtsteils im Internet, die eine Herausforderung an die Allgemeinverständlichkeit und die Reichweite der Texte stellt. Abschließend wird das veränderte Selbstverständnis der Schulinspektion diskutiert: Während im ersten Zyklus Sachlichkeit, die Bezugnahme auf objektivierbare Kriterien und die klare Rollenabgrenzung gegenüber den Schnittstellenpartnern dominierten, steht im zweiten Zyklus der Gedanke des Deutungsangebots an die Schulen im Vordergrund, welches im wechselseitigen Austausch vermittelt wird und die Revidierbarkeit von Urteilen einschließt.

Im Frühjahr 2013 hat die Schulinspektion Hamburg ihren ersten Zyklus nach 375 durchgeführten Inspektionen abgeschlossen. Parallel dazu hat sie im Jahr 2012 den zweiten Zyklus vorbereitet und pilotiert. Seit März 2013 werden nun alle Schulen nach dem Verfahren des zweiten Zyklus inspiziert. Nachfolgend werden die tief greifenden Veränderungen in den Bewertungsgrundlagen und die Veränderungen sowohl der normativen Bezugsquellen als auch des Verfahrens selbst dargestellt und mit Blick auf die bestehenden Herausforderungen diskutiert.

1. Bewertungsgrundlagen

Der grundlegende Auftrag der Schulinspektion hat sich auch im zweiten Zyklus nicht verändert: Nach wie vor obliegt es ihr, alle staatlichen Hamburger Schulen nach einem definierten Verfahren extern zu evaluieren. Dabei erfüllt sie weiterhin die vier Funktionen der Katalyse (Ermöglichung von Entwicklung), der Gewährleistung, der Erkenntnisgenerierung und der Normendurchsetzung (vgl. den Bei-

trag „Aufbau und Rolle der Schulinspektion Hamburg“ in diesem Band). Was sich jedoch verändert hat, sind die Bewertungsgrundlagen, auf die sich die Inspektion bei ihrem Vorgehen stützt. Damit war die Schulinspektion nach Abschluss des ersten Zyklus vor die Aufgabe gestellt, ihr gesamtes Instrumentarium anzupassen, da veränderte normative Grundlagen operativ anders umgesetzt werden müssen. Die Änderungen betrafen insbesondere den Orientierungsrahmen Schulqualität als Beschreibung dessen, was in Hamburg unter einer „guten Schule“ verstanden werden soll, aber auch die politischen Schwerpunktsetzungen, die die Schulinspektion zur vertieften Betrachtung ausgewählter Themenbereiche anhalten.

1.1 Neuer Orientierungsrahmen Schulqualität

Der erste Orientierungsrahmen Schulqualität ist im Jahr 2005 im Zusammenhang mit der erweiterten Selbstverantwortung von Schulen entstanden und 2006 der Drucksache zur Änderung des Hamburger Schulgesetzes, in deren Rahmen auch die Schulinspektion eingeführt wurde, als Anhang beigelegt worden (Drs. 18/3780). Seitdem hat sich die Hamburger Schullandschaft in vielerlei Hinsicht deutlich gewandelt: Neben strukturellen Entwicklungen im Rahmen der Hamburger Schulreform 2010¹ haben Themen wie Kompetenzorientierung, Ganzttag, Inklusion oder die veränderte Rolle der Schulleitung die Schulen geprägt. Unabhängig davon, dass es Konstanten für gute Schulen und guten Unterricht gibt, stellt sich die Frage, wie solche politisch initiierten Veränderungen nachhaltig in den grundlegenden normativen Regelwerken abgebildet werden. Um den Orientierungsrahmen in Einklang mit den prioritären Entwicklungsaufgaben der Schulen zu bringen und um die Erfahrungen, die im ersten Inspektionszyklus gesammelt wurden, angemessen berücksichtigen zu können, ist deshalb in den Jahren 2010 bis 2012 in einem umfassenden Prozess unter Beteiligung von Behördenvertreterinnen und -vertretern, von Schulaufsichtsbeamtinnen und -beamten, Vertreterinnen und Vertretern des Landesinstitutes, ausgewählten Schulleitungen und der Schulinspektion ein neuer Orientierungsrahmen Schulqualität erarbeitet worden, dem unverändert folgendes Grundverständnis zugrunde liegt:

- Schulqualität hat Kontext-, Input-, Prozess- und Ergebnismerkmale. Der Orientierungsrahmen leitet dazu an,
 - Zusammenhänge zu verstehen und zu gestalten,
 - Entwicklungsschwerpunkte einzuordnen.

1 Im Rahmen dieser Reform wurde u. a. ein Zweisäulenmodell in der Sekundarstufe I und II eingeführt, bestehend aus Stadtteilschule (Abitur nach neun Jahren) und Gymnasium (Abitur nach acht Jahren).

- Der Orientierungsrahmen ist mehrperspektivisch angelegt. Es muss deutlich sein, wer welche Erwartungen an die Qualität der Schule hat.
- Der Orientierungsrahmen setzt bei der Gewährleistungsverantwortung der Schulleitung an und versteht Qualitätsentwicklung als eine Leitungsaufgabe.
- Der Orientierungsrahmen ist ein Verständigungsmittel, das Qualitätsentwicklungen in den Schulen initiiert und anleitet. Schulqualität muss deshalb Schritt für Schritt konkretisiert werden.
- Der Orientierungsrahmen gibt Bezugspunkte für interne und externe Evaluationen, denn Schulqualität muss anhand transparenter Kriterien nachgewiesen werden.

Abbildung 1: Qualitätsdimensionen und -bereiche des zweiten Orientierungsrahmens
Schulqualität

Dimension 1: Führung und Management	Dimension 2: Bildung und Erziehung	Dimension 3: Wirkungen und Ergebnisse
1.1 Führung wahrnehmen	2.1 Lehr- und Lernprozesse gestalten	3.1 Kompetenzen
1.2 Die Entwicklung der Schule und des Lernens steuern	2.2 Zusammenarbeiten	3.2 Bildungslaufbahnen und Schulabschlüsse
1.3 Qualitätsmanagement etablieren	2.3 Zusätzliche Förderung gewährleisten	3.3 Anschlüsse und nachhaltige Wirkungen
1.4 Verantwortung für das Personal wahrnehmen	2.4 Erziehungsprozesse gestalten	3.4 Einverständnis und Akzeptanz
1.5 Die Finanz- und Sachmittel gezielt einsetzen	2.5 Lernentwicklung begleiten und Leistungen beurteilen	3.5 Arbeitszufriedenheit der Mitarbeiterinnen und Mitarbeiter
1.6 Kommunikation nach innen und außen gewährleisten	2.6 Den Unterricht kontinuierlich weiterentwickeln	
1.7 Regionale Bezüge ermöglichen und die Vernetzung sichern	2.7 Beratungsangebote gestalten	
1.8 Organisatorische Rahmenbedingungen sichern	2.8 Die Schulgemeinschaft beteiligen	
	2.9 Regionale Vernetzung gestalten	

Eine Herausforderung beim Schreiben des neuen Orientierungsrahmens bestand darin, trotz vielfältiger Neuerungen eine Kontinuität zur ersten Fassung herzustellen. Dies ist durch die gewählte Struktur erfolgt: So ist die Unterscheidung in die drei Dimensionen „Führung und Management“, „Bildung und Erziehung“ sowie „Wirkungen und Ergebnisse“ erhalten geblieben. Innerhalb dieser Qualitätsdimensionen werden nun jedoch weitaus mehr Qualitätsbereiche unterschieden, nämlich 22 statt bislang 14. Abbildung 1 gibt die Struktur wieder. Die Differenzierungen im Verhältnis zur ersten Fassung ergeben sich zunächst vor allem in der Leitungsdimension. Hier ist insbesondere die Verantwortung für die Steuerung der schulischen Qualitätsentwicklung stärker konturiert. Insgesamt ist als neues Thema die Bedeutung regionaler Vernetzung in Erscheinung getreten, darüber hinaus wird den Aspekten Erziehung und Zusammenarbeiten mehr Gewicht verliehen. Weniger offensichtlich, weil nicht als eigener Qualitätsbereich ausgewiesen, sind die Themen ganztätige Bildung und Inklusion. Diese sind jedoch als durchgängige Gestaltungsprinzipien an verschiedenen Stellen verortet: Gute Schulen werden als ganztägige, inklusive Schulen konzipiert.

Alle 22 Qualitätsbereiche werden durch zahlreiche Qualitätsmerkmale beschrieben. Insgesamt ist der neue Orientierungsrahmen im Verhältnis zum ersten aufgrund der weiteren Ausdifferenzierung sehr viel umfangreicher geworden. Zudem hat er durch die Unverbundenheit der einzelnen Qualitätsbereiche einen eher aufzählenden Charakter. Um eindeutige Schwerpunktsetzungen zu ermöglichen und um innere Zusammenhänge der einzelnen Qualitätsbereiche zu verdeutlichen, ist er durch einen sogenannten *Leitfaden* ergänzt worden. Der Leitfaden erfüllt drei Funktionen:

- Er soll *fokussieren* und *zusammenfassen*. Dadurch dient er vor allem der Kommunikation mit einer interessierten Öffentlichkeit, die sich rasch über Aspekte guter Schulen informieren will.
- Er *identifiziert* die wesentlichen *Kriterien*, an denen sich Schule messen lassen muss, nämlich Kompetenzerwerb und guter Unterricht. Schule ist nämlich darauf ausgerichtet, dass Schülerinnen und Schüler Wissen erwerben und Kompetenzen aufbauen, und dieses Ziel wird vor allem durch guten Unterricht erreicht. Alle anderen Qualitätsmerkmale müssen sich auf diese beiden Kriterien beziehen.
- Er stellt ein *Beziehungsgefüge* zur Systematisierung der Qualitätsbereiche her. Die unverbundene Aufzählung der einzelnen Qualitätsbereiche wird in einen inneren Zusammenhang gebracht, sodass ihr Einfluss auf die zentralen Kriterien „Kompetenzerwerb“ und „guter Unterricht“ sichtbar wird.

Abbildung 2: Beziehungsgefüge der Qualitätsbereiche des Orientierungsrahmens Schulqualität



Der Leitfaden signalisiert so gegenüber den Schulen, dass manche Qualitätsbereiche auf direktem Wege zum Ziel führen, während andere eher einen indirekten Einfluss auf die Zielerreichung haben. Damit sind Wege der Schulentwicklung zwar in einem gewissen Maße vorgezeichnet, die Schulen werden jedoch in der Gestaltung ihrer Entwicklungsprozesse nicht beschnitten. Um es an einem Beispiel zu konkretisieren: Durch die Stärkung der regionalen Vernetzung im Stadtteil wird vermutlich keine direkte Verbesserung der Unterrichtsqualität erreicht. Gleichwohl kann eine Schule aber hier einen Einstieg in die Qualitätsentwicklung wählen, wenn es ihr – beispielsweise aufgrund einer herausfordernden Schülerklientel – da-

rum geht, strategische Partner für eine Veränderung der schulischen Kulturen und die Gestaltung außerschulischer Lernsettings zu gewinnen.

Die hier skizzierte Schwerpunktsetzung hat erhebliche Implikationen für die Schulinspektion, denn sie muss sich in den Instrumenten niederschlagen, die bei den Inspektionen eingesetzt werden. Dementsprechend bilden der Unterricht und die Unterrichtsentwicklung neben dem Steuerungshandeln die Schwerpunkte der Ausrichtung des zweiten Zyklus. Insbesondere die Unterrichtsentwicklung erlangt im Verhältnis zum ersten Zyklus eine stärkere Bedeutung. Aufgrund der starken Betonung des Kompetenzerwerbs der Schülerinnen und Schüler wäre es folgerichtig, auch diesen Bereich verstärkt bei den Schulinspektionen zu berücksichtigen. Hierzu bedarf es jedoch entsprechender Daten, die eine wertende Aussage über die Qualität der Schule mit Blick auf das Lernen der Schülerinnen und Schüler zulassen. Diese Daten liegen jedoch gegenwärtig noch nicht vor. Leistungsaspekte können frühestens dann einbezogen werden, wenn bei den Auswertungen der Hamburger Vergleichsarbeiten KERMIT (Kompetenzen Ermitteln) Aussagen zu Kompetenzentwicklungen getroffen werden können.

1.2 Bildungspolitische Schwerpunktsetzungen

Neben dem Auftrag, die Qualität der Einzelschule zu bewerten, leistet die Schulinspektion auch einen Beitrag zum Systemmonitoring, indem sie auf aggregierter Ebene Entwicklungsnotwendigkeiten identifiziert und den Steuerungsinstanzen so Hinweise auf Handlungserfordernisse gibt. Damit richtet sie sich auch und vor allem an die politische und administrative Steuerungsebene, deren Anliegen es ist, Rückmeldungen zum aktuellen Entwicklungsstand des Schulsystems zu erhalten. Da von ihr weitgehende Impulse zur Veränderung und Weiterentwicklung ausgehen, liegt es nahe, das Schulsystem eben mit Blick auf genau solche Schwerpunktsetzungen zu analysieren, von denen sich die politischen und administrativen „Steuerleute“ eine Impulswirkung versprechen. Vor diesem Hintergrund hat es sich die Schulinspektion im zweiten Zyklus zur Aufgabe gemacht, folgende drei Themen der aktuellen schulpolitischen Diskussion aufzugreifen und operativ umzusetzen: 1. Inklusion, 2. ganztägige Bildung, 3. Berufs- und Studienorientierung. Bei der übergreifenden Berichterstattung wird jeweils jahresweise einer dieser Bereiche vertiefend betrachtet und analysiert.

1.3 Empirische Fundierung

Die Schulinspektion ist ein wissenschaftlich fundiertes, auf empirischen Methoden basierendes Verfahren (vgl. dazu auch den Beitrag „Aufbau und Rolle der Schul-

inspektion Hamburg“ in diesem Band). Hierzu gehört, die Vorgaben des Orientierungsrahmens und der politischen und administrativen Steuerungsebene mit den Befunden der empirischen Bildungsforschung abzugleichen. Ob der Fülle der Qualitätsbereiche und -merkmale des Orientierungsrahmens neben der Fokussierung, die durch den Leitfaden entsteht, muss eine Auswahl und Engführung getroffen werden, die der Leitfaden allein nicht leisten kann. Diese Auswahl kann in der Rationale der Schulinspektion sinnvoll nur empirisch begründet werden. Dementsprechend hat die Schulinspektion im zweiten Zyklus – wie auch schon im ersten – Befunde der Schuleffektivitäts- und Schulqualitätsforschung herangezogen, um diejenigen Bereiche zu identifizieren, die einer unbedingten Aufmerksamkeit bedürfen (vgl. Scheerens & Bosker 1997; Teddlie & Stringfield 2003; Scheerens et al. 2007).

Abbildung 3: Schwerpunkte und Qualitätsbereiche im zweiten Zyklus der Hamburger Schulinspektion

Steuerungs- handeln	<ul style="list-style-type: none"> • Führung wahrnehmen • Die Entwicklung der Schule & des Lernens steuern / Qualitätsmanagement etablieren • Verantwortung für das Personal übernehmen • Organisatorische Rahmenbedingungen sichern • Die Schulgemeinschaft beteiligen
Unterrichtsentwicklung: Zusammenarbeiten	<ul style="list-style-type: none"> • Zusammenarbeiten
Unterrichtsentwicklung: Feedback	<ul style="list-style-type: none"> • Lernentwicklung begleiten und Leistung beurteilen • Den Unterricht kontinuierlich weiterentwickeln
Unterrichts- qualität	<ul style="list-style-type: none"> • Die Lehr- und Lernprozesse gestalten • Zusätzliche Förderung gewährleisten • Erziehungsprozesse gestalten
Wirkungen und Ergebnisse	<ul style="list-style-type: none"> • Bildungslaufbahnen und Schulabschlüsse / Anschlüsse und nachhaltige Wirkungen • Einverständnis und Akzeptanz • Arbeitszufriedenheit der Mitarbeiterinnen und Mitarbeiter

Die skizzierten Erwägungen haben dazu geführt, dass die Schulinspektion im zweiten Zyklus die in der Abbildung 3 aufgeführten Qualitätsaspekte betrachtet: 1. Steuerungshandeln, 2. Unterrichtsentwicklung durch Zusammenarbeit, 3. Un-

terrichtsentwicklung durch Feedback, 4. Unterrichtsqualität, 5. Wirkungen und Ergebnisse. Diese werden durch insgesamt 13 Qualitätsbereiche beschrieben (vgl. Abbildung 3). Der Qualitätsbereich „Bildungslaufbahnen und Abschlüsse/Wirkungen und nachhaltige Ergebnisse“ kann – wie oben ausgeführt – erst bei Vorliegen entsprechender Daten bewertet werden. Damit bilden der Unterricht und die Maßnahmen zur Unterrichtsentwicklung den Schwerpunkt der Bewertungen.

Daneben wird jährlich ein bildungspolitischer Schwerpunkt vertiefend betrachtet. Im Schuljahr 2014/15 wird der Bereich Inklusion/Umgang mit Heterogenität aufgenommen. Ziel dabei ist, den Schulen eine Rückmeldung dazu zu geben, wie sie in diesem Bereich „aufgestellt“ sind und welche Handlungserfordernisse aus Sicht der Schulinspektion bestehen.

2. Instrumentierung

Die oben skizzierten Bewertungsgrundlagen galt es, in angemessene Erhebungsinstrumente zu überführen. Ihrem Grundprinzip der Triangulation (vgl. „Aufbau und Rolle der Schulinspektion Hamburg“ in diesem Band) ist die Schulinspektion dabei treu geblieben. So bemüht sie sich weiterhin, ihre Bewertungen auf mehrere Datenarten aus unterschiedlichen Datenquellen zu stützen. Konkret bedeutet dies, dass weiterhin Dokumente hinzugezogen, mündliche und schriftliche Befragungen aller schulischen Gruppen sowie Unterrichtsbeobachtungen durchgeführt werden.

2.1 Dokumentenanalyse

Im Bereich der schulischen Programmatik werden weiterhin ähnliche Dokumente wie im ersten Zyklus betrachtet. Sie dienen jedoch vor allem als Hintergrundwissen und werden nicht mehr schematisch ausgewertet. Weiterhin werden das Leitbild und, sofern aus Sicht der Schulleitung für den aktuellen Schulalltag noch von Relevanz, das Schulprogramm angeschaut, ferner die Dokumente zur Fortbildungsplanung, die Ziel- und Leistungsvereinbarungen zwischen Schule und Behörde sowie einige schulinterne Curricula. Letztere werden vor allem unter dem Blickwinkel betrachtet, inwiefern verbindliche Maßnahmen zu ihrer Umsetzung getroffen werden und es zwischen Fächern vergleichbare Standards gibt. Neu ist die Betrachtung des Förderkonzepts (bzw. von Dokumenten, die die Förderung dokumentieren) und des Ganztagskonzepts. Zusätzlich werden nun vor Ort Berufswahlpässe und Förderpläne eingesehen, um einen vertiefenden Einblick in die Aspekte Berufs- und Studienorientierung sowie Förderung nehmen zu können.

Ein neues Dokument, das ausschließlich für die Schulinspektion erstellt wird, ist der sogenannte Entwicklungsbericht. In ihm legen die Schulleiterinnen und

Schulleiter da, welche Veränderungen es seit der ersten Inspektion gegeben hat und welche Entwicklungsschritte bzw. Maßnahmen die Schule eingeleitet und umgesetzt hat. Damit erhält die Schulinspektion bereits vor ihrem Schulbesuch einen vertieften Einblick in die Entwicklung der Schule in den vergangenen Jahren aus Sicht der Schulleitung.

2.2 Schriftliche Befragungen

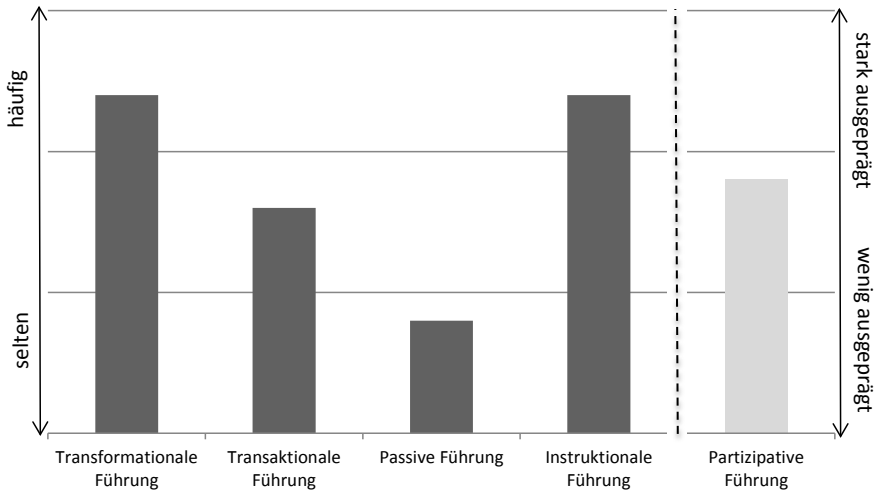
Schülerinnen und Schüler, Pädagoginnen und Pädagogen, Eltern und an den berufsbildenden Schulen die Ausbildungsbetriebe werden weiterhin schriftlich befragt. Neu ist die Einbeziehung aller pädagogisch tätigen Personen (also beispielsweise auch der Erzieherinnen und Erzieher oder der therapeutisch Tätigen). Die Änderungen in diesem Fragebogen beruhen nahezu ausschließlich auf etablierten Instrumenten, die sich in Studien der empirischen Bildungsforschung bewährt haben. Ziel der Überarbeitung und Ergänzung des Fragebogens ist es, empirisch abgesicherte Verfahren einzusetzen, die neben ihrer wissenschaftlichen Güte insbesondere auch einen Vergleich mit existierenden Befunden zulassen, was im ersten Zyklus so nicht möglich gewesen ist. Dementsprechend erfolgt die Rückmeldung nicht mehr auf der Ebene einzelner Items, sondern auf der Ebene von Konstrukten, also durch mehrere Items beschriebene Merkmalsbündel. Trotz einer dadurch methodisch verbesserten Vorgehensweise handelt sich die Schulinspektion mit diesem Vorgehen gleichwohl die Problematik ein, dass Konstrukte – beispielsweise im Bereich des Schulleitungshandelns, s. u. – in der Regel auf einem spezifischen theoretischen Hintergrund basieren, den es in der Rückmeldung mitzuteilen gilt, damit die Ergebnisse nachvollzogen werden können. Gleichzeitig entzieht sich die Schulinspektion dadurch aber ein Stück weit der Anschaulichkeit und intuitiven Plausibilität. Diese (Opportunitäts-), „Kosten“ werden aber in Anbetracht eines methodisch seriösen Vorgehens in Kauf genommen. Die Berichte enthalten deshalb weiterführende Erläuterungen zu den Konstrukten. Selbstverständlich können die Schulen aber auch einen Überblick auf der Ebene der einzelnen Items bekommen, wenn sie dies wünschen.

Inhaltlich legt der Fragebogen einen deutlichen Schwerpunkt auf die Wahrnehmung des Steuerungshandelns in der Schule. Damit folgt die Schulinspektion Befunden, die die herausgehobene Stellung der Schulleitung für die Qualitätsentwicklung der Schule belegen (vgl. Huber 2005; Pietsch 2014a, 2014b; Scheerens 2012). Aspekte guten Führungshandelns sind dabei eine aktive, motivierende Personalführung, die Ausrichtung des Leitungshandelns auf die gezielte Verbesserung von Lehr- und Lernbedingungen und die Verteilung von Führungsverantwortung und -aufgaben im Kollegium. Um ein solches Führungshandeln zu beschreiben, werden fünf Führungsstile unterschieden und mithilfe etablierter Befragungsinstru-

mente erhoben (vgl. Hallinger 2003; Huber 2005). Für eine „gute“ Führung bedarf es einer situativ unterschiedlich ausgeprägten Mischung dieser unterschiedlichen Stile (s. Abbildung 4):

- *Transformationale Führung*: Die Schulleitung vertritt gegenüber ihren Mitarbeiterinnen und Mitarbeitern eine langfristig ausgerichtete, sinnstiftende Vision für die gesamte Organisation, die als Ordnungsprinzip für alle Mitarbeiterinnen und Mitarbeiter gilt. Darüber hinaus bietet sie Mitwirkungsmöglichkeiten; die Mitarbeiterinnen und Mitarbeiter werden aktiv in die Umsetzung und Erreichung von Visionen einbezogen. Die Schulleitung unterstützt die Mitarbeiterinnen und Mitarbeiter entsprechend aktiv darin, an Veränderungen zu partizipieren.
- *Transaktionale Führung*: Die Schulleitung befindet sich mit ihren Mitarbeiterinnen und Mitarbeitern in einer vergleichsweise hierarchisch organisierten Austauschbeziehung. Erwartungen und Ziele werden klar formuliert, Prozesse und Zielrealisierungen kontrolliert und durch die Schulleitung belohnt oder, bei Abweichung von Regeln bzw. Nichterreichen, gegebenenfalls sanktioniert.
- *Passive Führung*: Die Schulleitung verhält sich passiv und greift in der Regel nicht oder zu spät in das Geschehen ein; sie verzichtet auf ihren Führungsanspruch und kommt ihrer Verantwortung als Schulleitung nicht nach.
- *Instruktionale Führung*: Die Schulleitung kontrolliert und koordiniert Aspekte des Schul- und Unterrichtsgeschehens, die den Lernfortschritt der Schülerinnen und Schüler betreffen, und nimmt entsprechend direkten Einfluss auf den Unterricht und das Curriculum.
- *Partizipative Führung*: Die Schulleitung führt die Schule kooperativ und/oder in geteilter Verantwortung und richtet dafür Strukturen ein, die eine breite Aufteilung der Führungsaufgaben und der Führungsverantwortung innerhalb der Schule garantieren.

Abbildung 4: Profil einer idealtypischen Führung von Schule



Der Elternfragebogen und der Fragebogen für die Schülerinnen und Schüler wurden gegenüber den früheren Fassungen deutlich gekürzt. Damit hat die Schulinspektion auf wiederholte Kritik der Schulen und der Eltern im ersten Zyklus reagiert, dass die Instrumente die Befragten angesichts ihrer Länge und Komplexität überforderten.

2.3 Mündliche Befragungen

Alle schulischen Gruppen werden weiterhin zu den einzelnen Qualitätsbereichen und -merkmalen mündlich befragt. Dabei wird der Befragung der Schulleitung und der Pädagoginnen und Pädagogen zeitlich ein größeres Gewicht als im ersten Zyklus eingeräumt. Anders als im ersten Zyklus wird nun der Tatsache Rechnung getragen, dass insbesondere in größeren Schulen Leitung vielfach im Team wahrgenommen wird, weshalb jetzt die gesamte erweiterte Schulleitung von Anfang an am Interview teilnimmt. Thematisch wird vor allem die Steuerungsdimension behandelt, wobei Fragen der Verantwortungsdelegation, Aufgabenwahrnehmung, Qualitätssorge, Schulentwicklung, Personalführung und Unterrichtsentwicklung im Zentrum stehen. Entsprechend werden in den übrigen Interviews die für die jeweilige Befragtengruppe dominanten Themen bearbeitet.

2.4 Unterrichtsbeobachtungen

Die Schulinspektion hält auch im zweiten Zyklus daran fest, in Abhängigkeit von der Schulgröße zwischen 40 und 100 Unterrichtssequenzen pro Schulbesuch zu beobachten (vgl. „Aufbau und Rolle der Schulinspektion Hamburg“ in diesem Band). Dabei ist auch der Unterrichtsbeobachtungsbogen gleich geblieben, da er sich bewährt hat und wissenschaftlich umfänglich fundiert wurde (vgl. Pietsch 2008, 2010). Verändert hat sich jedoch die Rückmeldung. Anders als noch im ersten Zyklus wird von einer Rückmeldung hierarchisch aufeinander aufbauender Stufen der Unterrichtsqualität abgesehen. Stattdessen werden die 30 Items 4 Bereichen zugeordnet, die gleichwertig nebeneinander gestellt werden:

1. Lernklima und Pädagogische Strukturen,
2. Klassenführung und Methodenvariation,
3. Schülermotivation, aktives Lernen und Wissenstransfer,
4. Differenzierung, wirkungs- und kompetenzorientierte Förderung.

Damit erhofft sich die Inspektion einen Akzeptanzgewinn, da die Schulen im ersten Zyklus häufig Kritik an der hierarchischen Darstellung geäußert haben.

2.5 Zum Verfahren an den berufsbildenden Schulen und an den Sonderschulen

Die berufsbildenden Schulen werden prinzipiell nach demselben Verfahren und mithilfe derselben Instrumente inspiziert wie die allgemeinbildenden Schulen. Der Orientierungsrahmen Schulqualität gilt für sie gleichermaßen. Allerdings gibt es einige Spezifika insbesondere in Bezug auf das Qualitätsmanagement, die auch die Schulinspektion angemessen berücksichtigen muss: Das Hamburger Institut für Berufliche Bildung (HIBB) hat 2007 ein Rahmenkonzept zum Qualitätsmanagement erlassen, das die Schulen zur Einführung eines umfassenden Qualitätsmanagementsystems verpflichtet, zu dem u. a. ein Qualitätshandbuch, die Vorkehr für umfängliche Feedbackmaßnahmen oder die Einrichtung einer Steuergruppe gehören. Die Schulinspektion achtet deshalb in verstärktem Maße auf den Entwicklungsstand in diesem Bereich, indem sie beispielsweise das Qualitätshandbuch in die Dokumentenanalyse einbezieht. Darüber hinaus berücksichtigt sie die betriebliche Sicht, unter anderem im Bereich Feedback, in der Lernbegleitung oder bezüglich der Einbindung von Schülerinnen und Schülern mit besonderem Förderbedarf. Im Sinne eines Feldversuchs erprobt die Inspektion in den berufsbildenden Schulen zusätzlich die Evaluation eines von der Schule selbst definierten schulspezifischen Entwicklungsprojekts. Das Thema Change Management wird

anhand dieses Schwerpunkts behandelt. Ziel ist eine stärker auf die Bedürfnisse bzw. Interessen der Schule ausgerichtete Vorgehensweise und eine Stärkung des dialogischen Charakters der Inspektion.

Auch organisatorisch gibt es bei den berufsbildenden Schulen eine Besonderheit: Aufgrund des neuen Schulentwicklungsplans sollen Schulen, die zusammengelegt werden, frühestens ein Jahr nach der Fusion inspiziert werden, es sei denn, sie äußern explizit einen anderen Wunsch. Ferner werden alle Schulen bereits ein Jahr vor der Inspektion über den Inspektionstermin informiert, um ihnen die Möglichkeit zu geben, sich hierauf vorzubereiten. Sofern sich dieses Vorgehen bewährt, soll es auf die allgemeinbildenden Schulen übertragen werden.

Bei den Sonderschulen ist gegenüber den allgemeinen Schulen beispielsweise der Aspekt Beratung als Qualitätsbereich zusätzlich aufgenommen worden, weil ihm in dieser Schulform eine besondere Bedeutung zukommt. Das Verfahren für die Inspektion der Regionalen Bildungs- und Beratungszentren, an denen ehemalige Einrichtungen der pädagogischen und psychologischen Beratung (Rebus) mit den Förder- und Sprachheilschulen an 13 Standorten zusammengelegt worden sind, befindet sich noch in der Entwicklung.

3. Berichtslegung

Mit dem zweiten Zyklus ist eine entscheidende Veränderung in der Berichtslegung der Schulinspektion eingeführt worden: Aufgrund hohen öffentlichen und politischen Interesses werden die Berichte nun einer breiten Öffentlichkeit zur Verfügung gestellt und im Internet publiziert. Die Schulinspektion war damit vor die Herausforderung gestellt, aussagekräftige Berichte zu verfassen, die den Schulen sowohl eine in Bezug auf die Schulentwicklung hilfreiche Rückmeldung geben als auch die schutzwürdigen Interessen der Beteiligten wahren. Den Berichten vorangestellt ist nun eine drei- bis vierseitige Zusammenfassung, in der die oben aufgeführten Bewertungsschwerpunkte möglichst allgemeinverständlich beschrieben werden. Ziel ist es, damit den Schulentwicklungsstand der Schule zusammenhängend zu beschreiben und somit gleichsam den roten Faden vorzugeben, anhand dessen innere Zusammenhänge der Schulentwicklung erkennbar werden. Um Stärken und Schwächen eindeutig identifizieren zu können, werden die Bewertungen der einzelnen Qualitätsbereiche pro Qualitätsschwerpunkt mit Zusatztexten ergänzt. Auf eine tabellarische bzw. grafische Darstellung wird in diesem öffentlichen Teil verzichtet, um einfaches Auszählen der Berichte mit dem Ziel eines Rankings zu erschweren. Dieser zusammenfassende Teil des Berichts wird acht Wochen nach der Rückmeldung an die Schulen im Internet unter der Adresse: <http://www.hamburg.de/inspektionsberichte/> veröffentlicht. Die Schulen haben die Möglichkeit, eine schriftliche Stellungnahme zum Bericht abzugeben

und diese an derselben Stelle zu veröffentlichen. Bislang wurde diese Möglichkeit allerdings kaum genutzt.

Die Schule erhält darüber hinaus einen sehr viel umfänglicheren Bericht, der zunächst einige Hinweise zum Bewertungshintergrund und zum Umgang mit den Daten gibt. Im Anschluss an die veröffentlichte schriftliche Zusammenfassung wird das Qualitätsprofil der Schule grafisch dargestellt. Danach wird für die einzelnen Qualitätsbereiche tabellarisch dargestellt, wie die Inspektion die dem Bereich jeweils zugehörigen Qualitätsaspekte eingeschätzt hat. Dabei wird jeder Aspekt auf einer vierstufigen Bewertungsskala von „–“ bis „++“ eingeordnet. Anschließend werden die Grundlagen des Berichts dargelegt, wozu eine Erläuterung des Inspektionsablaufs und der eingesetzten Instrumente gehört. Hier erhalten die Schulen auch umfassende Darstellungen der Einschätzungen des Kollegiums zum Führungshandeln (s. o.) nebst einer grafischen Aufbereitung der Ergebnisse zu den fünf Führungsstilen sowie eine detaillierte Aufbereitung der Unterrichtsbeobachtungen. Auch die Ergebnisse der Eltern- und Schülerfragebögen werden hier berichtet. Abschließend wird das detaillierte Bewertungsraster aufgeführt, das den einzelnen Bewertungen zugrunde liegt. Sämtliche Daten erhalten die Schulen auch auf einer Daten-CD.

4. Selbstverständnis

Aufgrund der Erfahrungen aus dem ersten Zyklus hat die Schulinspektion ihren Status, ihre Rolle und ihr Selbstverständnis umfassend reflektiert und weiterentwickelt. Deshalb tritt sie den Schulen im zweiten Zyklus mit einem veränderten Selbstanspruch entgegen, der sowohl nach innen als auch nach außen der Kommunikation und des Diskurses bedarf. Zu beachten ist, dass der Ausgang für diesen Prozess bereits im ersten Zyklus liegt: Es handelt sich um eine kontinuierliche Veränderung, die anlässlich des zweiten Zyklus explizit kommuniziert wurde. Die entscheidende Veränderung lässt sich mit der Wendung *„Vom Objektivitätsanspruch zum Deutungsangebot“* fassen. Dabei wird von der epistemologischen Grundüberzeugung ausgegangen, dass auch bei einem kriteriengeleiteten, zwischen verschiedenen Beobachtern konsensual ausgehandelten Beurteilungsverfahren keine absoluten Wahrheiten entstehen können, sondern bestenfalls eine spezifische Sicht auf die jeweilige Schule, die diese wiederum rezipieren und in Passung zu ihren eigenen Sichtweisen bringen muss. Dazu gehört unweigerlich der Abgleich, an welchen Stellen und in welchem Umfang die Schule die Einschätzungen der Schulinspektion teilt oder auch explizit nicht teilt. In diesem Sinne macht die Schulinspektion ein Angebot, wie Schulen und ihre Qualitäten auf der Grundlage nachvollziehbarer, transparenter Kriterien und entsprechender wissenschaftlich fundierter Datenerhebungen im Rahmen eines intersubjektiv geteilten Bewertungsverfahrens gedeutet

werden können – keinesfalls jedoch, wie sie „objektiv“ oder „tatsächlich“ sind. Empirisch bietet sich ein solcher Zugang allein deshalb an, weil auch innerhalb der Schule unterschiedliche Sichtweisen, wenn nicht gar unterschiedliche Wirklichkeiten bestehen, und es auch von daher unmöglich ist, der einen oder anderen Wahrnehmung einen ontologisch höheren Stellenwert zuzuschreiben. Vielmehr scheint es fruchtbar, diese unterschiedlichen Wahrnehmungen möglichst wertfrei nebeneinander zu stellen und nach Überschneidungen oder auch möglichen Ursachen für Abweichungen zu suchen.

Ein solcher Zugang lebt allerdings von Voraussetzungen sowohl auf Seiten der Schulen als auch der Schulinspektion selbst, die nicht ohne Weiteres gewährleistet sind. So setzt die Vorstellung eines Deutungsangebots implizit auf ein gleichberechtigtes, also nichthierarchisches Verhältnis, das seitens der Schulen nicht immer erlebt wird; vielmehr bestehen sie – auch unter Verweis auf die Veröffentlichung der Berichte – häufig darauf, dass zumindest für die Dauer des Inspektionsprozesses eine hierarchische Beziehung etabliert werde, in der die eine der anderen Instanz übergeordnet sei und an deren Ende ein richtiges oder falsches Urteil stehe. Tatsächlich laufen die Inspektionsberichte durch die Veröffentlichung Gefahr, von Außenstehenden als ein objektives Urteil wahrgenommen zu werden. Diesen Widerspruch zwischen behauptetem Deutungsangebot auf der einen Seite und wahrgenommener Hierarchie auf der anderen Seite aufzuheben, bedarf der Überzeugung und insbesondere eines als glaubhaft erfahrbaren Handelns der Inspektorinnen und Inspektoren. Diese wiederum werden durch die Vorstellung, keine richtigen oder falschen, sondern bestenfalls besser oder schlechter begründbare Urteile abzugeben, die unweigerlich auf eine innerschulische Verarbeitung angewiesen sind und dort dann eben auch zurückgewiesen werden können, verstärkt dazu gebracht, eigenes Handeln zur Disposition zu stellen und damit hinterfragbar zu werden. Vereinfacht gesagt: Während man sich als Inspektorin oder Inspektor im ersten Zyklus auf die Eindeutigkeit und Unfehlbarkeit der empirischen Instrumente zurückziehen konnte, stehen jetzt, im zweiten Zyklus, Diskurs, Überzeugungsarbeit und gegebenenfalls Revidierbarkeit der Urteilsbildung im Vordergrund. Dies setzt ein anderes professionelles Selbstverständnis voraus und bedarf einer veränderten Selbstreflexion der Betroffenen.

Eng mit diesem Wandel verbunden ist ein Bewusstsein für die Notwendigkeit, die Schulinspektion als ein Evaluationsinstrument unter vielen zu begreifen, das eine spezifische, aber eben nicht erschöpfende Sicht auf die Schule bieten kann. So stehen den Schulen potenziell Daten aus zahlreichen Verfahren zur Verfügung: Vergleichsarbeiten (in Hamburg KERMIT), Sprachfördermonitoring, Ergebnisse aus zentralen Abschlussprüfungen, Befunde zu ihrer speziellen Schülerklientel (Sozialindex und „Schule im Überblick“), aber auch diverse interne Evaluationsbefunde. Die Herausforderung liegt nun darin, diese Vielzahl von Daten in einen Zyklus des Qualitätsmanagements zu integrieren, das spezifische Potenzial der einzelnen

Befunde auszuschöpfen, etwaige Widersprüche zu verstehen und einzuordnen und darauf aufbauend die angemessenen Schlussfolgerungen für die weitere Entwicklung der Schule und des Unterrichts zu ziehen. Mit diesem Anspruch sind viele Schulen nach wie vor überfordert, zumal oftmals vermeintlich dringendere Aufgaben anstehen und der vollen Aufmerksamkeit bedürfen. Die Schulen bei ihrer Entwicklung und der Entwicklung ihres Unterrichts durch angemessene Datenaufbereitungen und weiterführende Strategien zu unterstützen, wird künftige Aufgabe sowohl der Schulinspektion als auch der weiteren datenrückmeldenden Stellen der Behörde für Schule und Berufsbildung, nicht zuletzt aber auch der Schulaufsicht und des Beratungs- und Unterstützungssystems sein. Auch deshalb setzt die Schulinspektion ihre im ersten Zyklus begonnene intensive Zusammenarbeit mit diesen Schnittstellenpartnern fort.

Bei den Bemühungen, ihre Ergebnisse für die weitere Schulentwicklung fruchtbar zu machen, darf die Schulinspektion allerdings eines nicht übersehen: Ob die Schulen mit ihren Daten arbeiten und sie nutzen, hängt letztlich nur bedingt von der Datenqualität und ihrer wissenschaftlichen Fundierung ab. Vielmehr bestimmen Offenheit, Bereitschaft und Erfahrung auf Seiten der Schulen, inwieweit sie für empirische Ergebnisse überhaupt zugänglich sind und sie für eine wertvolle Quelle für die Schul- und Unterrichtsentwicklung halten. Damit rückt der Fokus auf einen weiteren Aspekt des Inspektionsprozesses, dem im Vergleich zum ersten Zyklus nun noch mehr Aufmerksamkeit gewidmet wird. Die Qualität der Vermittlung der Inspektionsergebnisse dürfte entscheidende Weichen dafür stellen, ob und inwieweit sich Schulen über den Zeitpunkt der Rückmeldung hinaus mit den Berichten befassen. Entsprechend den oben zitierten Befunden kommt deshalb im zweiten Zyklus dem Rückmeldegespräch mit der Schulleitung eine herausgehobene Stellung zu. Dieses soll so gestaltet werden, dass nach Möglichkeit ein Dialog zwischen Schulleitung und rückmeldender Teamleitung entsteht. Der Prozess des Abgleichs von unterschiedlichen Sichtweisen, des Aufzeigens von Widersprüchen oder auch des Zurückweisens von bestimmten Einschätzungen soll so bereits in der Rückmeldesituation initiiert werden. Damit dies gelingen kann, kommt es vor allem auf das kommunikative Geschick und die Fähigkeit zur Herstellung einer gleichberechtigten und offenen Gesprächsatmosphäre an. Dies wiederum setzt einen entsprechend konstruktiven Verlauf der gesamten Inspektion voraus, in dem die jeweilige Schulleitung durch ein entsprechendes Agieren der zuständigen Inspektorin bzw. des zuständigen Inspektors davon überzeugt worden ist, dass es nicht primär um Beurteilung geht, schon gar nicht um Verurteilung, sondern um die Initiierung eines nutzbringenden Beitrags zur Schulentwicklung. Spätestens hier wird deutlich: Inspektion ist vor allem ein kommunikativer Prozess, der anspruchsvolle Qualifikationen auf beiden Seiten voraussetzt. Oder noch simpler gesagt: Inspektionsstätigkeit ist Beziehungsarbeit.

Literatur

- Hallinger, P. (2003). Leading educational change. Reflections on the practice of instructional and transformational leadership. *Cambridge Journal of Education*, 33 (3), 329–351.
- Huber, S. (2005). Führungskonzeptionen und Führungsmodelle im Überblick. In A. Bartz, M. Dammann, S.G. Huber, T. Klieme, C. Kloft, M. Schreiner (Hrsg.), *PraxisWissen Schulleitung* (Loseblattsammlung, o.S.). München: Wolters Kluwer. Verfügbar unter http://bildungsforschung.net/pdf_gesichert/Huber-2005-FuehrungskonzeptionenUndFuehrungsmodelleImUeberblick.pdf [21.10.2014]
- Pietsch, M. (2008). Beurteilereffekte bei der Messung von Unterrichtsqualität: Das Multifacetten-Rasch-Modell und die Generalisierbarkeitstheorie als Methoden der Qualitätssicherung in der externen Evaluation von Schulen. *Zeitschrift für Erziehungswissenschaft*, 11 (3), 430–452.
- Pietsch, M. (2010). Evaluation von Unterrichtsstandards. *Zeitschrift für Erziehungswissenschaft*, 13 (1), 121–148.
- Pietsch, M. (2014a). Was wissen wir über wirksame Schulleitungen? Eine Zusammenschau und praxisorientierte Einordnung von Best-Evidence-Forschungsbefunden der letzten zehn Jahre. *Journal für Schulentwicklung*, 18 (2), 15–23.
- Pietsch, M. (2014b). Wirksame Schulleitungen: Annahmen, Befunde und der Blick der Schulinspektion. *Hamburg macht Schule*, 26 (1), 6–11.
- Scheerens, J. (2012). *School Leadership Effects Revisited. Review and Meta-Analysis of Empirical Studies*. Dordrecht: Springer. Verfügbar unter http://www.schoolleadership.eu/sites/default/files/scheerens-schoolleadershipeffectsrevisited-march2012_3.pdf [21.10.2014]
- Scheerens, J. & Bosker, R.J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Scheerens, J., Luyten, H., Steen, R. & de Luyten-Thouars, Y. (2007). *Review and meta-analyses of school and teaching effectiveness*. Enschede: Universiteit Twente, Afdeling Onderwijsorganisatie en -management. Verfügbar unter <http://www.utwente.nl/gw/owk/employees/Emeriti/scheerens/rev-meta-analysis2007.PDF> [21.10.2014]
- Teddlie, C. & Reynolds, D. (2000). *The international handbook of school effectiveness research*. London: RoutledgeFarmer.

Die Perspektive der Stakeholder: Ein Interview von Martin Spiewak

Martin Spiewak, Angela Hellberg, Claudia Lemke, Gerrit Petrich, Norbert Rosenboom & Mina Zulal

Zusammenfassung

Das zentrale Ziel der Schulinspektion Hamburg ist es, dazu beizutragen, dass sich die Qualität von Schulen in Hamburg nachhaltig verbessern kann. Damit dies gelingt, braucht sie immer wieder die Rückmeldung der schulischen Interessengruppen, um sich auf dieser Grundlage kontinuierlich weiterzuentwickeln. Im Interview mit dem Bildungsjournalisten Martin Spiewak von der ZEIT berichten der Amtsleiter und Landesschulrat Norbert Rosenboom, Dr. Claudia Lemke als Vertreterin der Hamburger Lehrerkammer, Gerrit Petrich als Vertreter der Elternkammer, die Vorsitzende der Hamburger SchülerInnenkammer Mina Zulal sowie Angela Hellberg als Vertreterin von Handelskammer und Landesschulbeirat über Erfahrungen mit der Schulinspektion aus den vergangenen Jahren, diskutieren Stärken und Schwächen des Verfahrens und entwickeln Ideen für künftige Verbesserungen.

Martin Spiewak: Herr Rosenboom, die Schulinspektion gibt es jetzt schon seit sieben Jahren. Sie haben sie mit initiiert und mit gegründet. Im Rückblick: Was hat Sie am meisten überrascht? Was hätten Sie anders erwartet?

Norbert Rosenboom: Am meisten bin ich überrascht, dass die Inspektion ein voll akzeptiertes Instrument ist. Klagen über Inspektionen gab es eher in der Anfangsphase. Mittlerweile beziehen sich Klagen nur noch auf den Zeitpunkt der Inspektion. Gelegentlich wird dieser als ungünstig empfunden und dann bekomme ich beispielsweise zu hören: „Zurzeit ist es bei uns ganz schwierig, verschiebt doch bitte die Inspektion.“ Ansonsten melden die Schulen zurück: Inspektion ist für uns ein sinnvolles Verfahren, wir erfahren etwas, es wird öffentlich gemacht und wir können damit gut umgehen.

Martin Spiewak: Warum hat Sie das überrascht?

Norbert Rosenboom: Weil wir ab 2000 mit der empirischen Wende eine tief greifende Veränderung im Schulsystem hatten. Damals wurde alles standardisierte Messen

höchst kritisch gesehen, eben auch ein standardisiertes Controllingverfahren wie die Schulinspektion. Und deshalb hat es mich verblüfft, wie schnell dieses Verfahren Akzeptanz gefunden hat. Das deute ich so und das haben auch Schulleiter mir gegenüber gesagt: Die Schulinspektion sagt uns wirklich etwas.

Martin Spiewak: Gleich zu den Lehrern – Herr Rosenboom sagte, die Akzeptanz in den Schulleitungen ist überraschend groß oder extrem groß. Frau Lemke, ist das mittlerweile auch bei den Lehrern so?

Claudia Lemke: Nein, das ist bei den Lehrern ganz bestimmt nicht so. Die Lehrer fühlen sich ein Stück weit außen vor. Sie fühlen sich als Akteure vernachlässigt. Sie werden oftmals als Widerstand gegen Change-Prozesse dargestellt. Das widerspricht natürlich der Selbstwahrnehmung. Darum beteiligen sie sich häufig auch nicht an der Online-Befragung und die Interviews sind nicht wirklich repräsentativ. Die Lehrer nehmen wahr, dass bei der Inspektion die Instruktionstechniken im Vordergrund stehen. Wir sehen uns aber nicht nur als diejenigen, denen die Kompetenzen beigebracht werden müssen, um diese Instruktionstechniken anwenden zu können. Und insofern kann ich nicht so eindeutig sagen, die Schulinspektion findet breite Akzeptanz bei den Kolleginnen und Kollegen. Die empirische Wende hat vor allem den Druck auf die Lehrerinnen und Lehrer erhöht. Und es ist ein Druck, der von oben nach unten weitergegeben wird, und zwar mit der Veränderung der Schulleiterfunktion, die sich jetzt nicht mehr als Kolleginnen und Kollegen verstehen, sondern als Führungspersönlichkeiten und als jemand, der selber unter Druck steht von oben und das weitergeben muss. Es hat viel an Innovationen gegeben in den letzten Jahren, zum Beispiel Individualisierung, Kompetenzorientierung, Ganztag.

Norbert Rosenboom: Aber noch mal direkt zum Instrument Inspektion. Inwiefern verbinden Sie das mit einer Inspektion? Sie sagen, die Lehrer werden zu wenig gesehen.

Claudia Lemke: Richtig, und ein ganz großer Teil unserer Arbeit wird in der Inspektion überhaupt nicht abgebildet. Es ist überhaupt nicht repräsentativ für das, was wir machen. Wir haben einen Bildungs- und Erziehungsauftrag und wir sind nicht Ausführende irgendwelcher Instruktionstechniken. Wir verstehen uns als der Bildung verpflichtet. Und Bildung lässt sich nicht reduzieren auf Instruktionstechniken. Wir sehen nicht, wo die Inspektionsberichte wirklich ein Impuls für Schulentwicklung sein können, denn sie verschwinden in der Schublade der Schulleitung und sind manchmal nicht einmal in den Sekretariaten einsehbar.

Martin Spiewak: Das ist ein anderer Punkt, zu dem wir noch kommen: die Wirkung von Schulinspektion. Aber vielleicht zu den anderen: Was für eine Bilanz

würden Sie ziehen? Was ist aus Elternsicht positiv gelaufen und was läuft in Ihren Augen schief?

Gerrit Petrich: Ich würde das in Theorie und Praxis unterteilen. Die Idee der Schulinspektion wird von den Eltern begrüßt, denn wir verstehen sie als Teil des Qualitätsmanagementprozesses. Die Eltern erleben sie nicht als ein autoritäres System, in dem Druck von oben nach unten weitergegeben wird. Die Probleme stecken eher in der Praxis. Das beginnt schon bei der Elternbeteiligung: Bei den schriftlichen Befragungen sind die Rückläufe immer sehr gering, man erreicht selten eine repräsentative Quote. Ein zweiter schwieriger Bereich ist die Kommunikation der Ergebnisse. Bei der schulöffentlichen Präsentation wird nur ein Teil vorgestellt, nicht der gesamte Inspektionsbericht. Und am Ende, da gebe ich Ihnen recht, ist unklar: Was passiert mit den Ergebnissen? Wie kriegen wir das Zirkuläre im Qualitätsmanagementprozess hin, dass tatsächlich Veränderungen angestoßen werden? Die Schulinspektion war da, es gibt einen Bericht – und was folgt dann? Für Eltern sind die anschließenden Schritte nicht transparent. Man setzt sich eben nicht in der Schulkonferenz zusammen und bespricht die Ergebnisse und das was daraus folgt, sondern die Ergebnisse fließen bestenfalls in die Ziel-Leistungsvereinbarungen zwischen Schulaufsichtsbeamten und der Schule ein. Da sind Eltern dann außen vor.

Martin Spiewak: Gehen wir zur dritten wichtigen oder vielleicht wichtigsten Gruppe innerhalb der Schule: Was kommt denn bei den Schülern an von der Schulinspektion? Was würdest du sagen: Bekommen die Schüler das überhaupt mit?

Mina Zulal: Ja, an meiner Schule war das der Fall. Wir haben dazu eine Ganztagskonferenz gemacht, das heißt, dass alle Klassensprecher eingeladen wurden, um daran teilzunehmen. Allerdings glaube ich, dass es an jeder Schule anders gehandhabt wird, deswegen ist unsere Schule vielleicht nicht ganz so repräsentativ. Was wir sehen, ist, dass die Gespräche, die mit einer Schülergruppe geführt werden, nicht wirklich repräsentativ sind, da der Schulleiter ja die Möglichkeit hat, diese Schüler auszusuchen. Vielleicht könnte man ein Verfahren entwickeln, um es breiter zu streuen und die Schüler zufällig auszuwählen. Eine andere Sache ist uns aufgefallen, die betraf einige der Kategorien bei der Ergebnisvorstellung. Wir fanden sie nicht zutreffend für die Schulen. Beispielsweise wurden wir darin bewertet, wie wir als Ganztagschule sind, dabei sind wir gar keine Ganztagschule. Da wurde ein falsches Kriterium an die Schule angelegt. Wir finden deshalb, dass die Kategorien noch besser differenziert werden müssen.

Martin Spiewak: Und was bekommt der „normale“ Schüler von der Inspektion mit? Weisen die Lehrer vorher darauf hin? Habt ihr festgestellt, dass die Lehrer

anders unterrichten als sonst? Hast du schon mal gehört, dass Lehrer auf einmal besser vorbereitet sind? Das ist ja auch ein häufiger Vorwurf, für die Inspektion finde eine Showveranstaltung statt.

Mina Zulauf: Das ist zwar subjektiv, aber bei uns gab es keine drastische Veränderung des Unterrichts.

Martin Spiewak: Was würden Sie dazu sagen, Frau Lemke? Machen Kollegen von Ihnen auf einmal einen ganz anderen Unterricht?

Claudia Lemke: Ich erlebe Unterschiedliches. Es gibt viele Kolleginnen und Kollegen, denen ist es egal, und das ist gut so. Das trifft aber nicht auf alle zu. Manche denken, sie müssen bei der Inspektion einen besonders guten Unterricht machen und sich auf diese Stunden besonders gut vorbereiten. Und andere sagen: Ich muss sowieso mal zum Zahnarzt, dann mache ich das, wenn die Inspektoren kommen. Im Großen und Ganzen würde ich sagen, dass Unterricht so gemacht wird wie sonst auch. Allerdings muss ich sagen, dass diese zwanzig Minuten Hospitation wirklich gar nichts aussagen über die Qualität von Unterricht, und schon gar nicht über stattfindende Bildungsprozesse. In unseren Augen geht es ausschließlich um Instruktionstechniken, die Fachlichkeit ist komplett ausgeklammert.

Martin Spiewak: Sind denn Instruktionstechniken nicht relevant für den Unterricht?

Claudia Lemke: Instruktionstechniken sind sehr wohl relevant für den Unterricht, aber sie bilden nicht die Qualität ab.

Norbert Rosenboom: In meinen Augen geht es nicht um Instruktionstechniken, sondern um die Frage, ob Unterrichtsstrukturen so stimmen, dass ein Schüler dem dargebotenen Stoff folgen kann. Es geht um eine dritte, objektivierende Sicht auf die Frage, ob die Strukturen so stimmen, dass die Schüler sich anbinden können und theoretisch den Unterricht übernehmen könnten.

Martin Spiewak: Gut. Frau Hellberg, Sie haben ja eine noch stärkere Außensicht, Sie sind ja nicht selbst Betroffene. Wie kommt die Schulinspektion bei Ihnen an?

Angela Hellberg: Zunächst einmal: Wir sind ein absoluter Verfechter der Schulinspektion. Wir haben sie ja auch mit ins Leben getragen. Meine Wahrnehmung ist, dass es am Anfang schon sehr starke Ängste bei den Lehrkräften gab. Dazu muss man auch sagen, es war ja etwas ganz Neues, dass plötzlich jemand ins Klassenzimmer kommt und zuschaut und -hört. Das war einfach ungewohnt und angstbesetzt. Inzwischen erlebe ich es deutlich anders: Alle, auch die Lehrkräfte, haben die

Schulinspektion als Instrumentarium akzeptiert, sie haben sich daran gewöhnt und diese Angst ist nun einfach nicht mehr da.

Martin Spiewak: Auch von Lehrerseite?

Angela Hellberg: Auch von Lehrerseite. Vielleicht ist das an der einen oder anderen Stelle immer noch kritikbesetzt, aber nicht mehr angstbesetzt. Eher wird die Inspektion als Chance für die Weiterentwicklung gesehen. Mit der Schulinspektion sind ja auch Qualitätsteams in den Schulen eingerichtet worden, Qualitätsbeauftragte usw., sodass das ganze Thema auf eine andere Ebene gezogen wird. Inzwischen würde ich behaupten, ist das Ganze eine Selbstverständlichkeit, insbesondere bei den jüngeren, nachwachsenden Lehrern.

Martin Spiewak: Vielleicht noch mal die Eltern: Wie würden Sie das mit der Angst sehen?

Gerrit Petrich: Ich würde es positiv formulieren, so wie Frau Hellberg das auch gesagt hat. Es geht um das Aufbrechen des Einzelkämpfertums, darum dass der Lehrer noch immer zu häufig alleine vor der Klasse steht, darum dass man diesen „Closed Shop“ aufbricht. Es geht darum, dass man die Lehrer dazu bringt, sich mit anderen Lehrern zu vernetzen und auszutauschen: „Die kommen jetzt in meinen Unterricht, wie erlebst du das? Hast du denn dabei ein gutes Gefühl? Was klappt bei dir und was kann ich tun?“ Viele Schulen kommen jetzt dazu, Jahrgangsteams zu bilden, bei denen es dann wirklich nicht mehr so sehr darauf ankommt, welcher Lehrer vor dem Kind steht, sondern wie die Strukturen sind. Das empfinden wir Eltern als sehr positiv, weil es im Moment noch zu sehr auf den einzelnen Lehrer ankommt. Wenn man das aufbrechen kann und man nicht mehr so abhängig ist vom einzelnen Lehrer, der vorne steht, wenn das ein Resultat der Schulinspektion ist, dann können wir das nur begrüßen.

Martin Spiewak: Das ist ja ein Anspruch an ein gesamtpädagogisches Handeln. Sehen Sie denn schon ein Ergebnis in diesem Sinne?

Gerrit Petrich: Ja, ich denke schon, dass die Lehrer nicht mehr die Tür zu machen, sondern dass da eine Glastür ist, und dass sie auch nicht mehr so viel Angst davor haben, dass jemand in den Unterricht kommt, weil sie das als Unterstützungsmechanismus erfahren und nicht als Beurteilungsmechanismus.

Martin Spiewak: Und spüren Sie von Elternseite, dass die Schulen und Lehrer Angst haben?

Gerrit Petrich: Ja, ich kenne jetzt ein, zwei Schulen, bei denen die Schulinspektion vor der Tür steht. Die haben eigentlich weniger Angst davor, beurteilt zu werden, sondern eher Angst vor dem Verwaltungsaufwand. Die denken: „Oh je, es ist Anmelderunde, und es ist Schulinspektion, wie soll ich denn das alles schaffen?“ Aber keine Angst davor, dass jemand kommt und Zensuren verteilt. Das ist es nicht.

Angela Hellberg: Es hat sich enorm viel verändert beim Thema Hospitation, beim Thema kollegiale Unterstützung, aber auch beim Thema Feedback. Dass Schüler ihre Lehrer beurteilen, das gab es doch früher einfach nicht.

Martin Spiewak: Hat das mit der Inspektion zu tun?

Angela Hellberg: Ja, das ist ja auch eine Folge davon, ein Teil der neuen Feedbackkultur.

Mina Zulal: Für die SchülerInnenkammer ist die Feedbackkultur eine zentrale Forderung, die wir für ganz wichtig erachten. Da fühlen wir uns auf Augenhöhe mit den Lehrern. Und wer vor den Resultaten Angst hat, ist irgendwie fehl am Platz.

Martin Spiewak: Offensichtlich gibt es unterschiedliche Perspektiven auf das Thema Angst. Von Lehrerseite sieht es vielleicht ein bisschen anders aus. Zwei Dinge möchte ich noch mal aufgreifen. Einmal die zwanzig Minuten Unterrichtsbeobachtung: Was kann man da überhaupt sehen? Zum anderen die Fachlichkeit: Was kann ein Mathelehrer, der Teil des Inspektionsteams ist, überhaupt im Philosophieunterricht beurteilen?

Norbert Rosenboom: Ich möchte ganz deutlich sagen: Die zwanzig Minuten reichen durchaus aus. Mir ist viel wichtiger, dass ein Inspektionsteam durch so viele Klassen wie möglich geht und einen schnellen Eindruck bekommt: Ist der Unterricht strukturiert, läuft da etwas Organisiertes ab? Wenn ein Inspektor eine ganze Stunde beobachtet, benötigt man viel Zeit für relativ wenig Erkenntnis. Es geht doch darum, dass eine Schule mit so vielen Ebenen wie möglich erfasst wird. Und deswegen ist dieses vermeintliche Schnellverfahren ein für mich sehr angemessenes. Es ist übrigens bundesweit üblich und statistisch auch gar nicht hinterfragt. Es geht auch überhaupt nicht um den einzelnen Lehrer. Aus unserer Sicht geht es darum, ob die jetzige Steuerung funktioniert: Wir setzen Rahmenvorgaben für die selbstverantwortete Schule, die durch eine erweiterte Schulleitung repräsentiert und verantwortet wird. Wenn dann deutlich wird, dass eine Schule unter einer bestimmten Ergebniserwartung zurückbleibt, dann schließe ich daraus: Hier stimmt die Leitung nicht. Wenn also jemand Sorge haben müsste vor einem Inspektionsbericht, wäre es die Schulleitung. Im ersten Zyklus haben wir bei 14 bis 18 Schulen,

die besonders schlecht abgeschnitten hatten, Fallkonferenzen durchgeführt und ein Maßnahmenpaket beraten. Zur Fallkonferenz gehören das Landesinstitut mit den Fortbildnern, die Schulaufsicht, die Inspektoren und die Amtsleitung mit der leitenden Schulaufsicht. Am Ende steht eine Empfehlung an die Schulen zu den Maßnahmen, die sie durchführen sollen, und ein entsprechendes Unterstützungsangebot. Und nach einer angemessenen Frist kommt eine zweite Inspektion. Mit diesem Verfahren konnten wir in Hamburg in kürzester Frist die Unterrichtsqualität heben.

Angela Hellberg: Ich würde gerne zum Thema „Fachlichkeit“ etwas sagen. Ich bin bei uns im Hause auch interne Auditorin. Und da ist es sogar Prinzip, dass man fachfremd ist. Um bei der Schulinspektion zu bleiben: Es nützt nichts, wenn ich die mathematische Formel nachvollziehen kann; wichtig ist der Außenblick, als sozusagen Artfremde.

Martin Spiewak: Jetzt hat Herr Rosenboom ja deutlich gemacht, dass der einzelne Lehrer gar nichts zu befürchten hat. Deshalb noch mal die Frage, woher kommt die Angst?

Angela Hellberg: Ich glaube, es ist in erster Linie Druck auf die Schulleitung. Häufig wird der Druck nach unten weitergegeben, statt klug und bedacht im Sinne eines Qualitätsmanagementzyklus zu handeln. Der organisatorische Aufwand wird oft erst mal irgendwie verteilt. Aber bezüglich der Unterrichtshospitationen ist mein Eindruck, dass die meisten Kollegen das entspannt sehen und ihren Unterricht nicht ändern, wenn die Schulinspektion kommt.

Martin Spiewak: Stellen Sie öfter fest, dass eine Schulinspektion Aktivitäten der Schulleitung auslöst, die Druck machen, beziehungsweise nicht so sinnvoll sind?

Angela Hellberg: Ja, tatsächlich. Oft fehlt das Wissen, wie Schulleitung Interkollegialität unterstützen kann, wie eine Feedbackkultur etabliert werden kann, wie Strukturen der Entwicklung überhaupt etabliert werden können.

Martin Spiewak: Das hat aber nichts mit der Schulinspektion an sich zu tun, sondern eher mit dem Wunsch, dass die Schulleitung diese Qualitäten mitbringen sollte oder sie fortgebildet werden sollte, um so etwas besser umsetzen zu können.

Angela Hellberg: Richtig.

Martin Spiewak: Also konkret: Haben Sie das Gefühl, dass die Schulleitung oft, zu oft, mit den Ergebnissen falsch umgeht, zum Beispiel den Druck weitergibt? Ist das innerhalb der Lehrerschaft ein verbreitetes Gefühl?

Claudia Lemke: Das ist ein verbreitetes Gefühl.

Norbert Rosenboom: Hier entsteht meines Erachtens ein falscher Eindruck. Die Schulinspektion ist ein Teil einer bestimmten Steuerungslogik. Dazu gehören zunächst die Rahmenvorgaben der Behörde. Daneben ist es die Schulaufsicht, die vor allem berät, und die Schulinspektion mit einem standardisierten Vergleichsverfahren. Im Sinne der selbstverantworteten Schule nimmt die erweiterte Schulleitung die Ergebnisse der Schulinspektion auf und findet die entsprechenden Maßnahmen, die anschließend mit der Aufsicht in Ziel- und Leistungsvereinbarungen gegossen werden. Systemisch ist das völlig durchgehend gesteuert. Praktisch mag es vorkommen, dass Schulleitungen die falsche Maßnahme ergreifen. Es kann aber auch sein, dass ein Lehrerkollegium gar keine Maßnahme, gar keine Veränderung will.

Angela Hellberg: Ich habe auch Schulleiter erlebt, die gesagt haben: „Jetzt habe ich endlich schwarz auf weiß, was ich schon immer gewusst habe. Jetzt sehe ich, wo unsere Defizite sind, und jetzt kann ich Maßnahmen mit der Schule im Rahmen der Ziel- und Leistungsvereinbarungen erarbeiten.“

Martin Spiewak: Und dass dann Druck ausgeübt wird auf die Lehrer?

Angela Hellberg: Eine Führungskraft muss führen. Natürlich möglichst mit den Beteiligten zusammen, aber dass auch unangenehme Entscheidungen getroffen werden müssen, steht außer Frage.

Claudia Lemke: Noch mal zurück zu dem, was dem Schulinspektionsbericht folgt. Da sagen Sie, Herr Rosenboom, dass das Steuerungsinstrument die Ziel- und Leistungsvereinbarungen sind. Voraussetzung für gute Ziel-Leistungsvereinbarungen sind funktionierende Kommunikationsstrukturen in der Schule, inklusive ordentlicher Verfahren. Man braucht funktionierende Arbeitsgruppen und Teamsitzungen, die die ZLV ordentlich vorbereiten können. Denn Entwicklung findet in Arbeitsgruppen und Teamsitzungen statt. Und wenn dafür nicht genug Raum ist, wenn dafür nicht genügend Funktionszeiten vorhanden sind, dann funktioniert es nicht. Auf dieses Strukturproblem muss geschaut werden, denn man braucht entsprechende Strukturen, um nachhaltig Schulentwicklung stattfinden zu lassen.

Martin Spiewak: Haben die anderen auch das Gefühl, dass Schulen teilweise die Strukturen fehlen, um solche Ergebnisse in die Praxis umzusetzen?

Norbert Rosenboom: Ja, habe ich. Sicher gibt es Schulen, bei denen die Schulleitung nicht ausreichend kommunikative Strukturen schafft, die zu effektiven Ergebnissen führen. Und diese Schulen lassen in den Ergebnissen auch laufend

nach. Denn Lehrer alleine können, wenn die Schule eine bestimmte Größe erreicht, die Struktur nicht bestimmen. Das geht schlichtweg nicht. Und dann fangen sie an, sich zurückzuziehen, zu vereinzeln. Das heißt, eine Schulleitung mit ihrer ganzen erweiterten Mannschaft ist dafür da, genau diese Strukturen zu schaffen. Und wenn sie das nicht tut, dann muss man Maßnahmen ergreifen, damit sie es kann. Das ist die Aufgabe von Behörde an dieser Stelle.

Martin Spiewak: Können Sie in etwa sagen, wie viele Schulleitungen das betrifft?

Norbert Rosenboom: Das ist eine deutliche Minderheit, die das nicht kann. Die Mehrheit setzt den Entwicklungsprozess vernünftig um.

Gerrit Petrich: Ich würde auch sagen, es geht nicht um „Können“, sondern um „Haben“. Also: Gibt es eine Struktur, eine Institutionalisierung, um den Qualitätsprozess anzuschieben? In der Schule meiner Kinder gibt es ein Gremium, in dem sitzen Lehrer, Eltern, Schüler und natürlich der Schulleiter. Dieses Gremium nimmt den Faden auf: Was sind unsere drei top To-dos für die nächsten zwei Jahre, was ist an dieser Schule entwicklungsbedürftig und mit welchen Maßnahmen erreichen wir das? In verschiedenen Gruppen wird ein Konzept erarbeitet, man setzt es um, natürlich schaut man auch zwischendurch immer wieder darauf und nach zwei Jahren schließt man den Kreis und wertet aus, ob man die Ziele erreicht hat oder ob man noch weiter daran arbeiten muss.

Martin Spiewak: Das ist die Idealform. Würden Sie sagen, das ist Standard?

Gerrit Petrich: Nein, das ist nicht Standard, das ist Glückssache.

Martin Spiewak: Wie professionell gehen Schulen mit den Ergebnissen insgesamt um? Wie häufig passiert es, dass sich Schulen die Ergebnisse anschauen und entscheiden, die dringenden Punkte mit entsprechenden Maßnahmen anzugehen?

Claudia Lemke: Also ich würde sagen, es passiert, aber wie immer im Leben, es passiert nicht überall flächendeckend.

Norbert Rosenboom: Ja, genau, zehn Prozent der Schulen machen es nicht.

Gerrit Petrich: Ich würde sagen, weniger als die Hälfte der Schulen macht das professionell.

Mina Zulal: Die Qualität dieser Gruppen ist ziemlich unterschiedlich. Bei mir in der Schule gibt es zum Beispiel eine Schulentwicklungsgruppe, die sich ge-

nau damit beschäftigt. Ich glaube aber, das ist längst nicht die Norm an allen Schulen. Natürlich gibt es auch Schulen, bei denen sind die Ergebnisse aus Schülersicht etwas verloren gegangen. Vielleicht tragen sie irgendwann Früchte, aber der Weg dahin ist nicht transparent. Deswegen halte ich es für den richtigen Ansatz, wenn man solche Schulentwicklungsgruppen zur Norm an Schulen macht. Das kann ruhig verordnet werden. Da sind Schüler sehr bereit und offen für.

Martin Spiewak: Noch mal zur Steuerung: Bei den ganz schwierigen Schulen greifen Sie ein. Was passiert denn mit den Ergebnissen der anderen Schulen? Sind Sie damit zufrieden, wie das umgesetzt wird? Was ist dabei die Rolle der Schulaufsicht?

Norbert Rosenboom: Die Schulaufsicht wertet diese Ergebnisse aus, genau wie ich. In der Regel sind wir zufrieden, weil wir einen Anstieg der Leistung sehen in den verschiedenen Feldern, und daraus erwächst für uns kein unmittelbarer Handlungsbedarf. Was wir allerdings nicht mitbekommen, ist der Kommunikationsprozess im Inneren der Schule nach der Inspektion.

Martin Spiewak: Und wie müsste das aussehen? Bei den schwierigen Schulen läuft es ja wirklich gut, nach dem Bericht erfolgt die Fallkonferenz, Maßnahmen werden durchgeführt und nach einem Jahr schaut die Schulinspektion erneut. Aber bei den anderen neunzig Prozent funktioniert es in meinen Augen nicht richtig. Auch da könnte nach zwei Jahren jemand nachschauen, was eigentlich passiert ist. Und das macht die Schulaufsicht in der Regel nicht, ist mein Eindruck.

Norbert Rosenboom: Eigentlich sollte die Schulaufsicht ihre Sicht auf die Inspektionsergebnisse in die nächste Ziel- und Leistungsvereinbarung einfließen lassen und darüber eine Verbesserung erreichen. Erst dann wäre das Inspektionsergebnis eines, das nachhaltig für Veränderungen sorgen würde. Ich will aber nicht behaupten, dass das immer geschieht. Die Aufsicht ist möglicherweise personell noch nicht genügend ausgestattet, um das zu leisten, was ich jetzt als Erwartung beschrieben habe.

Claudia Lemke: Ich möchte hinterfragen, ob die Ziel- und Leistungsvereinbarungen wirklich ein geeignetes oder ein hinreichendes Steuerungsinstrument sind, ebenso wie Bilanzierungsgespräche zwischen Schulen und Schulaufsicht, die seit einem Jahr laufen. Es sind viele kleine Dinge, die eine Schule sich vornimmt. Und wenn ich ein schlauer Schulleiter bin, dann nehme ich das Matheprojekt, das sowieso schon ein Kollege macht. Das sind nicht unbedingt die Knackpunkte, an denen Schule wirklich verändert werden muss. Im Bilanzierungsgespräch geht es in erster

Linie um die Ziel- und Leistungsvereinbarungen. Und das sind sicherlich nicht die dümmsten Schulleiter, die ein sehr gepflegt distanziertes Verhältnis zu ihrem Schulaufsichtsbeamten haben.

Gerrit Petrich: Aber die Ziel-Leistungsvereinbarungen werden doch beidseitig geschlossen. Und wenn der Schulleiter der Schulaufsicht versucht, etwas Nebensächliches unterzuschieben, erwarte ich von der Schulaufsicht, dass sie das merkt. Das System funktioniert insoweit gut, wie wir eine starke Schulaufsicht haben, und das muss nicht über Autorität gehen, sondern über die Beraterfunktion der Aufsicht. Beide Seiten müssen zur selben Meinung darüber kommen, was der Schule weiterhilft. Und idealerweise ist hinterher auch das Kollegium der Meinung, dass das, was an dieser Schule passieren soll, weiterhilft.

Martin Spiewak: Vielleicht gehen wir noch mal einen Schritt zurück, zur Kommunikation der Ergebnisse. Sie sagten, vieles verschwindet in der Schublade.

Gerrit Petrich: Nicht vieles, aber wir beobachten schon, dass auch manches in der Schublade verschwindet. Ein klassisches Beispiel ist, wenn Eltern eine Schule für ihr Kind suchen und bei der Schulleitung oder im Sekretariat nach dem Inspektionsbericht fragen. „Schulinspektionsbericht? Äh, nein, weiß ich jetzt auch gerade nicht.“ Das ist nichts, womit Schule wirklich offensiv wirbt und sagt: „Das ist der Bericht, da haben wir Stärken, da haben wir Schwächen.“ Es wurde jetzt verfügt, dass die Schulinspektionsberichte verkürzt im Internet veröffentlicht werden.

Martin Spiewak: Finden Sie das gut?

Gerrit Petrich: Die Elternkammer hat es gefordert, ich sehe aber keine Wirkung. Die Eltern wissen es nicht. Und auch wenn wir es kommunizieren über unsere Medien, nimmt es niemand wahr. Für mich bleibt die Frage, was wir mit dem Bericht machen. Wie kann man einem Laienpublikum transportieren, wo die Schule gut aufgestellt ist und wo sie Schwächen hat? Dieser Prozess scheint für mich im Moment noch nicht gut aufgesetzt zu sein. Zwar gibt es eine schulöffentliche Präsentation der Ergebnisse und die Eltern werden eingeladen. Aber häufig kommt keiner. Das ist so wie ein Produkt, das man ins Regal stellt, das aber keiner kauft.

Norbert Rosenboom: Dafür kann aber die Inspektion nichts. Die stellt die Ergebnisse vor und wenn keiner kommt, ist das so. Ich glaube, die andere Seite ist noch viel wichtiger. Wir müssen noch mal nachsteuern; wie beteiligt man alle Schulbeteiligten an der Auswertung des Inspektionsberichtes? Ich meine nicht die Vorstellung,

sondern die Frage, wie daraus Maßnahmen entstehen. Dazu gehören Lehrer, Eltern und Schüler. Man braucht dazu wenigstens eine Arbeitsgruppe, die so einen Bericht durcharbeitet und sich beraten lässt, die die Schulaufsicht mit an den Tisch holt und gegebenenfalls einen Inspektor, um dann zu entscheiden: Hier wollen wir starten und so fokussieren wir die Schule.

Martin Spiewak: Aber das erste Gremium wäre ja eigentlich die Schulkonferenz. Muss es nicht in der Schulkonferenz besprochen werden?

Mina Zual: Wenn ich es rückblickend betrachte, ist unsere Schulentwicklungsgruppe ziemlich deckungsgleich zu unserer Schulkonferenz. Deswegen glaube ich, dass die Schulkonferenz auch das richtige Gremium dafür wäre, um nachhaltige Qualitätssicherung zu ermöglichen. Sie könnte vielleicht auch ein kleines Kontrollorgan sein und die Themen auch wieder in die einzelnen Gremien zurückgeben. Das wäre der richtige Ansatz, um es verbindlich einzuführen.

Norbert Rosenboom: Das sehe ich auch so.

Martin Spiewak: Jetzt werden die Berichte ja veröffentlicht. Hat es Sie überrascht, dass so wenige Reaktionen kamen?

Norbert Rosenboom: Ja, das hat mich sehr überrascht, weil ich ein Ranking befürchtet habe. Das ist überhaupt nicht eingetreten, eher habe ich das Gefühl von großem Desinteresse. Es wurde nicht wahrgenommen. Vielleicht sollte man die Berichte künftig im Ganzen veröffentlichen, denn die Veränderungsnotwendigkeiten sind doch so stark, dass es auch Veränderungsdruck dazu braucht. Und wenn das dann den richtigen Druck erzeugen würde, dann hätte ich damit kein Problem.

Martin Spiewak: Wie sehen die Lehrer das?

Claudia Lemke: Die Veröffentlichung ist eindeutig abgelehnt worden von der Lehrerkammer. Es gibt die Angst vor dem Ranking.

Gerrit Petrich: Meines Erachtens kommt es stark darauf an, was man genau veröffentlicht. Eltern schauen auf Profile bzw. Stärken und Schwächen. Metriken im Sinne von Durchschnittswerten helfen da nicht weiter. Es geht darum, zu sehen, was eine Schule gut kann und was nicht. Schließlich sind Kinder unterschiedlich, dann dürfen die Schulen auch unterschiedlich sein. Darum gibt es auch nicht die beste Schule für alle Schüler, sondern die beste Schule für mein Kind. Für ein anderes Kind ist eine andere Schule die beste. Darum geht es: zu veröffentlichen, was

eine Schule besonders gut kann und wo sie noch Entwicklungsbedarf hat. Dann können Eltern sich wirklich entscheiden.

Martin Spiewak: Liegt es denn im Interesse der Schulbehörde, einen gewissen Wettbewerb zwischen den Schulen herzustellen?

Norbert Rosenboom: Nein, das ist gar nicht unser Interesse. Wir wollen Transparenz herstellen darüber, wie Schulen arbeiten. Wir wollen die Eltern so informieren, dass sie die richtige Schule für ihr Kind wählen können. Es geht darum, Meinungen ein wenig zu „objektivieren“, im Sinne von verlässlicher Information. Häufig geben Gerüchte den Ausschlag, die den Schulen gar nicht gerecht werden. Diese Gerüchteküche ist völlig unsteuerbar, was da umläuft, ist zum Teil vernichtend für Schulen.

Gerrit Petrich: Ich halte eine Konkurrenzsituation zwischen Schulen für völlig falsch. Wir Eltern wünschen uns ein ergänzendes Angebot und kein konkurrierendes. Ich möchte nicht drei Schulen haben, die darum konkurrieren, wer das bessere Sportprofil hat, sondern ich möchte ein Sportprofil, ein musikalisches Profil und ein naturwissenschaftliches Profil. Und das kann ich gar nicht miteinander vergleichen, deswegen nützt mir ein Durchschnittswert nichts.

Martin Spiewak: Wie sehen Sie das, Frau Hellberg?

Angela Hellberg: Transparenz ist ja gegeben, denn die Berichte kann man abrufen. Und für ein Ranking sind wir auch nicht, das macht keinen Sinn. Aber ideal wäre es, wenn die Schulen mithilfe der Berichte voneinander lernen könnten, im Sinne von Good Practice.

Martin Spiewak: Das heißt also, insgesamt gibt es ein überraschendes Desinteresse an den Berichten. Eine ganz andere Frage interessiert mich: Die Schulinspektion folgt einem sehr schematischen Verfahren, das gleiche Muster, die gleichen Zeiten, nur, wenn etwas ganz schiefliegt, dann kommt sie schneller wieder. Halten Sie dieses schematische Vorgehen für richtig oder sollte man stärker individualisiert vorgehen, vielleicht auch öfter, vielleicht auch anlassbezogen oder auch auf Wunsch der Schule?

Norbert Rosenboom: Mein Eindruck ist, dass man nach ein bis zwei Durchgängen genereller Art das Verfahren ändern sollte, im Sinne von Fokusinspektionen. Beispielsweise könnte man sich konzentrieren auf Feedbackverfahren in der Schule und diese gezielt in einem kurzen Abstand rückmelden. Denn ich glaube fest daran, dass eine Schule sich umso mehr einer Sache stellt, wenn sie weiß, dass die

Inspektion nach zwei Jahren mit der gleichen Frage wiederkommt. Damit hätte man einige der Probleme im Griff, die wir heute diskutiert haben, insbesondere mit Blick auf die Ableitung von Konsequenzen aus der Inspektion.

Martin Spiewak: Wie sehen die anderen Beteiligten diese Perspektive?

Angela Hellberg: Uns sind vier Jahre auf jeden Fall auch zu lang. Eigentlich müsste die Inspektion gekoppelt werden mit den Zielvereinbarungen, die in der Regel zwei Jahre laufen. Am Ende der Laufzeit müsste es möglich sein, auf die Entwicklung zu schauen und einzuschätzen, was sich verändert hat.

Gerrit Petrich: Ich stimme ebenfalls zu. Nach vier Jahren haben sich Schulen so stark verändert, dass man die Berichte gar nicht mehr nebeneinander legen kann, weil sich die Rahmenbedingungen geändert haben, die Schülerschaft, das Kollegium.

Claudia Lemke: Die zentrale Frage bei dem Ganzen ist doch: Wie schlägt sich das in nachhaltige Schulentwicklung um? Das ist aus Lehrersicht zentral. Alles was in der Beobachtung näher an die Schule heranrückt, würde ich begrüßen, also kleinere Beobachtungsschwerpunkte wie z. B. Feedbackkultur, Interkollegialität, Kommunikationsstrukturen. Dann muss aber auch gewährleistet sein, dass es nicht nur um Beurteilung geht, sondern dass auch Unterstützungsmaßnahmen vorhanden sind, ganz besonders für Schulen in Not. Es gibt Gesamtkollegien, die fühlen sich komplett alleine gelassen. Für die braucht es verlässliche Unterstützung.

Mina Zulal: Ich halte Inspektionen mit einem engeren Fokus auch für sinnvoll, zumal es überflüssig ist, einige Dinge zu begutachten, die an einer Schule schon gesichert sind. Vier Jahre sind da wirklich zu lang.

Martin Spiewak: Was Sie, Frau Lemke, angesprochen haben, ist ganz wichtig: die Frage der Hilfe. Bekommen die Schulen genug Hilfe, um das, was die Schulinspektion ihnen mitgibt, auch umzusetzen? Was ist Ihre Meinung, Herr Petrich?

Gerrit Petrich: Für Eltern ist das nicht transparent. Es ist nicht klar, was nach dem Bericht passiert. Es kann etwas Gutes in Gang kommen und es kommt auch vor, dass Eltern einbezogen werden, aber das ist dann keine Folge der Schulinspektion, sondern eine Folge von guter Schulleitung. Aber wenn es um Hilfen von außen geht, wie Beratung oder Weiterbildung oder Verabredungen zwischen Schulleitung und Schulaufsicht, dann weiß ich auch nicht, ob Eltern das unbedingt mitbekommen müssen. Solange das passiert, ist es gut.

Angela Hellberg: Das Landesinstitut hat entsprechende Angebote, die man abrufen kann. Vonseiten der Wirtschaft besteht das Angebot, einen Wirtschaftsvertreter ehrenamtlich eine Zeit lang an die Seite der Schulleitung zu holen.

Martin Spiewak: Eigentlich müsste das Verfahren doch klar sein. Die Inspektion meldet der Schule zurück, in einigen Bereichen ist sie gut aufgestellt und in einem hat sie ein Problem. Die Frage ist, was die Schule damit anfängt: Hat sie Hilfen und holt sie sich die? Gibt es genug Angebote?

Norbert Rosenboom: In meinen Augen gibt es genügend Unterstützungsangebote im Landesinstitut. Die Nachfrage, die zurzeit da ist, wird gedeckt. Und wir würden diesen Ansatz auch erhöhen, wenn sich die Nachfrage erhöhte.

Claudia Lemke: Das Problem liegt ja nicht bei den schwierigen Schulen, die in eine Fallkonferenz gehen, sondern bei den anderen: Was passiert mit der nachhaltigen Entwicklung an diesen Schulen?

Martin Spiewak: Es wird nicht draufgeschaut?

Norbert Rosenboom: Selbstverständlich wird da draufgeschaut. Es ist die Aufgabe der Schulaufsicht, hier zu beraten und ans Landesinstitut oder andere Unterstützungsstrukturen zu vermitteln.

Angela Hellberg: Die Frage ist doch, handelt es sich dabei um eine Holschuld oder eine Bringschuld? Ich meine, wenn es ein Angebot vom Landesinstitut gibt und die Anfragen bedient werden, dann müssen die Schulen die Anfrage auch stellen. Sie müssen sich die Hilfe holen.

Martin Spiewak: Ich glaube tatsächlich, dass es von beiden Seiten kommen muss. In dem Sinne geht es sowohl um Bring- als auch Holschuld.

Norbert Rosenboom: Ich sehe nicht, dass die Schulen sich zu wenig melden, um sich Hilfe beim Landesinstitut zu holen. Tatsächlich gehen Schulleitungen sehr oft zur Fortbildung.

Martin Spiewak: Das heißt, insgesamt würden Sie schon sagen, dass das Zusammenspiel von interner und externer Evaluation, aber auch interner und externer Schulentwicklung funktioniert?

Norbert Rosenboom: Insgesamt funktioniert es und an bestimmten Stolperstellen nicht.

Angela Hellberg: In meinen Augen gehört Evaluation zum normalen Arbeitsprozess dazu.

Claudia Lemke: Von den Kolleginnen und Kollegen wird das Thema Evaluation heiß diskutiert. Aus der Praxis kann ich sagen, wir sind ja gar nicht in der Lage, schulintern zu evaluieren. Dazu sind wir nicht ausgebildet. Außerdem fehlen Zeiten hierfür im Lehrerarbeitszeitmodell. Evaluation setzt aber auch voraus, dass entsprechende Ressourcen da sind, um aus dem Alltag herauszutreten und auf die eigene Praxis zu schauen. Wo sind diese Ressourcen?

Norbert Rosenboom: Ich glaube, bei den Kolleginnen und Kollegen, die in den letzten Jahren in die Schulen gekommen sind, ist das schon etwas anders. Wenn ich mir deren Ausbildung ansehe, stelle ich fest, dass ich als Lehrer drei Dinge beherrschen muss. Erstens: Gehe ich hörend und freundlich aufs Kind ein mit den Bedingungen, die es mitbringt? Zweitens: Bin ich fachlich und sozial steuerungsmäßig in der Lage, dieses Kind zu dem zu bringen, was in ihm angelegt ist? Und drittens: Wie reflektiere ich das, was ich tue? Das ist doch die Ebene, über die wir sprechen. Die würde ich nie Evaluation nennen, eher geht es darum, dass sich Lehrer dazu verabreden, in welchen Ritualen, auf welche Art und Weise geht man mit Kindern um, welchen fachlichen Fortschritt haben sie zu erleben und zu leisten. Wir sind auf dieses Gespür angewiesen, dass wir gemeinsam praktisch sagen, wo es hingeht.

Claudia Lemke: Da bin ich ganz bei Ihnen. Es ist ganz wichtig, dass eine gemeinsame Reflexion innerhalb der Kollegien zum Thema „guter Unterricht“ stattfindet. Und dafür muss es verbindlich verankerte Zeiten geben, auch für Teamsitzungen, für Tandems. Das ist eine Steuerungsaufgabe. Es kommt immer noch zu häufig vor, dass die Lehrerinnen und Lehrer alleine vor der Klasse stehen und dann doch die Tür zu machen.

Martin Spiewak: Würden Sie sagen, dass in der Lehrerbildung eine Haltung zur eigenen Reflexion verankert ist? Können junge Lehrer das?

Norbert Rosenboom: Ich gehe davon aus, dass sich Ausbildung gewaltig verändert hat. Reflexion gehört dazu. Auch Feedback ist ganz anderes verankert, einerseits durch die Inspektion, andererseits durch das Schülerfeedback. Da haben wir plötzlich ein Controlling auf der anderen Seite.

Martin Spiewak: Vielleicht schauen wir am Schluss noch einmal auf den dringenden Handlungsbedarf. Herr Rosenboom, was sind denn aus Ihrer Sicht die Hauptbaustellen, auf die Sie durch die Inspektion aufmerksam geworden sind?

Norbert Rosenboom: Die erste Baustelle: Schulleitung. Wenn eine Leitung nicht funktioniert, kann eine Schule sich nicht strukturieren und kann auch keine guten Leistungen erbringen. Die zweite Baustelle: Teamarbeit. Wenn man die vernünftig aufsetzt, würde das viele andere Probleme aufheben. Dazu braucht es Reflexion und Feedbackverfahren. Ich bin überzeugt, wenn die Struktur einer Schule sauber gelegt wird im Zusammenhang aller Beteiligten, dann kommt die Unterrichtsentwicklung von ganz alleine in den Gang.

Martin Spiewak: Dann möchte ich zum Abschluss jeden einmal bitten, zu benennen, was sich bei der Inspektion verändern sollte.

Mina Zulal: Ich würde klar fordern, dass die Qualitätssicherung durch die Schulkonferenz erfolgen soll. Ich bin überrascht, dass das noch gar nicht gesetzt ist, denn eigentlich ist das ein einfacher Mechanismus, der sehr effektiv ist.

Gerrit Petrich: Inspektion darf nicht mit der Präsentation der Ergebnisse aufhören, sondern der Qualitätskreislauf muss mit allen Beteiligten geschlossen werden. Es reichen nicht die ZLV zwischen Schulaufsicht und Schulleitung, sondern die Schulgemeinschaft mit allen Beteiligten soll über Maßnahmen und ihre Umsetzung beraten. Das kann auch heißen, dass nur ein Teil der Schulgemeinschaft diese Maßnahmen umsetzt. Aber es muss darum gehen, bei der nächsten Inspektion zu prüfen, was haben wir gemeinsam erreicht? Schule funktioniert nur mit allen Beteiligten. Das können nicht nur die Lehrer alleine schaffen und auch nicht nur die Behörde alleine.

Angela Hellberg: Ich spreche mich auf jeden Fall dafür aus, dass die Schulinspektion gestärkt wird. Dazu gehören die personelle Ausstattung und die finanzielle Ausstattung. Und die Akzeptanz der Schulinspektion muss erhöht werden. Der Schulinspektion würde ich einfach wünschen, dass sie sich in Schulen so selbstverständlich etabliert, dass wir gar nicht mehr darüber reden müssen.

Claudia Lemke: Es gibt ja durchaus die Forderung, die Schulinspektion abzuschaffen, so wie in Finnland. Jenseits davon ist mir ganz wichtig, dass die Lehrer sich als Akteure im Inspektionsprozess, als Stakeholder einbezogen fühlen. Beispielsweise könnten Lehrkräfte bei der Evaluation der Inspektion beteiligt werden. Der Prozess muss dichter an die Akteure rücken, dichter an die Lehrer und an die Schüler. Das kann so funktionieren, wie Herr Rosenboom vorhin vorgeschlagen hat, indem man auf Teilausschnitte schaut: ganz gezielt auf einzelne Aspekte, die dann in der Entwicklung auch unterstützt werden.

Norbert Rosenboom: Diese Forderung teile ich. Das Zweite ist eine gemeinsame Arbeitsgruppe nach der Schulkonferenz, die vorgeschrieben werden muss. Die Schulkonferenz kann das in meinen Augen nicht leisten, aber man kann eine gemeinsame Arbeitsgruppe mit allen Beteiligten, Lehrern, Schülern, Eltern einrichten, die die Ergebnisse in der Schule bearbeitet und in nachhaltige Entwicklung übersetzt.

Martin Spiewak: Herzlichen Dank.

Autorinnen und Autoren

Altrichter, Herbert

Jahrgang 1954, Dr. phil., o.Universitätsprofessor für Pädagogik und Pädagogische Psychologie an der Johannes-Kepler-Universität Linz. Arbeitsschwerpunkte: Schulentwicklung und Governance des Bildungswesens, Evaluation, qualitative Forschungsmethoden, neue Lernformen, Lehrerbildung.

Brüsemeister, Thomas

Jahrgang 1962, Diplom-Sozialwissenschaftler, Prof. Dr. phil.; Professor für Soziologie mit Schwerpunkt Sozialisation und Bildung, Justus-Liebig-Universität Gießen. Arbeitsschwerpunkte: Educational-Governance-Forschung, Schulforschung, Bildungssoziologie, qualitative Forschungsmethoden.

Diedrich, Martina

Jahrgang 1975, Diplom-Psychologin, Promotion in Erziehungswissenschaften; Leiterin der Abteilung Schulinspektion und Systemmonitoring im Institut für Bildungsmonitoring und Qualitätsentwicklung Hamburg. Arbeitsschwerpunkte: Empirische Schulqualitätsforschung, Schulinspektion, Evaluation im Bildungsbereich, Bildungsmonitoring.

Ehmke, Timo

Jahrgang 1971, Universitätsprofessor, Prof. Dr. habil.; Professor für Erziehungswissenschaft insbesondere Empirische Bildungsforschung am Institut für Bildungswissenschaft der Leuphana-Universität Lüneburg. Arbeitsschwerpunkte: Lehrerbildung, Large Scale Assessments, Kompetenzmessung.

Ehren, Melanie C. M.

Jahrgang 1975, Ph.D. Erziehungswissenschaften; Senior Lecturer am Institute of Education am University College London. Arbeitsschwerpunkte: Bildungsevaluation, Qualitätssicherung und -entwicklung im Bildungswesen.

Gärtner, Holger

Jahrgang 1974, Diplom-Psychologe, PD Dr. habil.; Projektleiter am Institut für Schulqualität der Länder Berlin und Brandenburg e.V. Arbeitsschwerpunkte: interne und externe Evaluation von Schulen, Ländervergleich, regionale Bildungsberichterstattung.

Graw, Stephanie

Jahrgang 1986, M.A. Soziologie; wiss. Mitarbeiterin im Institut für Bildungsmonitoring und Qualitätsentwicklung Hamburg. Arbeitsschwerpunkt: Schulinspektion.

Hellberg, Angela

Jahrgang 1962, stv. Geschäftsführerin der Handelskammer Hamburg und Leiterin der Abteilung Berufsorientierung Schule und Wirtschaft.

Köller, Olaf

Jahrgang 1963, Diplom-Psychologe, Dr. phil. habil.; Professor für empirische Bildungsforschung an der Christian-Albrechts-Universität zu Kiel und Geschäftsführender Wissenschaftlicher Direktor des Leibniz-Instituts für die Pädagogik der Naturwissenschaften und Mathematik (IPN). Forschungsschwerpunkte: Schulleistungsmessung, individuelle und institutionelle Determinanten von Schulleistungen.

Kuhn, Hans-Jürgen

Jahrgang 1953, Lehrer, MdA Berlin, Staatssekretär a.D., von 1991 bis 2013 Referatsleiter im Ministerium für Bildung, Jugend und Sport Brandenburg. Arbeitsschwerpunkte: Schulentwicklung, interne und externe Schulevaluation, Qualitätssicherung und -entwicklung, Schulforschung, Ganztagspädagogik.

Lemke, Claudia

Jahrgang 1966, Dr. phil. (Universität Hamburg), LA (Universität Hamburg), MA Anthropologie (University of Auckland, Neuseeland), BA Psychologie/Anthropologie (University of Auckland, Neuseeland). Lehrerin für Politik, Englisch und Philosophie, stellvertretende Vorsitzende der Hamburger Lehrerkammer.

Lücken, Markus

Jahrgang 1971, Diplom-Psychologe, Dr. phil.; wiss. Referent im Institut für Bildungsmonitoring und Qualitätsentwicklung Hamburg. Arbeitsschwerpunkte: Kompetenzmessung, Testentwicklung, Evaluation von Schulversuchen.

Maritzen, Norbert

Jahrgang 1952, Lehramt für Französisch, Deutsch und Philosophie; Direktor des Instituts für Bildungsmonitoring und Qualitätsentwicklung Hamburg. Arbeitsschwerpunkte: Schulentwicklung, empirische Bildungsforschung.

Müller, Sabine

Jahrgang 1970, Diplom-Pädagogin, Dr. päd.; wissenschaftliche Referentin für interne und externe Evaluation an der Qualitäts- und Unterstützungsagentur – Lan-

desinstitut für Schule des Landes Nordrhein-Westfalen (QUA-LiS NRW). Arbeitsschwerpunkte: schulinterne Evaluation, Schulinspektion, Schulentwicklung.

Peters, Susanne

Jahrgang 1960, Diplom-Pädagogin, Erstes und Zweites Staatsexamen für das Lehramt an der Grund- und Mittelstufe, Doktorandin an der Universität Koblenz; Schulinspektorin am Institut für Bildungsmonitoring und Qualitätsentwicklung Hamburg. Arbeitsschwerpunkte: Evaluation von Schulen, Inklusion, qualitative Forschung, Systemtheorie.

Petrich, Gerrit

Jahrgang 1968, Diplom-Wirtschaftsinformatiker; von 2013 bis 2015 Vorsitzender der Elternkammer Hamburg.

Pietsch, Marcus

Jahrgang 1974, Diplom-Pädagoge, Dr. phil.; Stipendiat für Empirische Bildungsforschung an der Leuphana-Universität Lüneburg, wiss. Referent im Institut für Bildungsmonitoring und Qualitätsentwicklung Hamburg. Arbeitsschwerpunkte: Empirische Bildungsforschung, Educational Measurement, Evaluation von Schulen und Bildungssystemen.

Rosenboom, Norbert

Jahrgang 1948, Lehrer, Landesschulrat und Leiter des Amtes für Bildung der Freien und Hansestadt Hamburg.

Schachtmeyer, Christiane von

Jahrgang 1962, Gymnasiallehrerin mit den Fächern Deutsch und Geschichte; seit 2000 Schulleiterin am Gymnasium Marienthal, seit 2008 Schulformexperten der Schulinspektion Hamburg.

Scheerens, Jaap

Jahrgang 1946, Prof. Dr.; Professor für Sozialwissenschaften; Gastprofessor an der Universität Bristol, Prof. em. an der Universität Twente. Arbeitsschwerpunkt: Bildungseffektivität

Scholand, Barbara

Jahrgang 1959, M.A. Erziehungswissenschaft; wiss. Mitarbeiterin an der Universität Hamburg. Arbeitsschwerpunkte: Ethnographische Bildungs- und Sozialisationsforschung.

Schulte, Klaudia

Jahrgang 1980, Diplom-Psychologin, Dr. rer. nat.; wiss. Referentin für Empirische Analysen im Institut für Bildungsmonitoring und Qualitätsentwicklung Hamburg. Arbeitsschwerpunkte: Erstellung von Sozialindikatoren, datengestützte Schulentwicklung.

Schulze, Peter

Jahrgang 1973, Diplom-Pädagoge und Coach; Schulinspektor im Institut für Bildungsmonitoring und Qualitätsentwicklung Hamburg. Arbeitsschwerpunkt: Externe Evaluation von Schulen.

Schwippert, Knut

Jahrgang 1965, Prof. Dr.; Professor für Erziehungswissenschaft insbesondere Empirische Bildungsforschung an der Universität Hamburg. Arbeitsschwerpunkte: Effektive Schulen, Evaluation, Large-Scale-Untersuchungen, Methoden in Large-Scale-Untersuchungen, Systemmonitoring.

Sowada, Moritz G.

Jahrgang 1980, Diplom-Pädagoge, M.A.; wiss. Mitarbeiter am Institut für Erziehungswissenschaft der Westfälischen Wilhelms-Universität Münster. Arbeitsschwerpunkte: Governance, Evaluation und Entwicklung von Schulen und Schulsystemen.

Spiewak, Martin

Jahrgang 1964, Redakteur mit dem Schwerpunkt Bildung im Hauptstadtbüro der ZEIT in Berlin.

Ulrich, Susanne

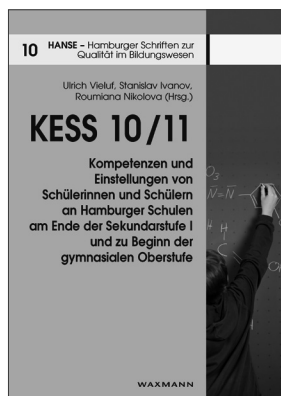
Jahrgang 1960, Leiterin der Schulinspektion Hamburg. Arbeitsschwerpunkt: Schulinspektion.

van den Ham, Ann-Katrin

Jahrgang 1987, M.Sc. Pädagogikwissenschaften; wiss. Mitarbeiterin im Institut für Bildungswissenschaft an der Leuphana-Universität Lüneburg. Arbeitsschwerpunkte: Äquivalenz- und Validitätsanalysen im Zusammenhang mit Schulleistungstudien.

Zulal, Mina

Jahrgang 1998, Schülerin am Carl-von-Ossietzky Gymnasium Hamburg, Landesvorsitzende der Kammer der Schülerinnen und Schüler Hamburg.



Ulrich Vieluf, Stanislav Ivanov,
Roumiana Nikolova (Hrsg.)

KESS 10/11

Kompetenzen und Einstellungen von
Schülerinnen und Schülern an Hamburger
Schulen am Ende der Sekundarstufe I
und zu Beginn der gymnasialen Oberstufe

2012, 280 Seiten, br., 29,90 €,

ISBN 978-3-8309-2615-3

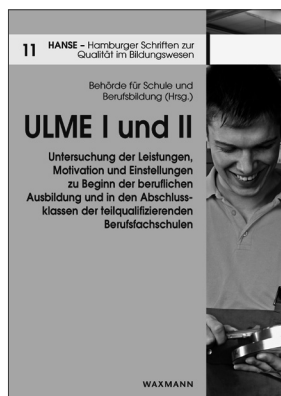
E-Book: 26,90 €, ISBN 978-3-8309-7615-8

Die Hamburger Untersuchung Kompetenzen und Einstellungen von Schülerinnen und Schülern (KESS) ist eine längsschnittlich angelegte Schulleistungsstudie, in deren Rahmen systematisch und kontinuierlich Leistungsdaten eines kompletten Schülerjahrgangs erfasst werden. Gegenstand dieses Berichts ist die vierte Erhebung der Studie (2009), und zwar am Ende der Jahrgangsstufe 10 (Juni) und zu Beginn der Jahrgangsstufe 11 (September).

Der Schwerpunkt der Untersuchung liegt auf der systematischen Beschreibung der Lernausgangslagen, der Lernentwicklungen sowie der fachbezogenen Einstellungen in zentralen Leistungsbereichen, insbesondere in Mathematik, Naturwissenschaften, Deutsch-Leseverständnis, Deutsch-Rechtschreibung sowie Englisch.



WAXMANN
www.waxmann.com



Behörde für Schule und Berufsbildung
(Hrsg.)

ULME I und II

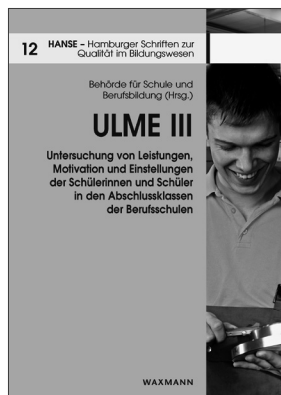
Untersuchung der Leistungen, Motivation
und Einstellungen zu Beginn der beruflichen
Ausbildung und in den Abschlussklassen
der teilqualifizierenden Berufsfachschulen

2013, 432 Seiten, br., 42,90 €,
ISBN 978-3-8309-2849-2

ULME I, im Jahr 2002 durchgeführt, befasst sich mit den Lernausgangslagen, Fachleistungen, Motivationen, Einstellungen, Erwartungen und fachübergreifenden Kompetenzen von Berufs- und Berufsfachschülern in Hamburg am Anfang ihrer beruflichen Ausbildung. An ULME II, durchgeführt im Jahr 2004, nahmen ausschließlich die Schülerinnen und Schüler in den Abschlussklassen der teilqualifizierenden Berufsfachschulen am Ende ihrer zweijährigen Ausbildung teil. Analysiert wurden in ULME II die Lernentwicklung und der Lernerfolg dieser Schülergruppe am Ende ihrer Schulform.



WAXMANN
www.waxmann.com



Behörde für Schule und Berufsbildung
(Hrsg.)

ULME III

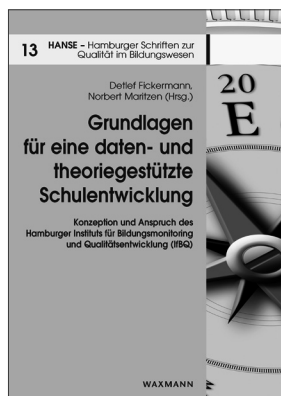
Untersuchung von Leistungen,
Motivation und Einstellungen der
Schülerinnen und Schüler in den
Abschlussklassen der Berufsschulen

2013, 360 Seiten, br., 42,90 €,
ISBN 978-3-8309-2860-7

Hier werden die Ergebnisse von ULME III vorgestellt. An ULME III, durchgeführt im Jahr 2005, beteiligten sich Schülerinnen und Schüler aus 17 Ausbildungsberufen in den Abschlussklassen der Berufsschulen am Ende des dritten und letzten Jahres ihrer Ausbildung. Berichtet werden Ergebnisse zu den kognitiven, metakognitiven und motivationalen Merkmalen der Schülerinnen und Schüler sowie aus berufsübergreifenden und berufsbezogenen Fachleistungstests.



WAXMANN
www.waxmann.com



Detlef Fickermann, Norbert Maritzen (Hrsg.)

Grundlagen für eine daten- und theoriegestützte Schulentwicklung

Konzeption und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)

2014, 304 Seiten, br., 29,90 €,

ISBN 978-3-8309-3053-2

E-Book: 26,99 €, ISBN 978-3-8309-8053-7

Angehörige des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ) diskutieren im ersten Teil des Bandes Elemente einer Theorie des Bildungsmonitorings und stellen konzeptionelle und wissenschaftliche Grundlagen einiger Arbeitsbereiche des Instituts vor. Im zweiten Teil werden aus einer Außenperspektive von Wissenschaftlerinnen und Wissenschaftlern sowie aus der Sicht des Amtschefs (a.D.) eines Kultusministeriums und zweier Schulleitungsmitglieder „Wirkungen“ des Bildungsmonitorings exemplarisch beschrieben und diskutiert. Der dritte Teil widmet sich der Frage, ob Schulen im Zeichen der „empirischen Wende“ zwischenzeitlich durch wissenschaftliche Untersuchungen überbeansprucht sind.



WAXMANN
www.waxmann.com



Behörde für Schule und Berufsbildung
(Hrsg.)

Bildungsbericht Hamburg 2014

2014, 136 Seiten, br.,
mit zahlreichen farbigen Abbildungen, 29,90 €,
ISBN 978-3-8309-3171-3

Mit dem Hamburger Bildungsbericht 2014 liegt zum dritten Mal eine aktuelle Bestandsaufnahme des Hamburger Bildungswesens vor. Der Bildungsbericht bietet einen indikatorengestützten Überblick über Bedingungen, Strukturen und Ergebnisse von Bildung und fokussiert dabei auf das vorschulische und schulische Bildungssystem. Insbesondere der Schulbereich ist in den vergangenen Jahren zahlreichen Veränderungen unterworfen gewesen, so dass die fortgesetzte Hamburger Bildungsberichterstattung eine wertvolle Gesamtschau der Entwicklungen bietet.



WAXMANN
www.waxmann.com