12. Beiheft, 2013

Die Deutsche Schule

Zeitschrift für Erziehungswissenschaft Bildungspolitik und pädagogische Praxis

DDS

Herausgegeben von der Gewerkschaft Erziehung und Wissenschaft

12

Isabell van Ackeren, Martin Heinrich, Felicitas Thiel (Hrsg.)

Evidenzbasierte Steuerung im Bildungssystem?

WAXMANN

Die Deutsche Schule

Zeitschrift für Erziehungswissenschaft, Bildungspolitik und pädagogische Praxis

Herausgegeben von der Gewerkschaft
Erziehung und Wissenschaft

12. Beiheft

Waxmann 2013
Münster / New York / München / Berlin

Isabell van Ackeren, Martin Heinrich,
Felicitas Thiel (Hrsg.)

Evidenzbasierte Steuerung im Bildungssystem?

Befunde aus dem BMBF-SteBis-Verbund

Waxmann 2013
Münster / New York / München / Berlin

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.

ISBN 978-3-8309-7962-3

© Waxmann Verlag GmbH, 2013

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Christian Averbeck, Münster
Umschlagbild: © Woodapple – Fotolia.com Satz:
Stoddart Satz- und Layoutservice, Münster

Alle Rechte vorbehalten
Printed in Germany

DDS – Die Deutsche Schule
Zeitschrift für Erziehungswissenschaft,
Bildungspolitik und pädagogische Praxis
12. Beiheft, 2013

INHALT

EDITORIAL

Evidenzbasierte Steuerung im Bildungssystem?
Befunde aus dem BMBF-Förderschwerpunkt Steuerung
im Bildungssystem (SteBis) 11

STEUERUNGSINSTRUMENTE UND EVIDENZQUELLEN IM VERGLEICH

Sebastian Wurster/Dirk Richter/Anna Schliesing/Hans Anand Pant
Nutzung unterschiedlicher Evaluationsdaten an
Berliner und Brandenburger Schulen
Rezeption und Nutzung von Ergebnissen aus Schulinspektion,
Vergleichsarbeiten und interner Evaluation im Vergleich 19

Isabell van Ackeren/Carmen Binnewies/Marten Clausen/Denise Demski/
Christian Dormann/Anna Rosa Koch/Bastian Laier/Peter Preisendörfer/
Daja Preuße/Christoph Rosenbusch/Uwe Schmidt/Martin Stump/
Olga Zlatkin-Troitschanskaia (Autorengruppe EviS)
Welche Wissensbestände nutzen Schulen im Kontext von Schulentwicklung?
Theoretische Konzepte und erste Befunde des EviS-Verbundprojektes
im Überblick 51

VERGLEICHSARBEITEN / LERNSTANDSERHEBUNGEN

Uwe Maier/Carolin Ramsteck/Annette Frühwacht
Lehr-lerntheoretische Argumentationsmuster bei der
Interpretation und Nutzung von Vergleichsarbeitsrückmeldungen
durch Gymnasiallehrkräfte 74

Barbara Muslic/Carolin Ramsteck/Harm Kuper
Das Verhältnis von Schulleitung und Schulaufsicht
im Kontext testbasierter Schulreform
Kontrastive Fallstudien zur Rezeption von Lernstandsergebnissen
im Mehrebenensystem der Schule 97

SCHULINSPEKTIONSVERFAHREN

Oliver Böhm-Kasper/Odette Selders
„*Schulinspektionen sollten regelmäßig durchgeführt werden*“*?*
Ländervergleichende Analyse der Wahrnehmung und Akzeptanz von Schulinspektionsverfahren 121

Bianca Preuß
Akteurkonstellation zwischen Schulträger und Schule
Empirische Analysen zur Governance von Schulinspektion 154

ERWEITERTE SCHULISCHE VERANTWORTUNG UND INTERNE EVALUATION

Stefan Brauckmann/Christoph Herrmann
Belastungserleben von Schulleiterinnen und Schulleitern im Rahmen erweiterter schulischer Eigenständigkeit
Erste empirische Befunde aus der SHaRP-Studie 172

Katja Thillmann/Anabel Bach/Christina Gerl/Felicitas Thiel
Schulische Selbststeuerung durch interne Evaluation
Eine Analyse von Evaluationsprogrammen Berliner und Brandenburger Gymnasien 197

STEUERUNG UND UMSTRUKTURIERUNG VON SCHULEN IN HERAUSFORDERNDEN LAGEN

Kathrin Racherbäumer/Christina Funke/Isabell van Ackeren/Marten Clausen
Datennutzung und Schulleitungshandeln an Schulen in weniger begünstigter Lage
Empirische Befunde zu ausgewählten Aspekten der Qualitätsentwicklung 226

Laura Fölker
„Da ist einfach ’ne gewisse Leistungs- und auch Schulentwicklungsgrenze erreicht!“
Schulfusionen als zusätzliche Reformimpulse im Zuge der Integration von Haupt- und Realschulbildungsgängen 255

BERICHT

Felicitas Thiel/Katja Thillmann/Jasmin Tarkian
Die Förderung des *Practice Impact* Empirischer Bildungsforschung als Herausforderung für die Koordination von Forschungsverbünden
Bericht der Koordinierungsstelle des Forschungsverbunds „Steuerung im Bildungssystem“ 276

DDS – Die Deutsche Schule
Zeitschrift für Erziehungswissenschaft,
Bildungspolitik und pädagogische Praxis
Supplement 12, 2013

CONTENTS

EDITORIAL

Felicitas Thiel/Martin Heinrich/Isabell van Ackeren
Evidence-Based Governance in the Educational System?
Results from the BMBF Funding Priority Governance in
the Educational System (SteBis) 11

A COMPARISON OF GOVERNANCE INSTRUMENTS AND SOURCES OF EVIDENCE

Sebastian Wurster/Dirk Richter/Anna Schliesing/Hans Anand Pant
Use of Various Evaluation Data in Schools in the Federal States of Berlin and Brandenburg
A Comparison of the Feedback Reception and Use
of School Inspection, State-wide Tests and Self-Evaluation 19

Isabell van Ackeren/Carmen Binnewies/Marten Clausen/Denise Demski/Christian Dormann/Anna Rosa Koch/Bastian Laier/Peter Preisendörfer/Daja Preuße/Christoph Rosenbusch/Uwe Schmidt/Martin Stump/Olga Zlatkin-Troitschanskaia (Autorengruppe EviS)
What Kind of Knowledge Do Schools Use for School Development Purposes?
Theoretical Concepts and Preliminary Findings of the EviS-Project.....52

MANDATORY TESTING / ASSESSMENT OF LEARNING PROGRESS

Uwe Maier/Carolin Ramsteck/Annette Frühwacht
Instrumental Feedback Data Usage by German Grammar School Teachers.....74

Barbara Muslic/Carolin Ramsteck/Harm Kuper
The Relation between Principals and School Supervising Authorities in the Context of Test-Based School Reform
Contrastive Case Studies on the Usage of Mandatory Proficiency Test Data
in the Multi-Level Structure of Schools.....97

SCHOOL INSPECTION PROCEDURES

"***School Inspections Should Be Conducted Regularly*"*?***
Comparing the Perception and Acceptance of School Inspection Procedures across German Federal States 121

Bianca Preuß
The Constellation of Actors between Local Education Agencies and Schools
Empirical Analyses with Regard to the Governance of School Inspection 154

EXTENDED RESPONSIBILITY OF SCHOOLS AND INTERNAL EVALUATION

Stefan Brauckmann/Christoph Herrmann
Workload of School Principals in the Context of Greater School Autonomy
First Empirical Findings from the SHaRP-Study 173

Katja Thillmann/Anabel Bach/Christina Gerl/Felicitas Thiel
Schools' Self-Governance by Using Internal Evaluation
An Analysis of Evaluation Programs at Schools in Berlin and Brandenburg 197

GOVERNANCE AND RESTRUCTURING OF SCHOOLS IN CHALLENGING CIRCUMSTANCES

Kathrin Racherbäumer/Christina Funke/Isabell van Ackeren/Marten Clausen
Quality Development of Schools in Challenging Circumstances
Empirical Findings on Data Use and Quality Development 226

Laura Fölker
"A Certain Limit of Effort and School Development Has Been Reached!"
School Fusions as Additional Reform Impulses in the Course of the Integration of Various Secondary Schools (*Hauptschulen* and *Realschulen*) 256

REPORT

Felicitas Thiel/Katja Thillmann/Jasmin Tarkian
The Advancement of the Practice Impact of Empirical Educational Research as Challenge for the Coordination of Collaborative Research Projects
Report from the Coordination Board of the Collaborative Research Project "Governance in the Educational System" 276

Gutachter und Gutachterinnen von Beiträgen für das Beiheft 2013

Die Redaktion der DDS – Die Deutsche Schule bedankt sich herzlich für die Erstellung von Gutachten für die Beiträge dieses Beihefts bei folgenden Kollegen und Kolleginnen:

Prof. Dr. Herbert Altrichter (Linz-Auhof), Prof. Dr. Martin Bonsen (Münster), Dr. Frauke Choi (Mainz), Dr. Fabian Dietrich (Hannover), Prof. Dr. Birgit Eickelmann (Paderborn), Dr. Jana Groß Ophoff (Freiburg), Prof. Dr. Stephan Huber (Zug), Prof. Dr. Till Sebastian Idel (Bremen), Dr. Sieglinde Jornitz (Frankfurt a.M.), Prof. Dr. Rick Mintrop (Berkeley), Dr. Sabine Müller (MSW Düsseldorf), Dr. Marcus Pietsch (Kiel), Prof. Dr. Peter Posch (Klagenfurt), Prof. Dr. Carsten Quesel (Aarau), Prof. Dr. Knut Schwippert (Hamburg), Ulrich Steffens (Wiesbaden), Prof. Dr. Klaus-Jürgen Tillmann (Berlin), Dr. Julia Warwas (Bamberg), Prof. Dr. Jochen Wissinger (Gießen), Prof. Dr. Beate Wischer (Osnabrück)

EDITORIAL

Evidenzbasierte Steuerung im Bildungsystem?

Befunde aus dem BMBF-Förderschwerpunkt Steuerung im Bildungssystem (SteBis)

Evidence-Based Governance in the Educational System?

Results from the BMBF Funding Priority Governance in the Educational System (SteBis)

In den letzten Jahren ist die Vorstellung einer „evidenzbasierten Steuerung" des Bildungssystems zum vielfach formulierten bildungspolitischen Anspruch avanciert. Unklar ist hierbei jedoch, um *welche Form* der Evidenz es sich handeln soll und *wie* durch Evidenzen gesteuert werden könnte. Neuralgisch für die Klärung dieser Fragen ist der Blick auf Entscheidungsprozesse, da der Programmatik nach in ihnen maßgeblich die so genannte „Neue Steuerung des Bildungssystems" verankert ist. Der Blick richtet sich damit auf die Akteure, die durch ihre evidenzbasierten Entscheidungen (anders) steuern sollen, sowie auf die anvisierte neue Form der Entscheidungsprozesse und damit verbunden auf die Frage, inwieweit diese durch Evidenzen, d.h. in diesem Fall empirische Befunde der Bildungsforschung, geleitet sind.

Richtet man zunächst den Blick auf die Akteure im Schulsystem, so fällt auf, dass die „Neue Steuerung" im Bildungssystem einhergeht mit einer Umschichtung von Entscheidungskompetenzen zwischen Bildungspolitik, Bildungsadministration und Schulen. Das „Neue" der Steuerung lässt sich demnach nur zureichend bestimmen, wenn nicht mehr eine lineare Vorstellung einer hierarchischen Steuerung angenommen wird, sondern eine Vielfalt handlungskoordinierender und vermittelt darüber steuernder Akteure im Mehrebenensystem.

Die sich in der Praxis durchsetzende Logik der Umschichtungsprozesse beruht dabei auf zwei unterschiedlichen, aber nicht zufällig gemeinsam auftretenden, sondern aufeinander verwiesenen politischen Modernisierungskonzepten: Auf der einen Seite erfolgte im Rahmen des *New Public Management* in nahezu allen Bereichen der öffentlichen Verwaltung eine stärkere Differenzierung von strategischen und operativen Entscheidungen; auf der anderen Seite wurde mit der Definition und Operationalisierung von Bildungsstandards und der Entwicklung entsprechender Leistungstests *Accountability* zu einem zentralen Prinzip bildungspolitischer Steuerung. Die in diesen beiden Konzepten enthaltene Interdependenz be-

steht in einer Neuordnung der Entscheidungsprozesse: Während die Bildungspolitik in der Logik des neuen Steuerungskonzepts idealtypisch allgemeine Standards oder Leistungserwartungen definiert sowie das Monitoring des Gesamtsystems und seiner Teile verantwortet, sehen sich Schulen nun mit der Erwartung konfrontiert, für eine bedarfsgerechte Zielformulierung und Maßnahmenentwicklung sowie die entsprechende Überprüfung der Wirksamkeit von Maßnahmen in erheblichem Umfang selbst Sorge zu tragen. Schulaufsicht soll diese Entwicklungsprozesse durch geeignete Impulse zur Selbstverpflichtung sowie durch Beratungsleistungen unterstützen. Mit der Schulinspektion tritt zudem ein neuer Akteur auf den Plan, der die Schulen, und in einigen Ländern auch die Schuladministration, mit Daten über Prozesse und Wirkungen des Schulmanagements und des Unterrichts versorgt. Die Erweiterung der Entscheidungsspielräume der Schulen geht dabei nicht nur mit internen Reorganisationsprozessen, wie etwa einer Stärkung der strategischen Spitze oder der Ausdifferenzierung neuer Aufgabenbeschreibungen, einher, sondern auch mit einer Veränderung und Dynamisierung schulischer Umweltbeziehungen. Einerseits adressieren Eltern ebenso wie die Abnehmer von Qualifikationen ihre Erwartungen unter den Bedingungen der neuen Steuerung sehr viel direkter an Schulen und berufen sich dabei zunehmend auf unterschiedliche Datenquellen. Auf der anderen Seite verändert sich das Verhältnis zwischen Einzelschule und Schulaufsicht. Letztere wird nicht zuletzt durch die Etablierung der Schulinspektion zu einer Neujustierung ihrer Rolle im Spannungsbogen von Kontrolle und Beratung herausgefordert. Von der Schulaufsicht werden heute auch beratende und unterstützende Tätigkeiten erwartet, da die „Eingriffsaufsicht" unter Druck geraten ist. Es verstärken sich die Bestrebungen hin zu einer „Beratungsaufsicht"; damit soll Schulaufsicht zunehmend zur Qualitätssicherung beitragen. Deutlich wird im Überblick, dass sich damit durch die Umschichtung der Entscheidungskompetenzen eine ganz neue – und nicht immer ganz einfach zu sortierende – Akteurkonstellation im steuernden Zusammenwirken ergibt.

Als zentrales Ordnungsprinzip innerhalb jener „neuen Unübersichtlichkeit" in der Akteurkonstellation der Neuen Steuerung soll im Sinne einer Rationalisierung der gemeinsamen Handlungen die Vorstellung einer Evidenzbasierung dienen. Durch die Rückbindung der Entscheidungen auf und in allen Ebenen des Mehrebenensystems an Evidenzen soll es möglich werden, alle Akteure auf ein rationales Handeln zu verpflichten, um hierüber im Gesamtsystem eine kohärente und vermittelt darüber auch effektivere und effizientere Handlungskoordination zu evozieren. Die Vorstellung des *data-based decision making* (vgl. Schildkamp/Ehren/Lai 2012) wird damit zum zentralen Prinzip des neuen Steuerungskonzepts. Dass Planungsentscheidungen mit Daten unterlegt sind, wird ebenso selbstverständlich erwartet wie die kontinuierliche datengestützte Überprüfung aller Interventionen und Maßnahmen. Diese Erwartung gilt für Bildungspolitik, Bildungsadministration, Schulen und Lehrkräfte gleichermaßen. Zwar bleibt die Nutzung von Leistungsdaten (vgl. z.B. O'Day 2002), Ergebnissen externer Evaluation (im Überblick: Husfeldt 2011) sowie interner Evaluation (vgl. z.B. Hofman/Dijkstra/Hofman 2009) im Allgemeinen hinter den hohen Erwartungen zu-

rück; gleichwohl belegen einige Studien, dass Nutzen und Güte der extern generierten Daten von Schulleitungen und teilweise auch von Lehrkräften unter bestimmten Bedingungen positiv eingeschätzt werden (vgl. Maier 2008; Gärtner/Hüsemann/Pant 2009; Dedering/Müller 2011), auch wenn damit noch keine verbindliche Aussage über die tatsächliche Nutzung bzw. letztlich die Verbesserung von Unterricht und Lernerträgen gemacht werden kann.

Die unklare Befundlage erzeugt einerseits einen Bedarf an differenzierenden Untersuchungen, die nationale und föderale Implementationsbedingungen von Instrumenten der neuen Steuerung im Bildungssystem in Rechnung stellen – so unterscheiden sich beispielsweise Ziele von Leistungstests unter *low-stakes-* und *high-stakes*-Bedingungen erheblich; ebenso macht ein systematischer Vergleich der Schulinspektion oder der internen Evaluation im föderalen Rahmen große Unterschiede zwischen den landesspezifischen Rahmenbedingungen deutlich. Andererseits besteht ein großer Bedarf an Studien, die die Handhabung der unterschiedlichen Instrumente im Zusammenhang untersuchen.

Die Enttäuschung überschießender Erwartungen legt es außerdem nahe, dem Verhältnis von Aufwand und Ertrag, das sich aus den verschiedenen Akteursperspektiven ganz unterschiedlich darstellt, stärkere Beachtung zu schenken. Dass neue Entscheidungskompetenzen unter Umständen auch als Entscheidungslasten wahrgenommen werden, ist sowohl für Schulleitungen als auch für Lehrkräfte augenfällig. Untersuchungen, die der Theorieperspektive des Neoinstitutionalismus folgen, haben zudem deutlich gemacht, dass sich der Umgang mit Steuerungszumutungen nur als ein Prozess des „*sense-making*" (vgl. Coburn 2004) wirklich verstehen lässt. Daraus lässt sich die Aufforderung an die Forschung ableiten, die praktische Nutzung (oder Nicht-Nutzung) von Daten stärker aus der Perspektive der Akteure zu rekonstruieren.

Ein weiterer Aspekt, dem die bisherige Forschung in den deutschsprachigen Ländern bislang zu wenig Aufmerksamkeit geschenkt hat, ist der Tatsache geschuldet, dass sich die Implementation des neuen Steuerungskonzepts für Schulen in schwieriger Lage anders darstellt als für die durch Kompositions- und Umweltfaktoren weniger belasteten Schulen. Zu wenig Beachtung findet bisher auch, dass sich die Implementation des neuen Steuerungskonzepts in einigen deutschen Bundesländern gleichzeitig mit einem Umbau der Systemstruktur und daraus resultierenden Schulfusionen vollzieht.

Obwohl im deutschen Sprachraum in den vergangenen Jahren einige Einzelstudien zur Implementation der neuen Steuerungsinstrumente vorgelegt wurden, besteht nach wie vor ein hoher Forschungsbedarf. Dies gilt insbesondere für Studien, die die in den vorausgehenden Abschnitten identifizierten Anforderungen an die Steuerungsforschung aufnehmen:

- die kontrastive Untersuchung einzelner Instrumente unter unterschiedlichen (föderalen) Implementationsbedingungen,
- die Untersuchung der Handhabung unterschiedlicher Instrumente (Schulleistungstests, Inspektion, interne Evaluation …) und die Nutzung unterschiedlicher Daten (Schulleistungsdaten, Inspektionsdaten, Daten aus interner Evaluation, andere Evidenzquellen) im Zusammenhang,
- die Untersuchung der Veränderung der Beziehungen zwischen Akteuren, insbesondere der Beziehung zwischen Schulleitung und Schulaufsicht, infolge der Implementation des neuen Steuerungskonzepts (Interdependenzmanagement),
- die Rekonstruktion von Akteursperspektiven – sowohl was Belastungsfaktoren als auch was *Sense-making*-Prozesse betrifft,
- die Beachtung der konkreten Schulsituation, insbesondere der Anforderungen, die aus einer spezifischen Komposition der Schülerschaft sowie schulstrukturellen Rahmenbedingungen resultieren.

Mit diesem Heft präsentieren wir eine Reihe von double-blind-reviewed Beiträgen, die diese Desiderate bearbeiten. Alle Studien werden im Schwerpunkt „Steuerung im Bildungssystem" (SteBis) des Rahmenprogramms Empirische Bildungsforschung durch das Bundesministerium für Bildung und Forschung (BMBF) gefördert. Erstmalig wurde mit diesem Förderschwerpunkt im deutschsprachigen Raum ein koordiniertes Forschungsprogramm zum Thema Steuerung und Qualitätsentwicklung im Schulsystem aufgelegt. Damit wird einem im internationalen Vergleich deutlich zu Tage tretenden Nachholbedarf Rechnung getragen (vgl. hierzu den Bericht von Thiel, Thillmann und Tarkian in diesem Heft).

Das Beiheft wird eingeleitet durch zwei Beiträge, in denen die Wirkung mehrerer neuer Steuerungsinstrumente und die Nutzung entsprechender Daten durch Schulen im Gesamtzusammenhang *vergleichend* untersucht werden.

Sebastian Wurster, Dirk Richter, Anna Schliesing und Hans Anand Pant beschäftigen sich mit der Frage, wie Schulinspektion, Vergleichsarbeiten und interne Evaluation von den Schulleitungen hinsichtlich Aufwand, Güte und Nutzen eingeschätzt werden und welche verfahrensübergreifenden Nutzungskonfigurationen sich zeigen. Ein wichtiges Ergebnis der Studie ist, dass Schulinspektionsrückmeldungen, VERA-Daten und interne Evaluation in einem Teil der Schulen miteinander im Zusammenhang stehen, offensichtlich weil interne Evaluation als eine Zielvorgabe im Schulqualitätsrahmen der Schulinspektion genannt wird.

Isabell van Ackeren und ihre Kolleginnen und Kollegen der Autorengruppe EviS (Evidenzbasierte Schulentwicklung) fragen in ihrem Beitrag, welche unterschiedlichen Daten und „Evidenzquellen" Schulen nutzen und wie die Daten von Schulleitungen und Lehrkräften hinsichtlich ihrer Nützlichkeit eingeschätzt werden. Die Ergebnisse zeigen, dass Daten aus Lernstandserhebungen und Daten der Schulinspektion im

Vergleich mit internen Informationsquellen weniger genutzt werden. Die Autoren und Autorinnen markieren weiteren Forschungsbedarf insbesondere hinsichtlich der Frage, ob die von den Schulen stärker genutzten internen Informationsquellen den Gütestandards entsprechen und ob sie systematisch für die Entwicklung genutzt werden.

Die nachfolgenden Beiträge widmen sich dann vertiefend *einzelnen Instrumenten* der neuen Steuerungslogik. Dabei geht es zunächst um die schulische Rezeption von *Vergleichsarbeiten* bzw. *Lernstandserhebungen*, auch unter der Perspektive der Handlungskoordination verschiedener Akteure im Schulsystem.

Uwe Maier, Carolin Ramsteck und Annette Frühwacht beschäftigen sich in ihrem Beitrag mit der Interpretation der Vergleichsarbeitsrückmeldungen durch Lehrkräfte. Sie schließen mit ihrem rekonstruktiven Ansatz eine Lücke in der Rezeptionsforschung, indem sie das lehr-lerntheoretische Argumentationsmuster, das den Horizont für die Interpretation der Daten absteckt, im Zusammenhang mit der praktischen Nutzung untersuchen. Es wird gezeigt, dass dieselben Daten vor dem Hintergrund unterschiedlicher lehr-lerntheoretischer Konzepte in ganz unterschied licher Weise für die Unterrichtsentwicklung genutzt werden.

Barbara Muslic, Carolin Ramsteck und Harm Kuper gehen – ebenfalls im Kontext der Lernstandserhebungen – der Veränderung der Beziehungen zwischen Schulen und Schulaufsicht nach. Die Autorinnen und der Autor zeigen, wie die schulische Organisation und der Managementstil der Schulaufsicht ineinandergreifen. Lernstandserhebungen können vor allem in einer gemanagten professionellen Organisation produktiv verarbeitet werden, wenn diese von einer Schulaufsicht des Typs Neues Management unterstützt wird. In autonomen professionellen Organisationen läuft die schulaufsichtliche Strategie des neuen Managements dagegen ins Leere. Eine Schulaufsicht, die distanziertes Management betreibt, lässt dagegen die Unterstützungserwartungen einer gemanagten professionellen Organisation oft unerfüllt.

Das Instrument der *Schulinspektion* wird in zwei weiteren Texten einerseits unter der Fragestellung der Wirkungen für Schul- und Unterrichtsentwicklung in einem Bundesländervergleich, andererseits in der Fokussierung der Akteursperspektiven von Schule und Schulträger analysiert.

Oliver Böhm-Kasper und Odette Selders untersuchen die Wahrnehmung und Akzeptanz von Schulinspektionsverfahren in vier Bundesländern, die sich im Hinblick auf zentrale Merkmale unterscheiden. Ein Ergebnis der Untersuchung lautet, dass die Schulinspektion in allen vier Bundesländern kaum Auswirkungen auf die Schul- und Unterrichtsentwicklung hat. In einem Bundesland, in dem eine Nachinspektion von leistungsschwachen Schulen vorgesehen ist, fällt die Akzeptanz des Verfahrens bei der Schulleitung deutlich geringer aus und auch der Nutzen wird geringer veranschlagt als

in den Ländern ohne Nachinspektion. Die Androhung von Konsequenzen führt nicht zu messbar stärkeren Aktivitäten der Schul- und Unterrichtsentwicklung.

Auch Bianca Preuß analysiert die Governance-Beziehung von Schulträger und Schule; den spezifischen Handlungskontext stellt hier die Einführung von Schulinspektionen dar. Auf der Grundlage von qualitativen Experteninterviews wird herausgearbeitet, dass für den Schulträger die Frage nach seinem Steuerungseinfluss eine Grundsatzfrage ist; dabei überwiegt noch die Wahrnehmung der Rolle als Beobachter des Geschehens. Insgesamt weist die Untersuchung aber auf veränderte Deutungs-, Legitimations- und Herrschaftsmuster hin, indem sich Schulträger nicht mehr allein auf die gesetzlich determinierte Rolle berufen, sondern ein Bündnis mit Schule angestrebt wird, das auf die Schaffung von Bedingungen guter Schule zielt.

Im Anschluss an die Analyse zentraler neuer Steuerungsinstrumente, die der externen Evaluation zuzuordnen sind, wird im weiteren Verlauf die ebenfalls geforderte Perspektive der Stärkung *interner Evaluationsmaßnahmen* in den Blick genommen; damit ist zugleich die *erweiterte Verantwortung von Schule* als Linie der Schulentwicklung angesprochen.

Der Beitrag von Stefan Brauckmann und Christoph Hermann ist der Frage gewidmet, wie die Erweiterung des Aufgabenspektrums von Schulleitungen und die entsprechende zeitliche Beanspruchung sowie Belastung durch einzelne Tätigkeiten von Schulleitungen wahrgenommen werden. Interessant ist, dass zeitliche Beanspruchung und empfundene Belastung in den meisten Fällen deutlich voneinander abweichen. Nicht die zeitliche Beanspruchung, sondern vielmehr die Bewertung der Tätigkeit und der schulspezifische Problemdruck scheinen das subjektive Belastungserleben der Schulleitungen zu erklären.

Katja Thillmann, Anabel Bach, Christina Gerl und Felicitas Thiel stellen Ergebnisse einer Dokumentenanalyse schulischer Evaluationsprogramme zweier Bundesländer dar, die sich hinsichtlich der Implementationsbedingungen interner Evaluation unterscheiden. Neben den Implementationsbedingungen spielen schulische Merkmale eine Rolle für die Qualität der Evaluationsprogramme. Die Autorinnen kommen zu dem Ergebnis, dass eine klare Steuerungsstrategie, die auf konkrete Vorgaben, gezielte Unterstützung und regelmäßige Rechenschaftspflicht setzt, die Qualität der Programme maßgeblich beeinflussen kann. Was die schulischen Merkmale betrifft, kommt insbesondere der Schulleitung eine wichtige Bedeutung zu.

Die letzten beiden Beiträge zum inhaltlichen Kern des SteBis-Verbundes explizieren schließlich die spezifische Bedeutung von *lokalen Rahmenbedingungen schulischer Entwicklungsarbeit* im Kontext neuer Steuerung.

Kathrin Racherbäumer, Christina Funke, Isabell van Ackeren und Marten Clausen untersuchten Schulen, die unter schwierigen sozialen Kontextbedingungen arbeiten. Sie zeigen, dass Schulleitungen erwartungswidrig guter Schulen ein deutliches fachliches Interesse an einer datengestützten Schulentwicklung haben, während Schulleitungen an erwartungskonform abschneidenden Schulen eher eine formale Auseinandersetzung mit Leistungsdaten initiieren. Was den Führungsstil der Schulleitungen betrifft, so konnte in erwartungswidrig guten Schulen eine signifikant stärkere Ausprägung eines partizipativen und organisationsbetonten Führungsstils ausgemacht werden als an erwartungskonform abschneidenden Schulen.

Im Beitrag von Laura Fölker kann gezeigt werden, wie Schulentwicklungsprozesse durch Schulreformprozesse überlagert werden. Die Autorin beschäftigt sich mit Schulfusionen und zeigt, dass konfliktreiche Fusionsprozesse die Schulentwicklung lähmen und eine integrative Fusion unter günstigen Bedingungen die Schulentwicklung vorantreiben kann.

Nach den referierten Beiträgen mit Ergebnissen aus den Forschungsprojekten schließt das Beiheft mit einem Bericht zum Förderschwerpunkt „Steuerung im Bildungssystem" aus der Perspektive der SteBis-Koordinationsstelle. Aufgaben, Strukturen und Wirkungen von Koordinationsstellen werden grundlegend diskutiert. Nicht zuletzt wird die Besonderheit des SteBis-Verbundes reflektiert: Das Modell einer evidenzbasierten Steuerungslogik wird in seinen unterschiedlichen Facetten erforscht, unterliegt aber zugleich auch selbst dem durch die Schaffung einer Koordinationsstelle markierten Anspruch der Erhöhung des *Science* und *Practice Impact* in diesem wissenschaftlichen Arbeitsfeld.

Felicitas Thiel/Martin Heinrich/Isabell van Ackeren

Literatur

Coburn, C.E. (2004): Beyond Decoupling. Rethinking the Relationship between the Institutional Environment and the Classroom. In: Sociology of Education 77, S. 211-244.

Dedering, K./Müller, S. (2011): School Improvement through Inspections? First Empirical Insights from Germany. In: Journal of Educational Change 12, H. 3, S. 301-322.

Gärtner, H./Hüsemann, D./Pant, H.A. (2009): Wirkungen von Schulinspektion aus Sicht betroffener Schulleitungen. Die Brandenburger Schulleiterbefragung. In: Empirische Pädagogik 23, H. 1, S. 1-18.

Hofman, R.H./Dijkstra, N.J./Hofman, W.H.A. (2009): School Self-Evaluation and Student Achievement. In: School Effectiveness and School Improvement 20, H. 1, S. 47-68.

Husfeldt, V. (2011): Wirkungen und Wirksamkeit der externen Schulevaluation. Überblick zum Stand der Forschung. In: Zeitschrift für Erziehungswissenschaft 14, H. 2, S. 259-282.

Maier, U. (2008): Vergleichsarbeiten im Vergleich – Akzeptanz und wahrgenommener Nutzen standardbasierter Leistungsmessungen in Baden-Württemberg und Thüringen. In: Zeitschrift für Erziehungswissenschaft 11, H. 3, S. 453-474.

O'Day, J. (2002): Complexity, Accountability, and School Improvement. In: Harvard Educational Review 72, H. 3, S. 239-329.

Schildkamp, K./Ehren, M./Lai, M.K. (2012): Editorical Article for the Special Issue on Data-based Decision Making around the World: from Policy to Practice to Results. In: School Effectiveness and School Improvement 23, H. 2, S. 123-131.

STEUERUNGSINSTRUMENTE UND EVIDENZQUELLEN IM VERGLEICH

DDS – Die Deutsche Schule
Beiheft 12, S. 19-50
© 2013 Waxmann

Sebastian Wurster/Dirk Richter/Anna Schliesing/Hans Anand Pant

Nutzung unterschiedlicher Evaluationsdaten an Berliner und Brandenburger Schulen

Rezeption und Nutzung von Ergebnissen aus Schulinspektion, Vergleichsarbeiten und interner Evaluation im Vergleich

Zusammenfassung
Evaluationen, wie Schulinspektion, Vergleichsarbeiten und interne Evaluation, sollen als Impuls für Qualitätsentwicklung in Schulen dienen. Die vorliegende Studie analysiert, wie die verschiedenen Evaluationsverfahren von Schulleitungen (N=329) hinsichtlich der Dissemination und Nutzung der Ergebnisse sowie des Aufwandes, der Nützlichkeit und der Diagnosegüte eingeschätzt werden. Es werden ebenso verfahrensübergreifende Nutzungskonfigurationen analysiert.
Schlüsselwörter: Evaluation, Schulinspektion, Vergleichsarbeiten, interne Evaluation, Datennutzung, Schulleitung

Use of Various Evaluation Data in Schools in the Federal States of Berlin and Brandenburg

A Comparison of the Feedback Reception and Use of School Inspection, State-wide Tests and Self-Evaluation

Summary
Evaluations, such as school inspection, state-wide tests, and self-evaluation, should serve as a stimulus for quality development in schools. This study investigates how principals (N=329) evaluate the various types of evaluation concerning the dissemination and use of the results as well as the effort, benefit and diagnostic quality of the procedures. Furthermore, configurations of feedback use across evaluation procedures are analyzed.
Keywords: evaluation, school inspection, state-wide tests, self-evaluation, feedback data use, principals

1. Einleitung

In der vergangenen Dekade sind im deutschen Bildungssystem auf vielen Ebenen Qualitätsstandards formuliert und, darauf bezogen, neue Formen der Evaluation schulischer Qualität eingeführt worden. Diese Instrumente sind Ausdruck der Neuausrichtung der Steuerung im Schulwesen. Diese Entwicklung geht auf den so genannten Konstanzer Beschluss der Kultusministerkonferenz (KMK) von 1997 zurück, der u.a. vorsieht, Instrumente zur Evaluation zu entwickeln und regelmäßig länderübergreifende Vergleichsuntersuchungen zum Leistungsstand von Schülerinnen und Schülern durchzuführen. Er stellt eine Reaktion auf die Ergebnisse internationaler Schulleistungsstudien dar, die die KMK veranlasst haben, Qualitätssicherung zum zentralen Thema zu machen (vgl. KMK 2006). Neben einem nationalen Bildungsmonitoring in Form einer Bildungsberichterstattung und der kontinuierlichen Teilnahme an nationalen und internationalen Schulleistungsstudien wurden verschiedene Instrumente zur Evaluation schulischer Arbeit in den Ländern eingeführt. Dazu zählen die von der KMK beschlossenen, bundesweit durchgeführten Vergleichsarbeiten (VERA) sowie die nicht zentral von der KMK eingeführte, in den einzelnen Bundesländern ausgestaltete, externe Evaluation von Schulen durch Schulinspektionen sowie die gesetzliche Verpflichtung zur Selbstevaluation. Die eingeführten Verfahren evaluieren auf Basis ebenfalls neu formulierter Qualitätskriterien unterschiedliche Aspekte schulischer Arbeit. Zum einen wurden die KMK-Bildungsstandards eingeführt, die beschreiben, welche Kompetenzen Schülerinnen und Schüler zu einem bestimmten Zeitpunkt in ihrer Schulkarriere erreicht haben sollen. Sie wurden von jedem Land spezifisch (z.B. in Form neuer Rahmenlehrpläne) implementiert. Zum anderen wurden in den Ländern Schulqualitätsrahmen entwickelt, die definieren, was unter „guter Schule" verstanden werden soll. Schulen bekommen folglich Rückmeldungen aus verschiedenen Informationsquellen, die potenziell alle als Evaluation ihrer Arbeit angesehen werden können (vgl. Maag Merki 2009) und damit ein möglicher Ausgangspunkt für Qualitätssicherung und Qualitätsentwicklung von Schulen sind.

Ausgehend von der zugedachten Funktion, Qualitätsentwicklung in Schulen anzustoßen, stellt sich die Frage, inwiefern dies den unterschiedlichen Evaluationsverfahren gelingt. Bislang wurden verschiedene Evaluationsinstrumente zumeist getrennt voneinander empirisch analysiert, ohne diese miteinander in Beziehung zu setzen und verfahrensübergreifend bzw. vergleichend zu untersuchen. Diese Studie betrachtet exemplarisch für die Länder Berlin und Brandenburg die Evaluationsverfahren VERA, Schulinspektion und interne Evaluation hinsichtlich ihrer Nutzung für Entwicklungsmaßnahmen sowie der Akzeptanz durch die Evaluationsbetroffenen.

2. Evaluation im Deutschen Bildungssystem

2.1 Evaluationsverfahren als Teil der Steuerung des Bildungssystems

Die Einführung von Schulinspektionen, VERA und interner Evaluation ist Ausdruck der „neuen Steuerung" im Schulsystem. Altrichter (vgl. 2010, S. 220f.) fasst die Leitideen der „neuen Steuerung" in den folgenden vier Merkmalen zusammen: 1. Evidenzbasierung, 2. Outputorientierung, 3. Unterrichtsfokussierung und 4. Zielorientierung (vgl. Klieme 2004). Entwicklungsentscheidungen sollen auf geprüften Informationen basieren (Evidenzbasierung), und die Umsetzung wird empirisch evaluiert. Die Erträge von Bildungseinrichtungen haben einen hohen Stellenwert (Outputorientierung). Diese werden durch Unterricht erbracht und durch die Messung von Schülerleistungen überprüft (Unterrichtsfokussierung). Die intendierten Entwicklungsziele sollen den Akteuren des Bildungssystems deutlich kommuniziert werden, und ihre Tätigkeit soll darauf ausgerichtet sein (Zielorientierung). Ein wesentliches Element in der „neuen Steuerung" ist dabei die Verknüpfung von Standards bzw. Qualitätskriterien mit einer Überprüfung ihrer Erfüllung, z.B. durch Tests (vgl. Altrichter 2010). Im Folgenden werden die Evaluationsverfahren kurz umrissen. Die Darstellung erfolgt für die Ausgestaltung der Verfahren in den Ländern Berlin und Brandenburg, aus denen die Teilnehmer und Teilnehmerinnen der vorliegenden Studie rekrutiert wurden.

2.1.1 Vergleichsarbeiten

Standardisierte Lernstandserhebungen bzw. Vergleichsarbeiten (VERA) ermöglichen es festzustellen, inwiefern die in den Bildungsstandards für das Ende der vierten Jahrgangsstufe bzw. das Ende der Sekundarstufe I definierten Kompetenzen bereits vorher von Schülerinnen und Schülern einer bestimmten Jahrgangsstufe erreicht wurden. Vergleichsarbeiten sollen folglich der Standortbestimmung dienen und damit Ausgangspunkt für kompetenzorientierte Entwicklungs- und Fördermaßnahmen sein. Bislang finden sie länderübergreifend in den Jahrgangsstufen 3 und 8 statt und werden in den Fächern Deutsch, Mathematik und in weiterführenden Schulen zusätzlich in der ersten Fremdsprache (Englisch oder Französisch) durchgeführt. Die Testung findet ein Jahr bzw. zwei Jahre früher statt, als die Erreichung der Bildungsstandards vorgesehen ist (Primarstufe in Klasse 4; Sekundarstufe I in Klasse 9 für den Hauptschulabschluss bzw. 10 für den Mittleren Schulabschluss), um den Lehrkräften frühzeitig diagnostische Anhaltspunkte über den erreichten Kompetenzstand der Lerngruppe zurückzumelden. Die Ergebnisse von VERA dienen primär der Unterrichtsentwicklung; sie sollen jedoch auch über die Klassenebene hinaus zur Schulentwicklung genutzt werden (vgl. KMK 2010). Auf Schulebene können Vergleiche zwischen Lerngruppen sowie Vergleiche mit Referenzgruppen Anlass für Schulentwicklungsmaßnahmen sein. In Berlin und Brandenburg bekommen Schulen

angepasste Rückmeldungen für verschiedene Ebenen (Individualebene, Klassenebene und Schulebene; vgl. Pant u.a. 2011). Außer in ihrer primären Entwicklungsfunktion werden die VERA-Ergebnisse in einigen Ländern darüber hinaus zum regionalen Monitoring des Schulsystems eingesetzt.

2.1.2 Schulinspektion

Die Schulinspektion stellt ein externes Evaluationsverfahren dar, bei dem schulfremde Inspektorinnen und Inspektoren Schulen begutachten. Im Zentrum des Inspektionsverfahrens steht die Erfassung der Unterrichts- und Schulqualität, also der schulischen Prozesse (vgl. Döbert/Dedering 2008; Dedering/Müller 2011). Die Bewertung der Schulen erfolgt auf Basis von Orientierungsrahmen für Schulqualität, die die (normative) Zielvorstellung einer „guten Schule" beschreiben (vgl. z.B. Brandenburger Orientierungsrahmen zur Schulqualität (MBJS 2008a) oder Berliner Handlungsrahmen (SenBJS 2006)). Die Orientierungsrahmen enthalten als Qualitätsdimensionen neben den Erträgen einer Schule (z.B. Kompetenzen der Schülerinnen und Schüler) hauptsächlich Prozessdimensionen wie z.B. Schulmanagement, Maßnahmen der Qualitätsentwicklung und insbesondere die Unterrichtsqualität. Als Grundlage für die Bewertung der Schulen werden Daten mit verschiedenen Methoden erhoben (z.B. Unterrichtsbeobachtung, Interviews, schriftliche Befragungen, Dokumentenanalyse), die zu einem Stärken-Schwächen-Profil der erfassten Qualitätskriterien zusammengeführt werden. Die Rückmeldung der Ergebnisse soll als Ausgangspunkt zur weiteren Qualitätsentwicklung dienen, z.B. in Zusammenarbeit mit der Schulaufsicht durch die gemeinsame Formulierung von Entwicklungszielen. Andere Evaluationsverfahren (z.B. Ergebnisse aus VERA, Durchführung interner Evaluation) sind Teil der Bestandsaufnahme der Schulinspektion.

2.1.3 Interne Evaluation

Unter interner Evaluation bzw. Selbstevaluation werden zahlreiche Verfahren subsumiert, die es der Schule ermöglichen sollen, Erkenntnisse über den Erfolg der eigenen Arbeit zu gewinnen (vgl. Berkemeyer/Müller 2010). Die Verantwortung für die Gestaltung und Durchführung liegt bei der einzelnen Schule. Interner Evaluation wird deshalb eine große Bedeutung hinsichtlich professionellen Lehrerhandelns und einer lernenden Organisation zugeschrieben (vgl. ebd.; Thiel/Thillmann 2012). Anwendungsfelder interner Evaluation sind beispielsweise Bestandsaufnahmen zu schulischen Prozessen (z.B. mittels standardisierter Instrumente wie „Selbstevaluation in Schulen" (SEIS); vgl. Stern/Ebel/Müncher 2008), Feedback zum Unterricht mittels standardisierter Instrumente (z.B. das „Selbstevaluationsportal" (SEP); vgl. Gärtner 2010) oder Selbstevaluation als eigenständige individuelle Reflexion und Weiterentwicklung des eigenen Unterrichts (vgl. Berkemeyer/Müller 2010).

Den genannten Evaluationsverfahren VERA, Schulinspektion und interne Evaluation ist gemeinsam, dass sie Wissen über Aspekte schulischer Arbeit generieren sollen, wobei das generierte Wissen differiert (vgl. Tab. 1) und Schulen unterschiedliche Informationen erhalten. Dies verdeutlicht für alle Verfahren gleichermaßen die Evidenzbasierung als Merkmal der „neuen Steuerung". Die Zielorientierung zeigt sich durch die Bildungsstandards (VERA) und Qualitätsrahmen (Schulinspektion), die jeweils klare Bezüge zu schulischen Erträgen (Outputorientierung) und ebenso – zumindest partiell – einen Fokus auf Unterrichtsprozesse enthalten. Tabelle 1 stellt zur besseren Vergleichbarkeit die Verfahren gegenüber, um Unterschiede in den *Durchführungsmodalitäten* und *Evaluationsgegenständen* aufzuzeigen.

Tab. 1: Unterschiede zwischen den Evaluationsverfahren am Beispiel Berlin und Brandenburg

	Vergleichsarbeiten	**Schulinspektion**	**Interne Evaluation am Beispiel SEP**[1]
Turnus	jährlich	ca. alle 5 Jahre	Entscheidung durch Schule; teilw. gesetzliche Vorgaben
Datenerhebung	standardisierte Leistungstests (Korrektur und Dateneingabe durch Lehrkräfte)	Beobachtung, standardisierte Fragebögen, Interviews, Dokumentenanalyse	Onlinebefragung mit standardisierten Fragebögen
Evaluationsgegenstand	Erreichung der Bildungsstandards	Erfüllung der Kriterien im Qualitätsrahmen	Selbst-/Fremdeinschätzung der Qualität des Schulmanagements und der Unterrichtsqualität
Rückmeldung *Form*	Schul-, Klassen-, Individualrückmeldung: Rückmeldung auf Aufgaben- und Kompetenzebene	Bericht über die erhobenen Merkmale, Stärken-Schwächen-Profil	Profilgrafik mit Vergleich der Selbst- und Fremdeinschätzung, Tipps zur Weiterarbeit
Rückmeldung *Vergleichsmöglichkeiten*	kriterial (Kompetenzstufenverteilung), sozial (Vergleichsgruppe)	kriterial (erfasste Bereiche Qualitätsrahmen); Vergleich eigene Schule mit aggregierten Werten in Landesberichten	kriterial (erfasste Qualitätsaspekte), sozial (bei ausreichend großer Vergleichsgruppe)
Rückmeldung *Stellungnahme*	nein	möglich	nein

1 Am Beispiel des vom Institut für Schulqualität der Länder Berlin und Brandenburg seit 2008 bereitgestellten „Selbstevaluationsportals" (SEP; vgl. Gärtner 2010) wird hier eine Möglichkeit zur Selbstevaluation dargestellt.

Quelle: eigene Darstellung

Die Verfahren unterscheiden sich in Bezug auf die evaluierten Inhalte und in der Durchführung. Wie im folgenden Abschnitt beschrieben, ist jedoch die Art und Weise, wie sie in den Schulen Entwicklungsprozesse stimulieren sollen, ähnlich.

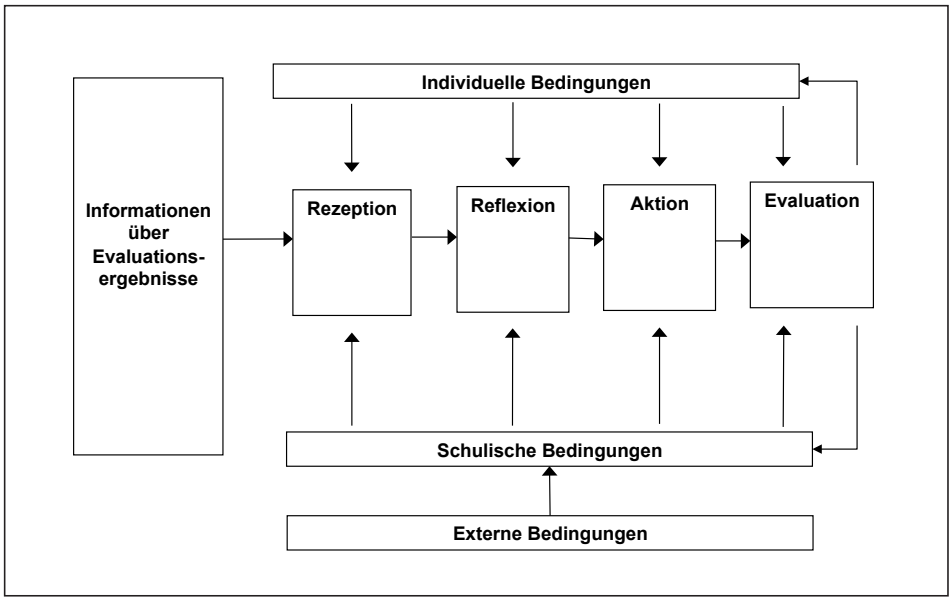
2.2 Intendierte Funktionen und Wirkungsweise der Evaluationsverfahren

Die Funktionen der Evaluationsverfahren sind als Mix aus Wissensgewinnung, Orientierung an Zielvorgaben, Rechenschaftslegung sowie Schul- und Unterrichtsentwicklung beschreibbar (vgl. Altrichter 2010; Landwehr 2011; Maritzen 2006). Die Funktion der *Wissensgewinnung* zeigt sich darin, dass alle Evaluationsverfahren Informationen über ihren Evaluationsgegenstand generieren. Die *Orientierung an Zielvorgaben* wird durch die der Evaluation zugrunde liegenden Bezugsrahmen deutlich. Während die Schulinspektion auf den länderspezifischen Qualitätsrahmen Bezug nimmt und VERA auf die nationalen Bildungsstandards, basiert interne Evaluation in ihrer sehr heterogenen Ausgestaltungsmöglichkeit nicht zwangsläufig auf expliziten Bezugsrahmen. Die *Rechenschaftslegung* bzw. *Kontrollfunktion* soll durch die Kontrolle der Erreichung extern gesetzter Kriterien bzw. Zielvorgaben erfüllt werden, indem die Bildungsverwaltung bzw. schulische Akteure über die in der einzelnen Schule erreichten Leistungen und die Erfüllung von Qualitätskriterien informiert werden. Die Evaluationsverfahren unterscheiden sich hinsichtlich ihrer Adressaten. VERA-Ergebnisse können innerhalb der Schule Einzelpersonen zugeordnet werden; bei der Schulinspektion werden auf Schulebene aggregierte Ergebnisse zurückgemeldet. Schulgenaue Ergebnisse von Schulinspektion und VERA werden zudem an die Bildungsadministration weitergegeben.

Gemeinsam ist allen Verfahren, dass die an die Schule zurückgemeldeten Ergebnisse von den schulischen Akteuren verarbeitet und anschließend, wenn notwendig, Maßnahmen ergriffen werden sollen. Dies entspricht der Funktion der *Schul-* bzw. *Unterrichtsentwicklung*. Eine Grundlage zur Veranschaulichung der Entwicklungsfunktion und zur Auswahl von Einflussfaktoren und Nutzungsvariablen ist das Rahmenmodell zur Nutzung von Vergleichsarbeiten (vgl. Helmke 2004), das für das Instrument Schulinspektion adaptiert wurde (vgl. Sommer 2011). Das Modell ist prinzipiell auch zur Beschreibung der Nutzung von Selbstevaluationsdaten geeignet, weshalb es hier zur Beschreibung der Nutzung aller Evaluationsverfahren verwendet werden soll.

Das Zyklenmodell von Helmke (2004), dargestellt in Abbildung 1, nimmt eine sequenzielle Verarbeitung von Evaluationsergebnissen an: von der Rezeption der Ergebnisse über die Reflexion hin zur Aktion und Evaluation. Dieser Informationsverarbeitungszyklus wird von verschiedenen individuellen, schulischen und externen Bedingungen beeinflusst und bildet die idealtypische Wirkungsweise von Evaluationen ab.

Abb. 1: Zyklenmodell schulischer Evaluation



Quelle: in Anlehnung an Helmke (2004)

Die erste Verarbeitungsphase, die *Rezeption*, beschreibt das Wahrnehmen und Verstehen der Informationen aus Evaluationen; *Reflexion* bezieht sich auf die Analyse von Ergebnissen und deren Ursachen. Ausgehend vom konkreten Ergebnis können Entwicklungsmaßnahmen eingeleitet werden, sofern Handlungsbedarf festgestellt wird. Die Entscheidung kann bspw. auf einem Vergleich zwischen Soll-Zustand (Bildungsstandards, Qualitätsrahmen) und Ist-Zustand basieren. *Aktion* umfasst Maßnahmen, die infolge der Ergebnisse zur Verbesserung der Qualität von Schule und Unterricht initiiert werden. Die Beschreibung der Ergebnisverarbeitung durch Kommunikation und Reflexion, also der Auseinandersetzung mit den Ergebnissen, sowie die Ableitung von Maßnahmen finden sich in ähnlicher Form auch in anderen Modellen (vgl. Ehren/Visscher 2006; KMK 2010; Schildkamp u.a. 2012; Sommer 2011). Diese Modelle verbindet, dass sie auf der Kollektivebene (z.B. Fachgruppe, Schule als Organisation) und auf der Individualebene anwendbar sind. In der letzten Phase, der *Evaluation*, werden die durchgeführten Maßnahmen auf ihre Wirkung hin überprüft. Nach Helmke (2004) kann dies beispielsweise durch eine Wiederholung der Vergleichsarbeiten geschehen, wodurch ein Zyklus entsteht. Dies ist auf andere Evaluationsverfahren übertragbar. Resultat dieser datengestützten Entwicklungskreisläufe soll eine Verbesserung von Prozessen sein, die Einfluss auf das Lernen der Schülerinnen und Schüler haben.

Im Zyklenmodell (Helmke 2004) werden verschiedene den Verarbeitungsprozess beeinflussende Faktoren aufgeführt. *Individuelle* Faktoren sind u.a. die Akzeptanz von

Evaluation, das Vorwissen oder die Selbstwirksamkeitserwartung, Veränderungen auch erfolgreich initiieren zu können. *Externe* Bedingungen umfassen beispielsweise Art und Umfang schulexterner Unterstützung, z.B. durch Schulaufsicht oder Weiterbildung. *Merkmale der Schule* betreffen u.a. die Schulausstattung oder das Evaluations- und Kooperationsklima. Diese Faktoren lassen sich auch in anderen Modellen identifizieren (vgl. Ehren/Visscher 2006; Sommer 2011; Stamm 2003). So wird eine positive Einstellung gegenüber den Evaluationsverfahren regelmäßig als determinierend für die Ergebnisnutzung verstanden (vgl. zu VERA: Maier 2008; Tresch 2007; zu Schulinspektion: Ehren/Visscher 2006; Husfeldt 2011; zu interner Evaluation: Meuret/Morlaix 2003; Vanhoof/van Petegem/de Maeyer 2009).

2.3 Forschungsstand

In der vergangenen Dekade gab es einige verfahrensspezifische Forschungsarbeiten für die Evaluationsverfahren VERA, Schulinspektion und interne Evaluation. Anknüpfend an das beschriebene Modell wird im Folgenden der Forschungsstand, insbesondere zur schulinternen Auseinandersetzung mit Evaluationsergebnissen (Rezeption und Reflexion), zur Nutzung für Entwicklungsmaßnahmen und zur individuellen Einstellung gegenüber den Verfahren, zusammengefasst.

2.3.1 Vergleichsarbeiten

Mit Ergebnissen aus Vergleichsarbeiten bzw. Lernstandserhebungen findet in Schulen durchaus eine Auseinandersetzung durch Schulleitungen und Lehrkräfte statt (vgl. Koch u.a. 2006; Maier 2008; Nachtigall/Jantowski 2007). Häufig werden Ergebnisse darüber hinaus Schülerinnen und Schülern sowie Eltern gegenüber kommuniziert (vgl. Koch u.a. 2006; Tresch 2007). Trotz der schulischen Auseinandersetzung mit Ergebnissen aus Lernstandserhebungen mangelt es an daraus resultierenden Maßnahmen der Schul- und Unterrichtsentwicklung. Werden Maßnahmen abgeleitet, so sind diese vor allem im Bereich der Leistungsbeurteilung zu finden. Es werden eher Einzelmaßnahmen, wie z.B. die Übung bestimmter Aufgabentypen, eingeleitet als umfassende Vorhaben der Unterrichtsentwicklung. Die eingeleiteten Maßnahmen gehen selten über die Klassenebene hinaus (vgl. Groß Ophoff/Hosenfeld/Koch 2007; Koch u.a. 2006; Maier 2008; Nachtigall/Jantowski 2007).

Die bislang in Studien erfasste Einstellung der Schulleitungen und Lehrkräfte gegenüber VERA bzw. Lernstandserhebungen ist überwiegend positiv. Schulleitungen stehen Lernstandserhebungen tendenziell noch positiver gegenüber als Lehrkräfte (vgl. Maier 2008; Nachtigall/Jantowski 2007; Schrader/Helmke 2003; Tresch 2007). Die Bereitschaft zur Auseinandersetzung mit den Ergebnisrückmeldungen ist abhän-

gig von der Akzeptanz und der eingeschätzten Nützlichkeit und weniger von der Verständlichkeit der Rückmeldungen (vgl. Kühle/Peek 2007). Hosenfeld (2010) weist ebenfalls auf den Einfluss der wahrgenommenen Nützlichkeit und der Akzeptanz auf die tatsächliche Nutzung hin.

Groß Ophoff, Hosenfeld und Koch (2007) beschreiben *Rezeptionstypen* von Lehrkräften im Umgang mit Vergleichsarbeiten. Diese unterscheiden sich hinsichtlich der aus VERA-Ergebnissen abgeleiteten Maßnahmen, der wahrgenommenen Verständlichkeit und Nützlichkeit sowie der Auseinandersetzung mit den Ergebnissen. Ein Rezeptionstyp zeigt auf allen erfassten Merkmalen hohe Ausprägungen, während bei einem zweiten Typ alle Merkmale im Vergleich zum ersten Typ niedriger ausgeprägt sind. Der dritte Rezeptionstyp zeichnet sich zwar durch intensive Auseinandersetzung aus, schätzt Verständlichkeit und Nützlichkeit jedoch gering ein und zeigt die geringste Aktivität.

2.3.2 Schulinspektion

Im Bereich der Forschung zur Schulinspektion liegen in Deutschland bislang vornehmlich Befragungen von Schulleitungen und Lehrkräften zu innerschulischen Verarbeitungsprozessen, Einstellungen gegenüber dem Verfahren und Folgemaßnahmen bzw. anderen Reaktionen auf die Inspektionen vor (vgl. Dedering 2012; Husfeldt 2011). Die Auseinandersetzung mit Inspektionsergebnissen erfolgt in der Mehrheit der Schulen unter Beteiligung verschiedener Akteure, insbesondere des Kollegiums. Die Ergebnisse werden in der Lehrerkonferenz und teilweise in der Schulkonferenz kommuniziert (vgl. Böttcher/Keune/Neiwert 2010; Dedering/Müller 2011; Gärtner/Wurster 2009a, 2009b; Huber 2006).

Studien zur Nutzung der Ergebnisse für Entwicklungsaktivitäten zeigen, dass in der Mehrheit der untersuchten Schulen Maßnahmen in den Bereichen Schulprogrammarbeit, interne Evaluation und Unterrichtsentwicklung eingeleitet und umgesetzt werden (vgl. Dedering/Müller 2011; Gärtner/Hüsemann/Pant 2009; Gärtner/Wurster 2009a, 2009b; Huber 2006). Von der Inspektion festgestellte Schwächen werden in die Zielvereinbarungen mit der Schulaufsicht aufgenommen und haben so Einfluss auf die eingeleiteten Maßnahmen (vgl. Gärtner/Wurster 2009a, 2009b; Huber 2006). Studien zur Qualität der Maßnahmen liegen derzeit noch nicht vor (vgl. Husfeldt 2011). Studien zur Einschätzung des Verfahrens deuten darauf hin, dass die Schulleitungen inspizierter Schulen mehrheitlich der Inspektion gegenüber positiv eingestellt waren und dass – ähnlich wie bei der Akzeptanz von Vergleichsarbeiten – Schulleitungen tendenziell positiver eingestellt sind als Lehrkräfte (vgl. Böttcher/Keune/Neiwert 2010; Dedering/Müller 2011; Gärtner/Wurster 2009a, 2009b; Huber 2006; Sommer 2011).

Eine differenzielle Betrachtung zeigt, dass auch auf Schulebene verschiedene *Typen* im Umgang mit Schulinspektionen und deren Ergebnissen beschrieben werden können (vgl. Wurster/Gärtner 2013). So finden sich Schulen, die sehr aktiv Maßnahmen ergreifen und eine positive Einstellung gegenüber dem Verfahren haben, sowie Schulen, bei denen das Gegenteil zutrifft. Daneben gibt es Schulen, die trotz Unzufriedenheit mit dem Verfahren aktiv werden. Ein gutes Inspektionsergebnis ist bei einem weiteren Typ mit Inaktivität verbunden. Zudem werden Schulen beschrieben, die vor allem durch Aktivitäten im Vorfeld der Inspektion charakterisiert werden können.

2.3.3 Interne Evaluation

Im deutschsprachigem Raum ist der Forschungsstand zur internen Evaluation sehr begrenzt (vgl. Berkemeyer/Müller 2010; Thiel/Thillmann 2012). Dies liegt zum einen an einer fehlenden eindeutigen Definition des Verfahrens, zum anderen an der Vielzahl eingesetzter Verfahren. Aus Deutschland gibt es bislang nur auf Fallstudien oder Pilotprojekten basierende empirische Studien, die von positiven Aspekten wie verstärkter Kommunikation, Offenheit, Selbstwirksamkeit und positiven Auswirkungen auf den Unterricht berichten. Diese sind jedoch nur bedingt verallgemeinerbar (vgl. Thiel/Thillmann 2012). Ergebnisse aus Schulinspektionen lassen darauf schließen, wie viele Schulen von interner Evaluation Gebrauch machen (vgl. Berkemeyer/Müller 2010). So stellt man beispielsweise in den Ländern Berlin und Brandenburg fest, dass im Schuljahr 2007 zwischen 60 Prozent und 80 Prozent der Schulen im Bereich interner Evaluation auf der vierstufigen Bewertungsskala der Schulinspektion als „eher schwach“ oder „schwach“ eingestuft wurden (vgl. MBJS 2008b; SenBWF 2009). Das zeigt eine relativ geringe Verbreitung von interner Evaluation im Sinne der Qualitätsrahmen.

Studien aus Belgien und den Niederlanden von Vanhoof, van Petegem und de Maeyer (2009) sowie Schildkamp und Visscher (2009) unterstreichen die Bedeutung einer positiven Einstellung gegenüber interner Evaluation für die Nutzung der Ergebnisse. So hat die wahrgenommene Nützlichkeit und Anwendbarkeit der Daten einen positiven Einfluss auf die Nutzung (vgl. Meuret/Morlaix 2003). In einer anderen Studie wird berichtet, dass in gut einem Viertel der Schulen Ergebnisse interner Evaluation umfassend kommuniziert und genutzt werden (vgl. Schildkamp/Visscher/Luyten 2009). Rund die Hälfte der Schulleitungen und ein Viertel der Lehrkräfte geben an, Qualitätsentwicklungsmaßnahmen auf Basis der Ergebnisse unternommen zu haben. Diese werden vor allem in den Bereichen Personalentwicklung und Kommunikation über die Arbeitsweise und Qualität der Schule berichtet.

2.3.4 Vergleichende Analyse verschiedener Evaluationsverfahren

Forschung mit einer vergleichenden Analyse verschiedener Evaluationsverfahren gibt es, abgesehen von einer Ausnahme, bislang nicht. Demski u.a. (2012) untersuchen die Einschätzung der Nützlichkeit, Nutzung und Auseinandersetzung mit verschiedenen Evidenzquellen durch Schulleitungen und Lehrkräfte. Evidenzquellen, die der internen Evaluation zuzuordnen sind (z.B. Schülerfeedback, kollegiale Unterrichtsentwicklung), weisen die höchste wahrgenommene Nützlichkeit bzw. den höchsten Grad der Auseinandersetzung und Nutzung auf. Vergleichsarbeiten und Schulinspektion werden dagegen als deutlich weniger nützlich eingeschätzt. Ergebnisse der Schulinspektion werden durch Schulleitungen jedoch in gleichem Ausmaß für Weiterentwicklungen genutzt wie die der internen Evaluation. Schulleitungen schätzen ihre Auseinandersetzung mit und Nutzung von VERA am geringsten ein. Die Auseinandersetzung mit der Schulinspektion ist von allen Evidenzquellen am höchsten. Lehrkräfte nehmen die Auseinandersetzung, Nützlichkeit und Nutzung aller Verfahren als geringer ausgeprägt wahr als Schulleitungen.

3. Fragestellungen

Für die Evaluationsverfahren Schulinspektion und Vergleichsarbeiten liegen bereits einige empirische Befunde vor, Verfahren der internen Evaluation waren in Deutschland bislang hingegen kaum Gegenstand empirischer Forschung. Ebenso wurden bislang fast ausschließlich verfahrensspezifische Aspekte untersucht, obwohl die Verfahren bei unterschiedlichen Durchführungsmodalitäten und Evaluationsgegenständen ähnliche intendierte Funktionen aufweisen. Der Fokus dieser Studie liegt deshalb auf der vergleichenden Betrachtung aller Evaluationsverfahren hinsichtlich ihrer Entwicklungsfunktion. In Anlehnung an das Modell von Helmke (2004) werden die schulinterne Kommunikation und Nutzung von Maßnahmen sowie, als mögliche Einflussfaktoren darauf, der eingeschätzte Aufwand und Nutzen und die attestierte Diagnosegüte analysiert. Anknüpfend an den Forschungsstand, der bei allen Verfahren auch differenzielle Nutzungsmuster zeigt, werden entsprechende verfahrensübergreifende Analysen durchgeführt.

Die Forschungsfragen lauten:

1) Wie werden interne Evaluation, Vergleichsarbeiten und Schulinspektion in Bezug auf folgende Aspekte beschrieben?
   - Kommunikation der Ergebnisse
   - wahrgenommener Aufwand und Nutzen
   - wahrgenommene Diagnosegüte
   - Nutzung von Evaluationsergebnissen
2) Gibt es hinsichtlich dieser Aspekte Unterschiede in der Perzeption zwischen den verschiedenen Evaluationsverfahren?
3) Gibt es auf Schulebene unterschiedliche Konfigurationen der Nutzung der Evaluationsverfahren?

4. Methode

4.1 Stichprobe

Für einen adäquaten Vergleich der Verfahren wurden an das Vorgehen bisheriger Studien anschließend schulische Akteure um Auskunft gebeten. Es wurden Schulleiterinnen und Schulleiter befragt, da diese Experten für die Prozesse an ihren Schulen sind. Die Datengrundlage der vorliegenden Studie stammt aus einer zwischen November 2011 und Februar 2012 flächendeckend durchgeführten Onlinebefragung in den Ländern Berlin und Brandenburg zur schulischen Evaluationspolitik und -praxis. Insgesamt nahmen 329 Schulleiterinnen und Schulleiter an der Befragung teil (27 Prozent Rücklauf). Davon waren 65 Prozent weiblich. Die durchschnittliche Dauer der Schulleitungstätigkeit beträgt 10,4 Jahre (SD=8,0). Mehr als die Hälfte der Befragten ist zwischen 51 und 60 Jahren alt (57 Prozent; 31 bis 40 Jahre: 2 Prozent; 41 bis 50 Jahre: 29 Prozent; über 60 Jahre: 12 Prozent). Die Verteilungen der Schularten und des geografischen Standorts der Schulen entsprechen weitgehend denen in der Grundgesamtheit aller Berliner und Brandenburger Schulen. Tabelle 2 beschreibt die Stichprobe für beide Bundesländer aufgeschlüsselt nach Schulart. Zur besseren Vergleichbarkeit der Ergebnisse zwischen den Bundesländern wurden die Schularten in Grundschulen, Gymnasien und andere weiterführende Schulen unterteilt.

Tab. 2: Stichprobe nach Bundesland und Schulart

	Brandenburg		Berlin		Insgesamt	
	N	in %	*N*	in %	*N*	in %
Gesamt	215	65	114	35	329	100
Schulart:						
Grundschule	140	65	63	55	203	62
Gymnasium	26	12	25	22	51	16
Andere weiterführende Schulen	49	23	26	23	75	23

Quelle: eigene Darstellung

4.2 Instrumente

Die Einschätzung der Evaluationsverfahren erfolgt durch Skalen zum perzipierten Aufwand, Nutzen und der attestierten Diagnosegüte der Evaluationsverfahren. Tabelle 3 enthält eine Übersicht der eingesetzten Skalen mit jeweils einem Beispielitem. Alle Skalen weisen eine gute bis sehr gute interne Konsistenz auf.

Tab. 3: Übersicht der eingesetzten metrischen Skalen mit Angabe der Reliabilität

Skalen	**Items**	**Cronbachs α**	**Evaluationsverfahren**
Wahrgenommener Aufwand	je 3	.85	VERA
Der zeitliche Aufwand war zu hoch.		.86	Schulinspektion
		.85	Interne Evaluation
Wahrgenommener Nutzen	je 3	.86	VERA
Für die pädagogische Arbeit waren die Ergebnisse sehr nützlich.		.82	Schulinspektion
		.81	Interne Evaluation
Eingeschätzte Diagnosegüte	je 4	.85	VERA
Die Stärken unserer Schule wurden in den Ergebnissen hinreichend abgebildet.		.83	Schulinspektion
		.93	Interne Evaluation

Anmerkungen. Antwortmöglichkeiten: „Stimme überhaupt nicht zu" = 1 bis „Stimme vollständig zu" = 6.

Quelle: eigene Darstellung

In Tabelle 4 werden weitere (kategoriale) Items vorgestellt. Die Nutzung von Evaluationsdaten in Form von konkreten Entwicklungsmaßnahmen wurde in ausgewählten Bereichen für VERA und die Schulinspektion erfasst. Zur detaillierten Erfassung der Nutzung waren offene Antworten zu konkret umgesetzten Maßnahmen möglich. Items zur Kommunikation der Ergebnisse in Konferenzen bilden die Rezeption und Reflexion der Ergebnisse ab. Die Veröffentlichung der Ergebnisse im Internet wurde als Indikator für die Kommunikation der Ergebnisse nach außen erfragt.

Tab. 4: Übersicht über die erhobenen kategorialen Indikatoren

Items	Antwortkategorien	Verfahren
Kommunikation der Evaluationsergebnisse		
Wie intensiv fand in den folgenden Gremien ein Austausch über die Ergebnisse der letzten Schulvisitation/-inspektion[1] statt? – in der Gesamt-/Lehrerkonferenz – in der Schulkonferenz	Kein Austausch Einmaliger Austausch Mehrfacher Austausch	VERA Schulinspektion Interne Evaluation
Veröffentlichung der Ergebnisse		
Hat Ihre Schule die folgenden Evaluationsergebnisse im Internet veröffentlicht? (dichotom)	Ergebnisse der Schulvisitation/-inspektion Schulergebnisse aus VERA Ergebnisse der schulinternen Evaluation	VERA Schulinspektion Interne Evaluation
Nutzung von Evaluationsergebnissen		
Wenn Sie an die letzte Schulinspektion denken: Konnten Sie – unabhängig von dem festgestellten Entwicklungsbedarf – die Inspektionsergebnisse für die Entwicklung von Maßnahmen nutzen? – (z.B. Lehr- und Lernprozesse (Unterricht))	Dichotom (ja/nein) und offene Antwort (Mehrfachantworten möglich)	Schulinspektion
Wenn Sie an die letzten Vergleichsarbeiten denken: Haben Sie auf der Grundlage der VERA-Rückmeldungen an Ihrer Schule Entwicklungsprozesse in den folgenden Bereichen angestoßen? – (z.B. Kollegiale Unterrichtsentwicklung)	Dichotom (ja/nein) und offene Antwort (Mehrfachantworten möglich)	VERA

1 In Brandenburg wird die Schulinspektion als Schulvisitation bezeichnet.

Quelle: eigene Darstellung

4.3 Auswertungsmethoden

Neben deskriptiven Analysen zur Beantwortung der ersten Forschungsfrage werden Varianzanalysen (ANOVA) für Mittelwertsvergleiche eingesetzt. Neben den Unterschieden zwischen den Evaluationsverfahren sollen gleichzeitig auch mögliche Unterschiede zwischen den beiden Bundesländern und den Schularten betrachtet werden. Deshalb werden 3-faktorielle ANOVAs durchgeführt. Um die Größe der berichteten Effekte interpretieren zu können, werden die Effektgrößen η^2 bei ANOVAs und Cohens *d* beim Vergleich zweier Mittelwerte berichtet.[1] Für die Berechnungen wurde die Software R 14.2 (vgl. R Development Core Team 2011) verwendet.

Um Konfigurationen der Nutzung von Ergebnissen zu analysieren (dritte Fragestellung) wurden Latent-Class-Analysen (LCA) für dichotome Items ein-

1 η^2>.14: großer Effekt, η^2>.06: mittlerer Effekt, η^2>.01: kleiner Effekt; *d*=.8: großer Effekt, *d*=.5: mittlerer Effekt, *d*=.2: kleiner Effekt (vgl. Cohen 1988).

gesetzt (vgl. Rost 2004). Die LCA ist ein Klassifikationsverfahren, das auf Basis der Antwortmuster einer Person diese einer (latenten) Klasse zuordnet, für die die Zuordnungswahrscheinlichkeit am größten ist. Jede Person wird nur einer einzigen Klasse zugeordnet. Die Bestimmung der Anzahl der Klassen erfolgte einerseits über die Informationskriterien Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICC), Bayesian Information Criterion (BIC) und Sample Adjusted Bayesian Information Criterion (aBIC). Zusätzlich wurden die Modellgeltungstests Vuong-Lo-Mendell-Rubin-Likelihood-Test (LMR; vgl. Lo/Mendell/Rubin 2001) und Bootstrap-Likelihood-Ratio-Differenzentest (BLRT) durchgeführt. Für die LCA wurde die Software Mplus 6.1 (vgl. Muthén/Muthén 1998-2010) verwendet.

5. Ergebnisse

5.1 Einschätzung und Nutzung der Evaluationsverfahren – Gibt es Unterschiede zwischen den Evaluationsformen?

5.1.1 Einschätzung des Aufwands, Nutzens und der Diagnosegüte der Evaluationsverfahren

Zur Beantwortung der ersten und zweiten Forschungsfrage, wie die Evaluationsinstrumente hinsichtlich des wahrgenommenen Aufwands, Nutzens und der attestierten Diagnosegüte eingeschätzt werden und ob sie sich unterscheiden, wurden 3-faktorielle Varianzanalysen durchgeführt (vgl. Tabelle 5). Die Haupteffekte zeigen mittlere signifikante Unterschiede zwischen den Evaluationsverfahren in allen drei abhängigen Variablen (η^2=.04 bis .05). Zudem unterscheiden sich die Befragten bei den zwei Variablen wahrgenommener Nutzen (η^2=.05) und attestierte Diagnosegüte (η^2=.02) nach Bundesland, aber nicht nach Schulart. Signifikante, aber kleine Interaktionseffekte gibt es zwischen Verfahren und Bundesland sowie zwischen Bundesland und Schulart.

Tab. 5: Wahrgenommener Aufwand und Nutzen und attestierte Diagnosegüte nach Evaluationsverfahren, Bundesland und Schulart (ANOVA) im Urteil von Schulleitungen (N=235)

		Aufwand		Nutzen		Diagnosegüte	
	df	*F*	η²	*F*	η²	*F*	η²
Evaluationsverfahren (A)	2	14.85***	.04	19.90***	.05	17.25***	.05
Bundesland (B)	1	0.01	.00	35.36***	.05	12.13***	.02
Schulart (C)	2	0.31	.00	1.55	.00	2.46	.01
A x B	2	2.87	.01	7.98***	.02	5.24**	.02
A x C	4	2.59*	.01	0.82	.00	0.28	.00
B x C	2	0.75	.00	8.49***	.02	4.48*	.01
A x B x C	4	1.06	.01	0.43	.00	0.04	.00

Anmerkungen. *** p<.001, ** p<.01, * p<.05; Evaluationsverfahren: VERA, Schulinspektion und interne Evaluation; Bundesland: Berlin und Brandenburg; Schulart: Grundschule, Gymnasium und andere weiterführende Schulen.

Quelle: eigene Darstellung

Ausgehend von den Ergebnissen der ANOVA werden in Tabelle 6 die Mittelwerte und Standardabweichungen für die Variablen wahrgenommener Aufwand, Nutzen und attestierte Diagnosegüte nach Evaluationsverfahren und Bundesländern differenziert dargestellt. Die Ausprägungen aller Evaluationsverfahren in den Variablen Nutzen und Diagnosegüte liegen oberhalb bzw. dicht am „theoretischen" Mittelwert der Skala von 3.5. Beim Aufwand ist es umgekehrt. Die Evaluationsverfahren werden folglich insgesamt im Durchschnitt eher positiv wahrgenommen, da der Aufwand tendenziell als eher gering, der Nutzen und die Diagnosegüte in der Tendenz als eher hoch bewertet werden. Unterschiede zwischen den Evaluationsverfahren zeigen sich bei allen drei Variablen (vgl. Tabelle 5). Der Aufwand, aber auch der Nutzen werden für die interne Evaluation im Vergleich zu VERA und der Schulinspektion als höher wahrgenommen (d=.39 bis .49). VERA und die Schulinspektion liegen hier auf gleichem Niveau. Die Diagnosegüte wird verfahrensübergreifend durchweg eher positiv beurteilt (M>3.5), wobei Schulinspektion und interne Evaluation noch positiver eingeschätzt werden als Vergleichsarbeiten (d=.41 und .47).

Tab. 6: Mittelwerte und Standardabweichungen für den wahrgenommenen Aufwand und Nutzen und die attestierte Diagnosegüte nach Evaluationsverfahren im Urteil von Schulleitungen

Skala		*N*	VERA *M*(*SD*)	SI *M*(*SD*)	IE *M*(*SD*)
Aufwand		244	3.08_{a} (1.47)	2.99_{a} (1.46)	3.64_{b} (1.39)
Nutzen		244	3.45_{a} (1.39)	3.55_{a} (1.33)	4.09_{b} (1.21)
Diagnosegüte		235	3.76_{a} (1.34)	4.29_{b} (1.21)	4.37_{b} (1.26)
Länderunterschiede					
Aufwand	BE	91	2.92_{a} (1.49)	$3.21_{a/b}$ (1.59)	3.52_{b} (1.44)
	BB	153	3.16_{a} (1.45)	2.87_{a} (1.37)	3.71_{b} (1.35)
Nutzen	BE	91	2.72_{a} (1.25)	3.42_{b} (1.39)	3.82_{b} (1.30)
	BB	153	3.85_{a} (1.30)	3.63_{a} (1.29)	4.26_{b} (1.12)
Diagnosegüte	BE	90	3.28_{a} (1.35)	4.27_{b} (1.30)	4.24_{b} (1.26)
	BB	145	4.03_{a} (1.26)	4.30_{b} (1.16)	4.45_{b} (1.25)

Anmerkungen. *N* enthält nur Fälle, die gültige Antworten bei allen Evaluationsverfahren aufweisen. a/b: Ergebnisse des Post-Hoc Tests (TukeyHSD); gleicher Buchstabe in einer Zeile bedeutet, dass es keine signifikanten Unterschiede zwischen beiden Mittelwerten gibt; ungleiche Buchstaben zeigen signifikante Unterschiede an; SI = Schulinspektion, IE = interne Evaluation, BE = Berlin, BB = Brandenburg.

Quelle: eigene Darstellung

Betrachtet man die Unterschiede zwischen den Bundesländern, lässt sich festhalten, dass die Antworten aus Brandenburg insgesamt tendenziell positiver ausfallen. Beim Nutzen und der Diagnosegüte sind einige der oben beschriebenen Wahrnehmungsunterschiede zwischen den Evaluationsverfahren auf besonders große Effekte in einem der beiden Bundesländer zurückführbar. Auffällig ist die besonders große Differenz zwischen VERA und interner Evaluation in Berlin (d=.86), die durch eine niedrige Bewertung des Nutzens von VERA zustande kommt. Dies zeigt sich auch beim Vergleich mit dem Nutzen der Schulinspektion (d=.53). Die im länderübergreifenden Vergleich gezeigten Unterschiede im Nutzen zwischen Schulinspektion und interner Evaluation gibt es vor allem in Brandenburg (d=.52). Die Diagnosegüte von VERA wird im Vergleich zu den anderen Verfahren als geringer eingeschätzt. Dies ist auf die Berliner Befragten zurückzuführen (d=.74 und .73). Der Nutzen und die Diagnosegüte von VERA werden vor allem in Berlin im Durchschnitt eher negativ eingeschätzt (M<3.5).

5.1.2 Kommunikation über Evaluationsergebnisse

Die Kommunikation über Evaluationsergebnisse kann als wichtig für die schulinterne Rezeption und Reflexion der Ergebnisse und damit als Voraussetzung für die Durchführung von Entwicklungsmaßnahmen auf Schulebene angesehen werden. In der institutionalisierten Kommunikation in Schul- und Lehrerkonferenzen

werden die Ergebnisse aller Evaluationsverfahren in der Regel mindestens einmal thematisiert (vgl. Tabelle 7). Dennoch gibt es Unterschiede zwischen den Verfahren. Die Ergebnisse der Schulinspektion werden in den Konferenzen in der Mehrheit der Schulen mehrfach thematisiert, während die Ergebnisse interner Evaluationen eher einmalig besprochen werden. Vergleichsarbeiten werden im Vergleich dazu öfter mehrfach thematisiert, allerdings immer noch weniger oft als die Ergebnisse aus Schulinspektionen. In der Schulkonferenz werden die Ergebnisse der Evaluationsverfahren insgesamt weniger häufig besprochen als in der Lehrerkonferenz, die jedoch in der Regel öfter stattfindet.

Tab. 7: Kommunikation über Evaluationsergebnisse in Konferenzen nach Angaben von Schulleitungen (in %)

Konferenz	Häufigkeit des Austausches	VERA	SI	IE	X²-Test
Lehrerkonferenz[a]/ Gesamtkonferenz[b]	Kein Austausch	7	1	10	X²=132.91, *df*=4, *p*<.001
	Einmaliger Austausch	45	21	63	
	Mehrfacher Austausch	48	79	27	
Schulkonferenz	Kein Austausch	16	0	24	X²=132.66, *df*=4, *p*<.001
	Einmaliger Austausch	60	50	68	
	Mehrfacher Austausch	24	50	9	

Anmerkungen. VERA = Vergleichsarbeiten, SI = Schulinspektion, IE = Interne Evaluation; [a]: Bezeichnung in Brandenburg, [b]: Bezeichnung in Berlin.

Quelle: eigene Darstellung

Evaluationsergebnisse können nicht nur schulintern kommuniziert werden, sondern auch an schulexterne Akteure. Die *Veröffentlichung der Ergebnisse im Internet* ist ein Indikator dafür. Am häufigsten wurden nach Angaben der Schulleitungen Ergebnisse der Schulinspektion im Internet veröffentlicht (43 Prozent der Schulen), gefolgt von Ergebnissen der Vergleichsarbeiten (20 Prozent) und denen interner Evaluationen (16 Prozent).

5.1.3 Nutzung der Evaluationsergebnisse

Für die Verfahren Schulinspektion und VERA liegen Informationen vor, in welchen Bereichen Evaluationsergebnisse zur Ableitung von Entwicklungsmaßnamen von den Schulen genutzt werden (vgl. Tabelle 8).

Tab. 8: Nutzung von Evaluationsergebnissen: abgeleitete Maßnahmen nach Angaben von Schulleitungen

Bereich	*N* (in %)	Beispielnennungen offene Antworten
	Schulinspektion	
Lehr- und Lernprozesse (Unterricht)	207 (66)	Binnendifferenzierung, schulinternes Curriculum wurde abgestimmt und erweitert, Fortbildungen
Schulkultur	146 (47)	Neue Strukturen bei der Schüler- und Elternbeteiligung werden geschaffen, Kooperationen, Etablierung schulischer Höhepunkte
Schulmanagement	163 (52)	Hospitation von der Schulleitung, bessere Aufgabenverteilung, Personalgespräche
Personalentwicklung	148 (47)	Fortbildungen, kollegiale Hospitationen, Personalgespräche
Ziele und Strategien der Qualitätsentwicklung	167 (54)	Überarbeitung Schulprogramm, Arbeit am schulinternen Curriculum, Einführung Selbstevaluation
	Vergleichsarbeiten	
Lehr- und Lernprozesse (Unterricht)	222 (75)	Binnendifferenzierung, Festlegung von Lernschwerpunkten, Leseförderung
Arbeit am Schulprogramm	102 (34)	Überarbeitung der Formulierungen schulinterner Ziele und Maßnahmen
Ausbau schulinterner Strukturen	117 (40)	Intensivierung der Arbeit der Fachkonferenzen, Arbeitsgruppe zur Erarbeitung des Übungsmaterials
Personalentwicklung	119 (40)	Fortbildungen; Hospitationen
Kollegiale Unterrichtsentwicklung	181 (61)	gegenseitige Unterrichtsbesuche, fächerverbindendes Arbeiten

Anmerkungen. Schulinspektion: *N*=312, VERA: *N*=296.

Quelle: eigene Darstellung

Schulleitungen geben an, dass sowohl die Ergebnisse der Schulinspektion (66 Prozent) als auch die der Vergleichsarbeiten (75 Prozent) vor allem im Bereich Lehr- und Lernprozesse (Unterricht) für Entwicklungsmaßnahmen genutzt werden. Für die anderen erfragten Bereiche, die auf dem Orientierungsrahmen Schulqualität basieren, gaben jeweils ca. die Hälfte der Schulleitungen an, die Schulinspektionsergebnisse für Maßnahmen genutzt zu haben. Vergleichsarbeiten werden nach Schulleitungsangaben erwartungsgemäß öfter für unterrichtsbezogene Entwicklungsmaßnahmen verwendet als für Maßnahmen der Schulentwicklung, wie beispielweise die Arbeit am Schulprogramm. Maßnahmen der Schulentwicklung werden dennoch in mehr als einem Drittel der Schulen auf Basis von VERA-Rückmeldungen durchgeführt.

Die offenen Antworten verdeutlichen, dass es eine große Bandbreite an abgeleiteten Maßnahmen gibt. Aus beiden Verfahren werden am häufigsten Maßnahmen zur Binnendifferenzierung abgeleitet. Sowohl Schulinspektions- als auch VERA-

Ergebnisse werden zur Einführung von Maßnahmen der internen Evaluation genutzt (z.B. Hospitationen in verschiedenen Konstellationen).

Im Zuge von Schulinspektionen erhalten Schulen Hinweise, in welchen Bereichen es in der Schule Entwicklungsbedarf gibt. Tabelle 9 zeigt, in welchen Bereichen Entwicklungsbedarf bestand und ob Schulen dort Aktivitäten eingeleitet haben.

Tab. 9: Anteil der Schulen mit diagnostiziertem Entwicklungsbedarf und Schulen mit eingeleiteten Entwicklungsmaßnahmen nach Angaben von Schulleitungen (in %; *N*=312)

	Entwicklungsbedarf			Kein Entwicklungsbedarf		
	Gesamt	Maßnahmen abgeleitet		Gesamt	Maßnahmen abgeleitet	
		ja	nein		ja	nein
Lehr- und Lernprozesse (Unterricht)	67	67	33	33	64	36
Schulkultur	13	53	48	87	46	54
Schulmanagement	28	60	40	72	49	51
Personalentwicklung	38	48	52	62	47	53
Ziele und Strategien der Qualitätsentwicklung	51	56	44	49	51	49

Quelle: eigene Darstellung

Von den 67 Prozent der Schulen, für die Entwicklungsbedarf im Bereich Lehr- und Lernprozesse durch die Schulinspektion festgestellt wurde, geben zwei Drittel an, entsprechende Maßnahmen eingeleitet zu haben. Ein Drittel der Schulen mit Entwicklungsbedarf wurde demzufolge nicht aktiv. Auf der anderen Seite haben von dem Drittel Schulen ohne Entwicklungsbedarf 64 Prozent dessen ungeachtet Maßnahmen abgeleitet. Qualtätsentwicklungsmaßnahmen bei festgestelltem Entwicklungsbedarf zeigen sich außer im Bereich Lehr- und Lernprozesse vor allem beim Schulmangement (60 Prozent), aber auch bei der Hälfte der Schulen in den anderen Bereichen. Das bedeutet im Umkehrschluss, dass in den Bereichen Schulmanagement, Personalentwicklung sowie Ziele und Strategien der Qualitätsentwicklung auch in etwa 40 bis 50 Prozent der Schulen trotz des Entwicklungsbedarfs keine Nutzung der Ergebisse für Entwicklungsmaßnahmen stattfindet. Qualitätssicherungsmaßnahmen ohne festgestellten Entwicklungsbedarf sind ebenfalls im Bereich Lehr- und Lernprozesse am häufigsten anzutreffen (64 Prozent). In den anderen Bereichen findet Qualitätssicherung in etwa der Hälfte der Schulen statt.

5.2 Konfigurationen der Nutzung von Schulinspektionen und Vergleichsarbeiten

In Schulen werden nach Angaben der befragten Schulleitungen auf Basis der Evaluationsergebnisse verschiedene Maßnahmen abgeleitet. Es wurden Latent-Class-Analysen durchgeführt, um unterschiedliche *Konfigurationen der Maßnahmen* basierend auf dem Feedback aus Schulinspektionen und Vergleichsarbeiten zu beschreiben. So kann nicht nur festgestellt werden, welche Maßnahmen durchgeführt werden, sondern auch, in welcher Kombination die Maßnahmen ergriffen werden.

Für die Auswahl der Anzahl der Antwortmuster (Klassen) wurden Modellgütetests und Informationskriterien herangezogen (vgl. Tabelle 10). Die Informationskriterien AIC und aBIC haben die niedrigsten Werte beim 6-Klassen-Modell und weisen damit für dieses Modell die beste Passung auf. Das Informationskriterium BIC und der Modelltest LMR weisen beim 3-Klassen-Modell die beste Passung auf, während der AICC und BLRT ein 5-Klassen-Modell indizieren. Die Unterschiede zwischen dem 6- und 5-Klassen-Modell sind beim AIC und aBIC nicht groß. Der BLRT ist zudem ein sehr zuverlässiger Indikator (vgl. Nylund/Asparouhov/Muthén 2007), weshalb das Modell mit fünf Klassen präferiert wird. Die mittlere Klassen-Zuordnungswahrscheinlichkeit ist mit .80 bis .97 zufriedenstellend hoch.

Tab. 10: Ergebnisse der Modellgütetests und Informationskriterien der Latent-Class-Analysen (N=295)

Modell	LL	AIC	BIC	aBIC	AICC	BLRT	LMR
2 Klassen	-1700	3442	3519	3452	3445	.00	.00
3 Klassen	-1587	3239	3357	3255	3247	.00	.00
4 Klassen	-1553	3192	3351	3214	3207	.00	.00
5 Klassen	-1535	3179	3378	3207	3203	.00	.31
6 Klassen	-1521	3172	3412	3206	3209	.10	.56

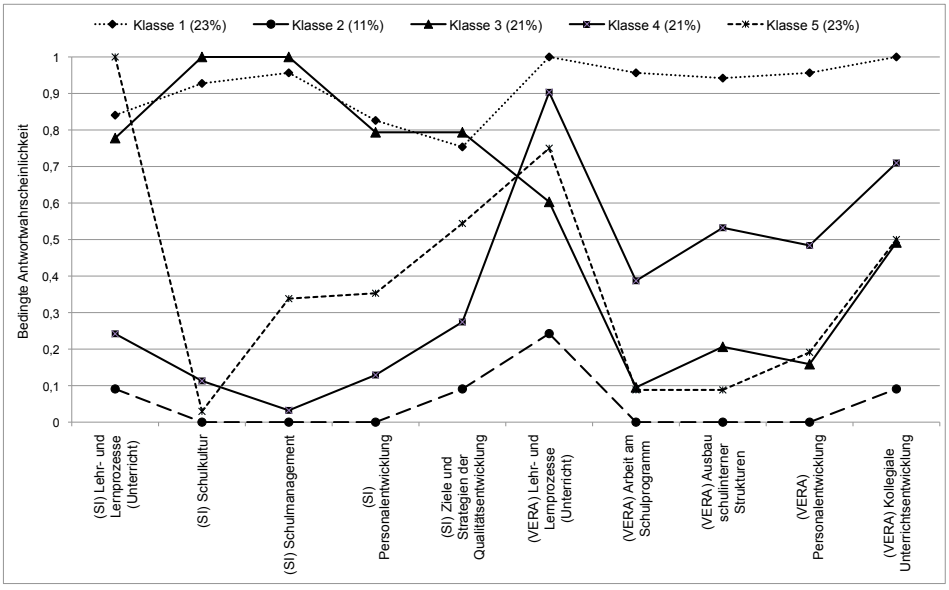
Anmerkungen. LL: Loglikelihood, BLRT: *p*-Wert des Bootstrap Likelihood Ratio Differenzentests, LMR: *p*-Wert des Vuong-Lo-Mendell-Rubin-Likelihood-Tests.

Quelle: eigene Darstellung

Die Itemprofile (vgl. Abb. 2) zeigen für jede identifizierte Gruppe (Klasse) für jedes Item die Wahrscheinlichkeit an, dieses mit ja zu beantworten. Es lassen sich zunächst zwei entgegengesetzte Muster der Nutzung von Ergebnissen im Itemprofil unterscheiden. Ein Itemprofil (Klasse 1, 23 Prozent) zeichnet sich durch eine in allen Bereichen umfangreiche Nutzung aus, da die Werte für alle Variablen nahe 1 liegen. Das heißt, Ergebnisse der Schulinspektion und der Vergleichsarbeiten werden in allen Bereichen für Entwicklungsmaßnahmen verwendet. Konträr dazu ist ein Itemprofil durch weitgehende Nicht-Nutzung in allen Bereichen gekennzeichnet (Klasse 2, 11 Prozent). Zwei weitere Itemprofile lassen sich ebenfalls als gegensätzlich beschreiben,

allerdings liegen die Unterschiede bei den Verfahren. In Klasse 3 (21 Prozent) werden primär die Schulinspektionsergebnisse für Entwicklungsmaßnahmen genutzt. In geringerem Umfang werden in diesem Nutzungsprofil auch VERA-Ergebnisse genutzt, allerdings nur für direkt unterrichtsbezogene Aktivitäten. Dem gegenüber gibt es eine Gruppe (Klasse 4, 21 Prozent), deren Schwerpunkt auf der Nutzung von VERA-Daten liegt, auch hier mit dem Fokus auf Unterricht. Das letzte Nutzungsprofil (Klasse 5, 23 Prozent) ist durch den primären Fokus auf Unterrichtsentwicklungsmaßnahmen charakterisierbar, unabhängig davon, ob die Ergebnisse aus der Schulinspektion oder VERA stammen.

Abb 2: Itemprofile der Nutzung von Schulinspektions- und VERA-Ergebnissen



Quelle: eigene Darstellung

Im Anschluss an die Identifikation unterschiedlicher Konfigurationen von Entwicklungsmaßnamen wird exploriert, ob es einen Zusammenhang zwischen diesen und der Schulformzugehörigkeit bzw. dem Bundesland gibt. Tabelle 11 zeigt, dass die Verteilung der latenten Klassen zwischen den Ländern und Schulformen variiert (χ^2=48, *df*=22, *p*<.001, ω=.27). In Berlin gehören Schulen unterproportional Klasse 1 (aktive Schulen) und überproportional Klasse 2 (inaktive Schulen) an. Auffällig ist insbesondere der geringe Anteil an Gymnasien in Klasse 1. Abweichungen gibt es weiterhin auch für die anderen latenten Klassen, insbesondere bei den weiterführenden Schulen, die unterproportional zu Klasse 4 (primär VERA-Nutzung mit Fokus Unterricht) zugeordnet sind. Diese Ergebnisse sind allerdings statistisch nicht signifikant (χ^2=3, *df*=8, *p*=.90, ω=.20). Hinsichtlich der Gymnasien gibt es in Brandenburg (χ^2=28, *df*=8, *p*<.001, ω=.28) ein sehr ähnliches Bild: Gymnasien gehören überpro-

portional Klasse 2 (inaktive Schulen) an und sind nahezu nicht in Klasse 1 (aktive Schulen) zu finden. Bei den Grundschulen ist das Gegenteil der Fall: Diese gehören überproportional Klasse 1 und unterproportional Klasse 2 an. Länderübergreifend lässt sich feststellen, dass sich die befragten Schulleitungen in den Angaben zur Nutzung von Evaluationsergebnissen zwischen den Schulformen unterscheiden. Insbesondere in Gymnasien wird zwar überproportional häufig keine Aktivität berichtet. Dies betrifft jedoch nur eine Minderheit der Gymnasien; in der Mehrzahl wird von der Nutzung der Ergebnisse in spezifischen Konfigurationen berichtet.

Tab. 11: Anteil der Klassen nach Bundesland und Schulform, absolut (in %)

		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Gesamt
Berlin	Grundschule	8 (14%)	9 (16%)	15 (27%)	9 (16%)	17 (29%)	58
	Gymnasium	1 (4%)	4 (17%)	6 (26%)	6 (26%)	6 (26%)	23
	Andere weiterführende Schulen	4 (17%)	4 (17%)	6 (26%)	3 (13%)	6 (26%)	23
Brandenburg	Grundschule	44 (36%)	3 (2%)	23 (19%)	30 (24%)	23 (19%)	123
	Gymnasium	1 (4%)	7 (29%)	6 (25%)	4 (17%)	6 (25%)	24
	Andere weiterführende Schulen	11 (25%)	6 (14%)	7 (16%)	10 (23%)	10 (23%)	44
	Gesamt	69 (23%)	33 (11%)	63 (21%)	69 (23%)	68 (23%)	295

Quelle: eigene Darstellung

6. Diskussion und Ausblick

Evaluationsverfahren, wie Vergleichsarbeiten, Schulinspektion und interne Evaluation, sollen als Impuls für Qualitätsentwicklung und -sicherung in Schulen dienen. Die vorliegende Studie untersucht, inwiefern dies gelingt und ob es Unterschiede zwischen den genannten Verfahren gibt. In Anlehnung an das Modell von Helmke (2004) wurden Angaben von Schulleitungen zur Nutzung der Ergebnisse der verschiedenen Evaluationsverfahren als Indikator für Qualitätsentwicklung analysiert. Zudem wurden die Verfahren hinsichtlich möglicher Einflussfaktoren darauf, wie Ergebniskommunikation sowie wahrgenommener Aufwand, Nutzen und Diagnosegüte, eingeschätzt. Auch hier wurde analysiert, ob es Unterschiede zwischen den Verfahren gibt.

6.1 Unterschiede zwischen Evaluationsverfahren in der Einschätzung und Nutzung

6.1.1 Einschätzung der Evaluationsverfahren

Die Einschätzung der Schulleitungen hinsichtlich des wahrgenommenen Aufwandes und Nutzens und der zugeschriebenen Diagnosegüte unterscheidet sich bedeutsam zwischen den Evaluationsverfahren. Analog zur Studie von Demski u.a. (2012) wird die Nützlichkeit von interner Evaluation am höchsten eingeschätzt. Dennoch zeigen die Ergebnisse, dass alle Evaluationsverfahren tendenziell eher positiv gesehen werden. Eine Ausnahme ist VERA in Berlin, wo der Nutzen und die Diagnosegüte als eher gering bewertet werden (zu den Länderunterschieden vgl. Maier u.a. 2011; Maier 2008). Dass VERA im Vergleich der Verfahren als am wenigsten nützlich eingeschätzt wird, zeigt sich auch bei Demski u.a. (2012).

Die positiven Einschätzungen der Diagnosegüte und des Nutzens interner Evaluation hängen möglicherweise damit zusammen, dass im Gegensatz zu den anderen Verfahren in den Schulen weitgehend selbst über Inhalte und Vorgehensweise bestimmt werden kann. Allerdings ist unbekannt, welche internen Evaluationsverfahren in den Stichprobenschulen eingesetzt werden. Weiterhin könnte auch die Implementation der Verfahren eine Rolle bei der Einschätzung spielen. VERA und die Schulinspektion sind top-down initiiert worden, ohne dass die Schulen beteiligt wurden, was eine mögliche Erklärung für niedrigere Akzeptanz sein kann. Diese These wird jedoch nicht durch das Ergebnis gestützt, dass sich insbesondere bei der Einschätzung der Diagnosegüte und des Nutzens von VERA bundeslandspezifische Unterschiede zeigen, obwohl VERA in beiden Ländern vom selben Institut durchgeführt und ausgewertet wird.

6.1.2 Kommunikation der Evaluationsergebnisse

Die hier präsentierten Befunde zur Kommunikation in Schul- und Lehrerkonferenzen zeigen, dass sich mit VERA-Ergebnissen weniger auseinandergesetzt wird als mit den Ergebnissen anderer Verfahren. Dieses Resultat in Bezug auf die Ergebniskommunikation deckt sich mit den Ergebnissen der Studie von Demski u.a. (2012). Informationen auf Basis der befragten Schulleitungen über die Kommunikation in Fachkonferenzen liegen nicht vor. Möglicherweise ist hier die Kommunikation über VERA-Ergebnisse anders ausgeprägt. Die Kommunikation über Schulinspektionsergebnisse ist im Vergleich am stärksten gegeben; dies wird ebenfalls von Demski u.a. (2012) berichtet. Die beschriebene umfangreiche Kommunikation deckt sich zudem mit bisherigen Befunden zur Auseinandersetzung mit Inspektionsergebnissen (vgl. zusammenfassend Dedering 2012). Die Unterschiede zwischen den Verfahren könnten damit zusammenhängen, dass die Schulinspektion

bis zum Befragungszeitpunkt lediglich einmal stattgefunden hatte und nicht wie VERA jährlich stattfindet. Zudem ist die Vorstellung der Inspektionsergebnisse zumindest für einen Teil der schulischen Akteure im Ablauf des Verfahrens vorgesehen.

6.1.3 Nutzung von Evaluationsverfahren

Die in dieser Studie untersuchten Evaluationsverfahren liefern Erkenntnisse, die in Schulen zur Ableitung von Entwicklungsmaßnahmen genutzt werden können. Die zu VERA und der Schulinspektion befragten Schulleitungen berichten, dass Nutzung auf Schulebene stattfindet und zwar vor allem im „Kerngeschäft" der Schule, dem Unterricht. Die Ableitung von Maßnahmen im Bereich des Unterrichts (z.B. Binnendifferenzierung) wird auch in anderen Studien berichtet (vgl. zur Schulinspektion: Dedering 2012; zu VERA: z.B. Groß Ophoff/Hosenfeld/Koch 2007; Koch u.a. 2006). Die in der Mehrheit der Schulen berichtete Nutzung zeigt, dass Unterrichtsentwicklung basierend auf Evaluationsergebnissen auch auf Schulebene initiiert wird. Dies unterstreicht die als Merkmal der „neuen Steuerung" formulierte Unterrichtsfokussierung der Verfahren. Nach Angaben der Schulleitungen wird das von der KMK intendierte Ziel der Unterrichtsentwicklung in der Mehrzahl der Schulen berichtet. Zudem findet die Ableitung von Schulentwicklungsmaßnahmen nach Auskunft der Schulleitungen je nach Verfahren in einem Drittel bis zur Hälfte der Schulen statt. Weitere beschriebene Maßnahmen sind solche zur internen Evaluation und Schulprogrammarbeit, deren Nennung ebenfalls mit anderen Studien übereinstimmt (zur Schulinspektion vgl. zusammenfassend Dedering 2012).

Trotz der insbesondere in Berlin negativeren Einschätzung von VERA geben Schulleitungen in über zwei Drittel der Schulen an, die Testergebnisse für Entwicklungsmaßnahmen im Bereich Unterricht genutzt zu haben. Den Zusammenhang zwischen der Einschätzung von Evaluationsverfahren und der Nutzung von Evaluationsergebnissen gilt es in weiterführenden Studien genauer zu untersuchen. Zudem ist bislang kaum erforscht, inwiefern festgestellter Entwicklungsbedarf bzw. das Abschneiden bei VERA zu Aktivitäten führen.

Die Ergebnisse der Schulinspektion in Form festgestellten Entwicklungsbedarfs sollen wie das Feedback aus VERA einen Anlass bieten, Maßnahmen einzuleiten. In etwa der Hälfte der Schulen werden auch ohne Entwicklungsbedarf in einem bestimmten Bereich Maßnahmen abgeleitet. Das bedeutet: Schulen nutzen die Inspektionsergebnisse auch in Bereichen, in denen kein akuter Handlungsbedarf besteht. Das kann als Hinweis auf die Orientierung an Zielvorgaben interpretiert werden. Zwischen einem Drittel und der Hälfte der Schulen ergreifen aber auch trotz diagnostizierten Entwicklungsbedarfs keine Maßnahmen. Gründe dafür könnten sein, dass Schulen die Ergebnisse ablehnen, der Zeitraum zwischen Inspektion und Umsetzungsmöglichkeit zu kurz war (vgl. Gärtner/Wurster 2009a, 2009b) oder

Schulen nicht dazu in der Lage sind, eigenständig geeignete Maßnahmen zu ergreifen (vgl. Ehren/Visscher 2006). Hier besteht Bedarf für weiterführende Forschung.

Ein weiterer Befund zur Nutzung der Verfahren ist, dass Schulinspektion und VERA mit der Einführung von Formen interner Evaluation zusammenhängen. Beispiele dafür sind der Einsatz von Hospitationen oder die Arbeit am Schulprogramm als eine mögliche Grundlage interner Evaluation. Andere Befunde bestätigen dies: In 25 Prozent der Schulen berichten Schulleitungen, dass die Inspektion zu einem anderen Umgang mit VERA und zentralen Abschlussprüfungen und in mehr als der Hälfte der Schulen zu interner Evaluation geführt hat (vgl. Gärtner/Hüsemann/Pant 2009). Die allgemeine Funktion der Zielvorgaben bzw. Normendurchsetzung wird durch den berichteten Zusammenhang von Schulinspektions- und VERA-Ergebnissen mit Maßnahmen interner Evaluation deutlich, da interne Evaluation als eine Zielvorgabe in den Qualitätsrahmen genannt wird.

6.1.4 Konfigurationen der Nutzung von Schulinspektion und Vergleichsarbeiten

Das Zusammenspiel und die gemeinsame Betrachtung verschiedener Evaluationsverfahren sind zur Beantwortung der Forschungsfrage nach Konfigurationen der Nutzung unterschiedlicher Verfahren bedeutsam. Zur Analyse wurden Latent-Class-Analysen durchgeführt. Unterschiedliche Nutzungskonfigurationen zeigen, dass es in substanziellem Ausmaß sowohl Schulen gibt, die Schulinspektionen und VERA gleichermaßen nutzen, als auch Schulen, die keines der beiden Evaluationsinstrumente bzw. nur eines nutzen. Dabei wurden auch Unterschiede zwischen Schularten sichtbar. Unterschiedliche Konfigurationen der Nutzung wurden bereits separat bei Lehrkräften für VERA (vgl. Groß Ophoff/Hosenfeld/Koch 2007) und auf Schulebene für die Schulinspektion (vgl. Wurster/Gärtner 2013) beschrieben. In diesen Studien, wie auch in der vorliegenden, wird deutlich, dass ein Teil der befragten Lehrkräfte bzw. Schulen die Ergebnisse nicht nutzt. Dies kann einerseits damit zusammenhängen, dass es keinen Bedarf an Veränderungsprozessen gibt (vgl. ebd.), andererseits aber auch damit, dass die Verfahren eher abgelehnt werden oder Schulen bzw. Lehrkräfte nicht dazu in der Lage sind, eigenständig geeignete Maßnahmen zu ergreifen.

6.2 Limitationen der Studie

Einschränkend muss festgehalten werden, dass die vorgestellten Ergebnisse auf einer Stichprobe basieren, die möglicherweise verzerrt ist. Die Ausschöpfungsquote der Vollerhebung liegt bei 27 Prozent. Ein Abgleich zwischen Stichprobe und Grundgesamtheit konnte nur für die regionale Verteilung der Schulen und die

Schulart durchgeführt werden, wobei keine Verzerrungen festgestellt wurden. Weiterhin sind hier Schulleitungen als Expertinnen und Experten ihrer Schule befragt worden. Mit Blick auf bisherige Forschungsergebnisse fallen Antworten von Schulleitungen tendenziell etwas positiver aus als Auskünfte von Lehrkräften (vgl. Böttcher/Keune/Neiwert 2010; Demski u.a. 2012; Gärtner/Wurster 2009a, 2009b; Maier 2008). Insgesamt muss bei der Interpretation der Ergebnisse dieser Studie beachtet werden, dass die Ergebnisse auf Selbstauskünften schulischer Akteure und auf Querschnittsdaten beruhen.

6.3 Implikationen der Studie

Die analysierten Evaluationsverfahren sind Thema aktueller bildungspolitischer Entscheidungen zur Qualitätssicherung im Bildungssystem. Die KMK hat jüngst die „Vereinbarung zur Weiterentwicklung von VERA" beschlossen (vgl. KMK 2012). In Berlin wurde 2011 ein Qualitätspaket (SenBWF 2011) verabschiedet, dass u.a. die Veröffentlichung der Schulinspektionsergebnisse, eine Verpflichtung der Schulen zur regelmäßigen Nutzung eines standardisierten Instruments zur internen Evaluation (SEP; vgl. Gärtner 2010) und für die VERA-Rückmeldung Vergleichsmöglichkeiten mit ähnlich zusammengesetzten Schulen vorsieht.

Anknüpfend an die aktuellen Entwicklungen lassen sich auf Basis der Ergebnisse dieser Arbeit unter Berücksichtigung der Limitationen erste Hinweise und Hypothesen zur Gestaltung und Weiterentwicklung schulischer Evaluationsverfahren formulieren. Zunächst zeigen die Befunde, dass insbesondere in Berlin die Akzeptanz von VERA eine Diskrepanz zur Bedeutung aufweist, die die Bildungsadministration VERA zumisst. Welche Maßnahmen konkret zu einer Verbesserung der Akzeptanz beitragen können, muss durch systematisch variierte Angebote ermittelt werden.

Die Bereitstellung von standardisierten Verfahren zur internen Evaluation könnte den im Vergleich zu den anderen Verfahren als höher eingeschätzten Aufwand reduzieren, da Schulen keine eigenen Prozeduren entwickeln müssen. Dabei sollte allerdings beachtet werden, dass nicht bekannt ist, welche Formen interner Evaluation konkret zu der im Vergleich als hoch wahrgenommenen Diagnosegüte und Nützlichkeit interner Evaluation geführt haben. Es ist folglich nicht auszuschließen, dass Schulen ein standardisiertes Instrument nicht so positiv wahrnehmen.

Weiterhin kann die gleichzeitige Verwendung von VERA und Schulinspektionsergebnissen in einem Teil der Schulen als Anlass genommen werden, perspektivisch die unterschiedlichen Verfahren stärker aufeinander zu beziehen. Dies könnte beispielsweise über eine gemeinsame Betrachtung aller Evaluationsergebnisse in der Maßnahmenplanung oder in Beratungssettings erfolgen. Dafür spricht auch der

Befund, dass aufgrund der VERA- und Inspektionsergebnisse Maßnahmen zur internen Evaluation ergriffen werden. Die Nicht-Nutzung in einem vergleichsweise kleinen Teil der Schulen könnte auf einen grundlegenden Unterstützungs- bzw. Fortbildungsbedarf im Umgang mit den Verfahren in diesen Schulen hinweisen.

6.4 Ausblick

Neben den oben genannten Aspekten lässt sich weiterführende Forschung aus den theoretischen Modellen zur Evaluationsnutzung ableiten. Beispielsweise kann das Zyklenmodell von Helmke (2004) als Heuristik zur Systematisierung möglicher Einflussfaktoren auf die Nutzung zu Grunde gelegt werden. Eine verfahrensübergreifende Betrachtung sollte um die organisationalen und externen Bedingungen der Nutzung von Evaluationsergebnissen erweitert werden, wie dies bereits in Ansätzen für einzelne Verfahren geschehen ist (vgl. Maier u.a. 2011; Schildkamp u.a. 2012). Außerdem gehen die Modelle von einer wechselseitigen Beeinflussung der Modellbestandteile aus. Diese wäre ebenfalls in zukünftigen Arbeiten in den Blick zu nehmen, z.B. über Analysen zum Einfluss von individuellen, organisationalen und externen Faktoren auf die Nutzung der Daten, die auf diese Weise erklärt werden kann. Die vorliegende Studie hat mit der Herausarbeitung von verfahrensübergreifenden Konfigurationen schulinterner Verarbeitung von Evaluationsergebnissen einen ersten Schritt in diese Richtung geleistet.

Literatur und Internetquellen

Altrichter, H. (2010): Schul- und Unterrichtsentwicklung durch Datenrückmeldung. In: Altrichter, H./Maag Merki, K. (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS, S. 219-254.

Berkemeyer, N./Müller, S. (2010): Schulinterne Evaluation – nur ein Instrument zur Selbststeuerung von Schulen? In: Altrichter, H./Maag Merki, K. (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS, S. 195-218.

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: Westfälische Wilhelms-Universität.

Cohen, J. (1988): Statistical Power Analysis for the Behavioral Science. Hillsdale, NJ: Erlbaum.

Dedering, K. (2012): Schulinspektion als wirksamer Weg der Systemsteuerung. In: Zeitschrift für Pädagogik 58, H. 1, S. 69-88.

Dedering, K./Müller, S. (2011): School Improvement through Inspections? First Empirical Insights from Germany. In: Journal of Educational Change 12, H. 3, S. 301-322.

Demski, D./Rosenbusch, C./van Ackeren, I./Clausen, M./Schmidt, U. (2012): Steuerung von Schule durch evidenzbasierte Einsicht? Konzeption und erste Befunde des Forschungsverbundes EviS. In: Hornberg, S./Parreira do Amaral, M. (Hrsg.): Deregulierung im Bildungswesen. Münster u.a.: Waxmann, S. 131-150.

Döbert, H./Dedering, K. (Hrsg.) (2008): Externe Evaluation von Schulen. Historische, rechtliche und vergleichende Aspekte. Münster u.a.: Waxmann.

Ehren, M.C.M./Visscher, A.J. (2006): Towards a Theory on the Impact of School Inspections. In: British Journal of Educational Studies 54, H. 1, S. 51-72.

Gärtner, H. (2010): Das ISQ-Selbstevaluationsportal. Konzepte eines Online-Angebots, um die Selbstevaluation in Schule und Unterricht zu unterstützen. In: Die Deutsche Schule 102, H. 2, S. 163-175.

Gärtner, H./Hüsemann, D./Pant, H.A. (2009): Wirkungen von Schulinspektion aus Sicht betroffener Schulleitungen. Die Brandenburger Schulleiterbefragung. In: Empirische Pädagogik 23, H. 1, S. 1-18.

Gärtner, H./Wurster, S. (2009a): Befragung zur Wirkung von Schulinspektion in Berlin. Ergebnisbericht. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg.

Gärtner, H./Wurster, S. (2009b): Befragung zur Wirkung von Schulvisitation in Brandenburg. Ergebnisbericht. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg.

Groß Ophoff, J./Hosenfeld, I./Koch, U. (2007): Formen der Ergebnisrezeption und damit verbundende Schul- und Unterrichtsentwicklung. In: Empirische Pädagogik 21, H. 4, S. 411-427.

Helmke, A. (2004): Von der Evaluation zur Innovation: Pädagogische Nutzbarmachung von Vergleichsarbeiten in der Grundschule. In: Das Seminar, H. 2, S. 90-112.

Hosenfeld, A. (2010): Führt Unterrichtsrückmeldung zu Unterrichtsentwicklung? Die Wirkung von videographischer und schriftlicher Rückmeldung bei Lehrkräften der vierten Jahrgangsstufe. Münster u.a.: Waxmann.

Huber, F. (2006): Konsequenzen aus der externen Evaluation an Bayerns Schulen. Auswertung einer Befragung von Schulleiterinnen und Schulleitern zu den Entwicklungen an ihren Schulen. München: Staatsinstitut für Schulqualität und Bildungsforschung (ISB).

Husfeldt, V. (2011): Wirkungen und Wirksamkeit der externen Schulevaluation. Überblick zum Stand der Forschung. In: Zeitschrift für Erziehungswissenschaft 14, H. 2, S. 259-282.

Klieme, E. (2004): Begründung, Implementation und Wirkung von Bildungsstandards. In: Zeitschrift für Pädagogik 50, S. 625-634.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (1997): Grundsätzliche Überlegungen zu Leistungsvergleichen innerhalb der Bundesrepublik Deutschland – Konstanzer Beschluss. (Beschluss der Kultusministerkonferenz vom 24.10.1997.) URL: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1997/1997_10_24-Konstanzer-Beschluss.pdf; Zugriffsdatum: 18.07.2013.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2006): Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. München: Wolters Kluwer.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2010): Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung. Köln: Wolters Kluwer.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2012): Vereinbarung zur Weiterentwicklung von VERA. (Beschluss der Kultusministerkonferenz vom 08.03.2012.) URL: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_03_08_Weiterentwicklung-VERA.pdf; Zugriffsdatum: 17.05.2013.

Koch, U./Groß Ophoff, J./Hosenfeld, I./Helmke, A. (2006): Qualitätssicherung: Von der Evaluation zur Schul- und Unterrichtsentwicklung – Ergebnisse der Lehrerbefragung

zur Auseinandersetzung mit den VERA-Rückmeldungen. In: Eder, F./Gastager, A./Hofmann, F. (Hrsg.): Qualität durch Standards? Beiträge zum Schwerpunktthema der 67. Tagung AEPF. Münster u.a.: Waxmann, S. 187-199.

Kühle, B./Peek, R. (2007): Lernstandserhebungen in Nordrhein-Westfalen. Evaluationsbefunde zur Rezeption und zum Umgang mit Ergebnisrückmeldungen in Schulen. In: Empirische Pädagogik 21, H. 4, S. 428-447.

Landwehr, N. (2011): Thesen zur Wirkung und Wirksamkeit der externen Schulevaluation. In: Quesel, C./Husfeldt, V./Landwehr, N./Steiner, P. (Hrsg.): Wirkungen und Wirksamkeit der externen Schulevaluation. Bern: hep, S. 35-69.

Lo, Y./Mendell, N.R./Rubin, D.B. (2001): Testing the Number of Components in a Normal Mixture. In: Biometrika 88, H. 3, S. 767-778.

Maag Merki, K. (2009): Evaluation im Bildungsbereich Schule in Deutschland. In: Widmer, T./Beywl, W./Fabian, C. (Hrsg.): Evaluation. Ein systematisches Handbuch. Wiesbaden: VS, S. 157-162.

Maier, U. (2008): Vergleichsarbeiten im Vergleich – Akzeptanz und wahrgenommener Nutzen standardbasierter Leistungsmessungen in Baden-Württemberg und Thüringen. In: Zeitschrift für Erziehungswissenschaft 11, H. 3, S. 453-474.

Maier, U./Bohl, T./Kleinknecht, M./Metz, K. (2011): Einflüsse von Merkmalen des Testsystems und Schulkontextfaktoren auf die Akzeptanz und Rezeption von zentralen Testrückmeldungen durch Lehrkräfte. In: Journal for Educational Research Online. Journal für Bildungsforschung Online 3, H. 2, S. 62-93.

Maritzen, N. (2006): Eine Trendanalyse. Schulinspektion zwischen Aufsicht und Draufsicht. In: Buchen, H./Horster, L./Rolff, H.-G. (Hrsg.): Schulinspektion und Schulleitung. Stuttgart: Raabe, S. 7-26.

MBJS (2008a): Orientierungsrahmen Schulqualität in Brandenburg. Qualitätsbereiche und Qualitätsmerkmale guter Schulen/Version 2.0. Potsdam: Ministerium für Bildung, Sport und Jugend.

MBJS (2008b): Schulvisitation im Land Brandenburg. Jahresbericht zum Schuljahr 2006/2007. Potsdam: Ministerium für Bildung, Jugend und Sport.

Meuret, D./Morlaix, S. (2003): Conditions of Success of a School's Self-Evaluation: Some Lessons of an European Experience. In: School Effectiveness and School Improvement 14, H. 1, S. 53-71.

Muthén, L.K./ Muthén, B.O. (1998-2010): Mplus. Statistical Analysis with Latent Variables. User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Nachtigall, C./Jantowski, A. (2007): Die Thüringer Kompetenztests unter besonderer Berücksichtigung der Evaluationsergebnisse zum Rezeptionsverhalten. In: Empirische Pädagogik 21, H. 4, S. 401-410.

Nylund, K.L./Asparouhov, T./Muthén, B. (2007): Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. In: Structural Equation Modeling 14, H. 4, S. 535-569.

Pant, H.A./Emmrich, R./Harych, P./Kuhl, P. (2011): Leistungsüberprüfungen durch Schulleistungsstudien und Vergleichsarbeiten. In: Sacher, W./Winter, F. (Hrsg.): Professionswissen für Lehrerinnen und Lehrer, Bd. 4: Diagnose und Beurteilung von Schülerleistungen. Grundlagen und Reformansätze. Baltmannsweiler: Schneider Verlag Hohengehren, S. 123-142.

R Development Core Team (2011): R: A Language and Environment for Statistical Computing. Wien: R Foundation for Statistical Computing. URL: http://www.R-project.org/; Zugriffsdatum: 10.07.2013.

Rost, J. ([2]2004): Lehrbuch Testtheorie – Testkonstruktion. Bern: Huber.

Schildkamp, K./Vanhoof, J./van Petegem, P./Visscher, A. (2012): The Use of School Self-Evaluation Results in the Netherlands and Flanders. In: British Educational Research Journal 38, H. 1, S. 125-152.

Schildkamp, K./Visscher, A. (2009): Factors Influencing the Utilisation of a School Self-Evaluation Instrument. In: Studies in Educational Evaluation 35, H. 4, S. 150-159.

Schildkamp, K./Visscher, A./Luyten, H. (2009): The Effects of the Use of a School Self-Evaluation Instrument. In: School Effectiveness and School Improvement 20, S. 69-88.

Schrader, F.-W./Helmke, A. (2003): Evaluation – und was danach? Ergebnisse der Schulleiterbefragung im Rahmen der Rezeptionsstudie WALZER. In: Schweizerische Zeitschrift für Bildungswissenschaften 25, H. 1, S. 79-110.

SenBJS (2006): Bildung für Berlin. Handbuch Schulinspektion. Berlin: Senatsverwaltung für Bildung, Jugend und Sport.

SenBWF (2009): Bildung für Berlin. Schulinspektionen im Schuljahr 2007/2008. Berlin: Senatsverwaltung für Bildung, Wissenschaft und Forschung.

SenBWF (2011): Qualitätspaket Kita und Schule. Berlin: Senatsverwaltung für Bildung, Wissenschaft und Forschung. URL: http://www.berlin.de/imperia/md/content/sen-bildung/schulqualitaet/qualitaetspaket.pdf?start&ts=1304676467&file=qualitaetspaket.pdf; Zugriffsdatum: 17.05.2013.

Sommer, N. (2011): Wie beurteilen schulische Gruppen die erlebte Schulinspektion? Ergebnisse einer Befragung. In: Müller, S./Pietsch, M./Bos, W. (Hrsg.): Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht. Münster u.a.: Waxmann, S. 137-164.

Stamm, M. (2003): Evaluation und ihre Folgen für die Bildung. Eine unterschätzte pädagogische Herausforderung. Münster u.a.: Waxmann.

Stern, C./Ebel, C./Müncher, A. (Hrsg.) (^{3}2008): Bessere Qualität in allen Schulen. Praxisleitfaden zur Einführung des Selbstevaluationsinstruments SEIS in Schulen. Gütersloh: Bertelsmann Stiftung.

Thiel, F./Thillmann, K. (2012): Interne Evaluation als Instrument der Selbststeuerung von Schulen. In: Wacker, A./Maier, U./Wissinger, J. (Hrsg.): Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung. Empirische Befunde und forschungsmethodische Implikationen. Wiesbaden: VS, S. 35-55.

Tresch, S. (2007): Potenzial Leistungstest. Bern: hep.

Vanhoof, J./van Petegem, P./de Maeyer, S. (2009): Attitudes Towards School Self-Evaluation. In: Studies in Educational Evaluation 35, H. 1, S. 21-28.

Wurster, S./Gärtner, H. (2013): Schulen im Umgang mit Schulinspektion und deren Ergebnissen. In: Zeitschrift für Pädagogik 59, H. 3, S. 425-445.

Sebastian Wurster, geb. 1984, wissenschaftlicher Mitarbeiter am Institut zur Qualitätsentwicklung im Bildungswesen (IQB).
E-Mail: sebastian.wurster@iqb.hu-berlin.de

Dirk Richter, Dr., geb. 1980, wissenschaftlicher Mitarbeiter am Institut zur Qualitätsentwicklung im Bildungswesen (IQB).
E-Mail: dirk.richter@iqb.hu-berlin.de

Anna Schliesing, geb. 1983, wissenschaftliche Mitarbeiterin am Institut zur Qualitätsentwicklung im Bildungswesen (IQB).
E-Mail: anna.schliesing@iqb.hu-berlin.de

Hans Anand Pant, Prof. Dr., geb. 1962, Direktor des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB).
E-Mail: IQBoffice@iqb.hu-berlin.de

Anschrift: Institut zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt Universität zu Berlin, Unter den Linden 6, 10099 Berlin

DDS – Die Deutsche Schule
Beiheft 12, S. 51-73
© 2013 Waxmann

Isabell van Ackeren/Carmen Binnewies/Marten Clausen/Denise Demski/
Christian Dormann/Anna Rosa Koch/Bastian Laier/Peter Preisendörfer/
Daja Preuße/Christoph Rosenbusch/Uwe Schmidt/Martin Stump/Olga
Zlatkin-Troitschanskaia (Autorengruppe EviS)

Welche Wissensbestände nutzen Schulen im Kontext von Schulentwicklung?

Theoretische Konzepte und erste Befunde des EviS-Verbundprojektes im Überblick[1]

Zusammenfassung
Zur Qualitätsentwicklung im Schulsystem wurden in den vergangenen Jahren erhebliche Anstrengungen unternommen, um wissenschaftliche Grundlagen für eine verlässliche Bestandsaufnahme und Perspektivenentwicklung zu schaffen. Zugleich wurden neue Instrumente einer outputorientierten und evidenzbasierten Steuerung implementiert. Damit liegt ein sehr umfangreicher Pool an Daten und Befunden vor. Verschiedene Berichts- und Feedback-Prozesse sollen dazu beitragen, dass diese Evidenzen einen Input für Akteure im Schulsystem darstellen und in die Prozesse der Leistungsentwicklung einfließen. Im Rahmen des EviS-Verbundes wird untersucht, inwieweit der Anspruch einer evidenzbasierten Schulentwicklung eingelöst wird. Neben zentralen theoretischen Konzepten werden im vorliegenden Beitrag erste empirische Befunde zum evidenzorientierten Klima und zur tatsächlichen Nutzung von Daten und Evidenzquellen auf der Grundlage einer großflächigen Erhebung in Rheinland-Pfalz präsentiert.
Schlüsselwörter: Evidenzbasierung, datengestützte Schulentwicklung, Neue Steuerung

1 Unser Dank gilt dem Bundesministerium für Bildung und Forschung für die Förderung des Verbundprojektes der Universitäten Mainz und Duisburg-Essen mit dem Titel „Evidenzbasiertes Handeln im schulischen Mehrebenensystem – Bedingungen, Prozesse und Wirkungen" (Laufzeit 2010-2013) im Rahmen der Förderung der empirischen Bildungsforschung im Bereich „Steuerung im Bildungssystem" (Förderkennzeichen 01 JG 1010 B).

What Kind of Knowledge Do Schools Use for School Development Purposes?

Theoretical Concepts and Preliminary Findings of the EviS-Project

Abstract

In recent years, considerable efforts have been made to improve the quality of the education system and to provide the scientific basis for a reliable stocktaking of the situation and the development of perspectives. At the same time, new instruments of an output-oriented and evidence-based control have been implemented. Thus, a very large pool of data and findings exists. Reporting and feedback processes have been designed to help ensure that this evidence serves as an input for actors in the education system and feeds into the process of performance development. In the context of the project EviS, we investigate to what extent the claim of evidence-based school development is met. In addition to essential theoretical concepts, this article presents first empirical results regarding the evidence-based climate in schools, the actual use of data and evidence sources. These results are based on an extensive survey in Rhineland-Palatinate.

Keywords: evidence-based / data-based school development, new public management, teacher training

1. Einleitung

Das Thema evidenzbasierter bzw. evidenzinformierter Steuerung und Entwicklung hat im Bildungsbereich an den Schnittstellen von Wissenschaft und Forschung, Politik und Administration, Lehrerbildung und Schulen national wie international Konjunktur. Trotz der vielfältigen und aufwändigen Instrumente Neuer Steuerung im deutschen Bildungssystem konnte bislang nur ein – gemessen an den bildungspolitischen Steuerungsansprüchen – verhaltenes Bild der tatsächlichen Nutzung einzelner Instrumente bzw. der dadurch generierten Ergebnisse in der schulischen Praxis gezeichnet werden. In vielen vorliegenden Studien wird berichtet, dass Lehrkräfte die ihnen zur Verfügung gestellten Evidenzen vergleichsweise wenig für die Unterrichts- und Schulentwicklung nutzen (vgl. z.B. Posch 2009) bzw. die Nutzbarmachung voraussetzungsreich ist (vgl. z.B. Maier/Kuper 2012). Dabei scheint es ein Spannungsfeld zwischen einem von außen an die Schulen herangetragenen Anspruch wissenschaftsorientierter Reflexion sowie einer auch erfahrungs- bzw. intuitionsgeprägten und tradierten Handlungspraxis in Einzelschulen zu geben (vgl. z.B. Berkemeyer/Bos 2009).

Allerdings beziehen sich bisherige Studien in der Regel auf *einzelne* Instrumente evidenzbasierter Steuerung (z.B. Vergleichsarbeiten, Zentrale Abschlussprüfungen und in jüngerer Zeit auch Schulinspektionen) und bieten demnach keine integrierende und vergleichende Gesamtsicht auf die potenziell verfügbaren und nutzbar zu machenden Evidenzquellen. Es stellen sich u.a. die Fragen, welche Evidenzquellen (z.B. Fachzeitschriften, systematisches Schülerfeedback zum Unterricht) vergleichsweise

häufiger genutzt werden und was ggf. statt oder auch in Kombination mit evidenzbasierten Wissensbeständen handlungsleitend ist.

Grundsätzlich lässt sich festhalten, dass im Rahmen der Arbeit mit den Instrumenten der Neuen Steuerung eine große Menge empirisch basierten schulentwicklungs- bzw. steuerungsbezogenen Wissens für die Schulakteure bereitgestellt wird. Die Wirksamkeit von darauf basierenden Schulentwicklungsmaßnahmen setzt dabei notwendigerweise nicht nur die Rezeption des neuen Wissensangebots, sondern gerade auch einen kompetenten Umgang mit dem evidenzbasierten Wissen und seiner Integration in eine reflexive Praxiserfahrung voraus.

Sowohl aus wissenschaftlicher als auch aus anwendungsbezogener Sicht stellt sich dabei die Frage, wie sich ein solches reflexives evidenzgestütztes Handeln entwickelt und durch welche individuellen Voraussetzungen und strukturellen Faktoren eine Förderung und gezielte Unterstützung erfolgen kann. Ein umfassender Literaturüberblick zeigt, dass die Forschungslage zu dieser Frage bislang defizitär ist (vgl. van Ackeren u.a. 2011). Dies betrifft sowohl das theoretische als auch das empirische Wissen über den professionellen Umgang mit (bestimmten) evidenzbasierten Erkenntnissen, über seine Voraussetzungen und seine Entwicklung. Dabei sollte der Genese der Wissensorientierungen und Handlungsschemata über die Lehrerausbildungsphasen hinweg besondere Aufmerksamkeit geschenkt werden.

Diesen Aspekten wird im Rahmen des EviS-Verbundprojektes[2] der Universitäten Duisburg-Essen und Mainz nachgegangen. Nachfolgend werden die zentralen theoretischen Konstrukte des Projektes unter Berücksichtigung des Forschungsstandes vorgestellt (Kapitel 2). Ein eigenes Kapitel ist der zentralen Frage des Transfers von Evidenz in die Schulpraxis und den spezifischen Voraussetzungen des kompetenten Umgangs mit Evidenz gewidmet (Kapitel 3). In Kapitel 4 werden nach einer knappen Vorstellung des EviS-Verbundes erste empirische Befunde präsentiert und im abschließenden Kapitel 5 diskutiert.

2. Wissensbestände: Theoretische Konzepte

Wissensbestände bringen aus einer soziologischen Perspektive betrachtet Ansprüche zum Ausdruck, die von bestimmten Gruppen in bestimmten Kontexten als angemessen erachtet werden (vgl. Maasen 2009, S. 58f.). Hierbei ringen stets verschiedene Wissensformen um Aufmerksamkeit und Dominanz (vgl. Doyle McCarthy 1996). So steht beispielsweise die Schülerbewertung durch eine Lehrkraft als individuelles professionelles Urteil unter Umständen im Widerspruch zu standardisiert erhobenen Testergebnissen. Das Bemühen um evidenzbasierte Bildungspolitik, Schul- und

2 EviS = Evidenzbasierte Schulentwicklung.

Unterrichtsentwicklung ist aus dieser Perspektive als der Versuch zu betrachten, in Schulen Aufmerksamkeit und Akzeptanz für *evidenzbasierte* Wissensbestände zu erhalten. Diese Wissensbestände sollen vorhandenes professionelles Wissen und vorhandene Praxiserfahrungen als bisherige Handlungsgrundlage ergänzen und so zu einer systematischen und intersubjektiven Weiterentwicklung vorhandener Strukturen und Prozesse im Schulsystem beitragen.

2.1 Evidenzbasiertes Wissen und Handeln

Der Begriff der *Evidenz* bezeichnet eine spezifische Form von Wissensbeständen. Um welche Wissensbestände es sich dabei genau handelt, wird allerdings in der aktuellen Diskussion um evidenzbasierte Bildungspolitik, Schul- und Unterrichtsentwicklung oftmals nicht näher definiert, sondern vereinzelt problematisiert und diskutiert, jedoch bisher nicht schlüssig operationalisiert (vgl. dazu Jornitz 2008; Bellmann/Müller 2011). Die ursprünglich vor allem im englischen Sprachraum für die Medizin und das allgemeine Management formulierte Forderung nach verstärkter *Evidenzbasierung* zielt im Kern darauf, organisationale Entscheidungen und Praktiken in Organisationen auf die am besten geeigneten, verfügbaren Fakten zu stützen: „Translating principles based on best evidence into organizational practices" (Rousseau 2006, S. 256). Dementsprechend wird im Bereich des organisationalen Managements vom *evidenzbasierten Handeln* bzw. *evidenzbasierten Management* (EBMgt) gesprochen (vgl. z.B. Briner 2000; Pfeffer/Sutton 2006). Im Wesentlichen ist in der Managementliteratur mit dem Begriff des EBMgt die Anwendung harter, mit wissenschaftlichen Methoden gewonnener, empirischer Befunde in der Praxis gemeint. Sofern die Evidenz mit validen Methoden gewonnen wurde, ist es irrelevant, ob diese Evidenzen durch Forschungsinstitute erhoben wurden oder z.B. durch systematische Befragungen durch Praxisakteure.

Vor diesem Hintergrund und mit Blick auf den Bildungsbereich werden Evidenzen im Folgenden als systematisch generierte, verobjektivierte und explizierte Informationen und Wissensbestände zur Wirksamkeit von Bildungsprozessen und ihren fördernden Rahmenbedingungen verstanden (vgl. Demski u.a. 2012, S. 132).

Evidenzen als spezifische Wissensbestände lassen sich hinsichtlich verschiedener Dimensionen und Charakteristika unterscheiden. So können sie im Hinblick darauf differenziert werden, ob sie *intern* in Schulen oder Klassen generiert werden oder jenseits davon durch *externe* Wissensproduzenten. Weiterhin stellt sich die Frage, inwieweit die betreffenden Wissensbestände durch Anwendung (sozial-)wissenschaftlicher Methoden zustande gekommen sind.

Das Spektrum reicht hier von wissenschaftlich-empirischen Studien bis hin zu kollegialen Hospitationen und Schülerfeedbacks, welche die einzelne Schule ggf. we-

niger standardisiert und nicht explizit unter Beachtung gängiger Gütekriterien der Datenerhebung durchführt. Da wissenschaftlich-methodisch gesammelte Daten im Zentrum der Forderung nach mehr Evidenzbasierung stehen, bietet es sich an, von *Evidenzquellen im engeren Sinne* und von *Evidenzquellen im weiteren Sinne* zu sprechen (vgl. Tabelle 1).

Tab. 1: Evidenzbasiertes Wissen: Herkunft des Wissens und wichtige Charakteristika

	Evidenzquellen im engeren Sinne	**Evidenzquellen im weiteren Sinne**
	eher ausgeprägte Wissenschaftsorientierung, Beispiele	eher geringere Wissenschaftsorientierung, Beispiele
eher formalisiert, explizierter Steuerungsanspruch	Schulleistungsstudien Bildungsberichterstattung Schulinspektion Vergleichsarbeiten Zentrale Abschlussprüfungen	Schülerfeedback zum Unterricht Gemeinsame Unterrichtsentwicklungsmaßnahmen (wie kollegiale Hospitationen) Schulinterne Befragungen Schulstatistische Auswertungen Parallelarbeiten
eher informell, implizites Steuerungspotenzial	externe Bereitstellung von Aufgabensammlungen vermittelte Evidenz über Artikel in schulfachbezogenen oder überfachlichen schulbezogenen Zeitschriften	Bildungsteile von Tages-/Wochenzeitungen oder Magazinen

Quelle: eigene Darstellung

Als ein weiteres wichtiges Klassifikationsmerkmal kann die Aussageebene bzw. der Fokus von Evidenzen gelten: Handelt es sich um Aussagen auf der Ebene des Bildungssystems, spezifischer Organisationen (z.B. Schulen), einzelner Gruppen (z.B. Klassen) oder Individuen (z.B. einzelne Schülerinnen und Schüler)? Darüber hinaus lassen sich Evidenzen bzw. Evidenzquellen bezüglich des Steuerungsanspruches differenzieren. Dieser ist beispielweise bei Informationsquellen wie Schulinspektionen oder auch landesweiten Vergleichsarbeiten vergleichsweise stark ausgeprägt, da die Bildungsadministration hiermit spezifische Steuerungserwartungen verknüpft. Deutlich unverbindlicher sind hingegen Artikel in pädagogischen oder schulfachbezogenen Fachzeitschriften oder auch im Bildungsteil von Zeitungen. Solche Quellen können als *vermittelte Evidenzen* bezeichnet werden (vgl. Demski u.a. 2012). Schließlich lassen sich evidenzbasierte Wissensbestände nach dem Grad der Konkretheit von Handlungsempfehlungen bzw. der intendierten oder tatsächlichen *Handlungsnähe* der Inhalte unterscheiden. Bislang ist weitgehend ungeklärt, welche potenziellen Evidenzquellen besondere Beachtung durch die schulischen Akteure er-

fahren und in welchem Verhältnis die einzelnen Quellen zueinander stehen (vgl. ebd., S. 132f.).

2.2 Zur Bedeutung von Praxis- bzw. Erfahrungswissen in Ergänzung evidenzbasierter Wissensbestände

Im Sinne erziehungswissenschaftlicher Professionsvorstellungen hat die umfassende Analyse konkreter und fallspezifischer pädagogischer Gegenstände und Situationen (idiographischer Ansatz) besondere Bedeutung. Dabei geht es weniger um die Ableitung allgemeingültiger Aussagen und die Abstraktion von einem spezifischen Phänomen (nomothetischer Ansatz). Im pädagogischen Handeln erweist sich insofern auch das *Erfahrungswissen* schulischer Akteure als relevante Voraussetzung der Handlungskompetenz von Lehrerinnen und Lehrern (vgl. Gruber 2006), bei dem erlebte Sachverhalte und Vorgehensweisen gleichwohl systematisch und reflexiv in einen Erkenntniszusammenhang gestellt werden. Beobachtungen und Wahrnehmungen, die sich auf konkrete Situationen beziehen, werden individuell konstruiert, interpretiert und in Folgerungen überführt. Folglich ist pädagogisches Handeln nicht allein auf der Grundlage quantifizierender Forschung zu fundieren, sondern auch aus der erfahrungsgesättigten, rekonstruierenden Reflexion des Handelns.

Erfahrungswissen kann gleichwohl in pädagogischen Kontexten durch evidenzbasiertes Wissen im Sinne eines integrierenden Ansatzes fundiert und erweitert werden. Dem Anspruch neuer Steuerung nach können evidenzbasierte Wissensbestände dazu beitragen, einer unmittelbaren, nicht weiter hinterfragten und nicht reflektierten Nachahmung von Vorgehensweisen und Strategien entgegenzuwirken (vgl. Dormann/Dormann 2010). Pfeffer und Sutton (2006) grenzen aus einer wirtschaftswissenschaftlichen Perspektive das von ihnen begrifflich eingeführte *Ersatzwissen* oder *substitutorientierte Wissen* (substitute) gegenüber einem evidenzbasierten Wissen im Sinne einer „science-based practice" ab. Das Ersatzwissen umfasst den Bezug auf Konzepte als Sammlung von Leitideen, die ohne exakten Nachweis ihrer Wirksamkeit mittels gezielter und systematischer Analyse angewendet werden. Evidenzbasiert ist demgegenüber Handeln, wenn es zugleich in hohem Maße auf Evidenzwissen und in geringem Maße auf dem in dieser Weise gekennzeichneten Substitut basiert. Es ist anzunehmen, dass in Organisationen bei vielen Entscheidungen – z.B. aufgrund hohen Zeit- und Handlungsdrucks, Tradition, fehlender Ressourcen oder auch geringer Kenntnisse – solches Ersatzwissen herangezogen wird. Dies kann geschehen, obwohl Faktenwissen prinzipiell vorhanden ist, etwa im Hinblick auf empirische Befunde der Lehr-Lernforschung oder Erkenntnisse zu effektivem Schulleitungshandeln, die gleichwohl situationsangemessen und reflexiv adaptiert werden müssen und sich nicht unmittelbar bzw. nur sehr eingeschränkt übertragen lassen.

2.3 Schulisches Handeln im Spannungsfeld von Evidenz- und „Ersatzwissen"-Orientierung

Unter *schulischer Evidenzorientierung* werden einem breiten Begriffsverständnis folgend (s.o.) die zumindest versuchten Vorgehensweisen und Praktiken in Schulen verstanden, die der Bereitstellung bzw. Beschaffung und Umsetzung verfügbarer, auf empirischen Belegen beruhender Wissensbestände für schulisches Entscheidungshandeln in spezifischen Handlungskontexten dienen (im Sinne eines Willens zum evidenzorientierten Handeln). Hingegen beschreibt *schulische Ersatzwissensorientierung* Vorgehensweisen und Praktiken in Schulen, durch die nicht gesicherte bzw. nicht systematisch generierte Wissensbestände ohne kritisch-reflexive Prüfung in schulisches Handeln umgesetzt werden. Der Soziologe Eröss spricht am Beispiel des Schul- und des Gesundheitswesens von der „Substitution (Ersatz) des […] Wissens durch andere Wissensformen, die […] für bestimmte Akteure zuverlässiger, ‚sicherer' (z.B. lokales Wissen), verfügbarer, konkreter […], legitimer […] oder authentischer (z.B. charismatisch inszeniertes, ‚verkörpertes' Wissen) erscheinen" (Eröss 2009, S. 373).

Schulisches evidenzbasiertes Handeln ist im Rahmen dieser Terminologie dann gegeben, wenn die schulische Evidenzorientierung der Organisation Schule insgesamt hoch und die schulische Ersatzwissensorientierung vergleichsweise gering ausgeprägt sind. Ist umgekehrt die Ersatzwissensorientierung hoch und wird gleichzeitig (weitgehend) jede Form von (verfügbarer) Evidenz ausgeblendet bzw. wird nicht nach verfügbaren Wissensbeständen dieser Art gesucht, kann man von einem *evidenzabgewandten Handeln* sprechen. Weiterhin wird hier davon ausgegangen, dass eine Schule prinzipiell sowohl eine hohe Evidenzorientierung als auch eine hohe Ersatzwissensorientierung aufweisen kann. Eine solche Konfiguration reflektiert das Bemühen, jede verfügbare Form von Information zu nutzen, um sie für Veränderungen in schulischen Prozessen und Vorgehensweisen einzusetzen. Hier kann von einem *volatilen Handeln* gesprochen werden. Schließlich kann es auch sein, dass beide Aspekte nur in geringem Ausmaß erkennbar sind; ein Bestreben, überhaupt Informationen zu beschaffen und für die Steuerung der Schule umzusetzen, ist hier nicht erkennbar (*sklerotisches Handeln)*.

Tab. 2: Unterschiedliche Entscheidungs- und Handlungsformen einzelschulischer Organisationen in Abhängigkeit vom Ausmaß der Evidenz- bzw. Ersatzwissensorientierung

		Evidenzorientierung (nachweisorientiertes Wissen) *(Beispielitem: „Unsere Unterrichtsgestaltung berücksichtigt stets die neuesten wissenschaftlichen Erkenntnisse.")*	
		gering	hoch
Ersatzwissensorientierung (Wissen aus alternativen Quellen) *(Beispielitem: „Die Aussage ‚Das haben wir immer schon so gemacht' ist Grundlage vieler Entscheidungen.")*	gering	*sklerotisch*	*evidenzbasiert*
	hoch	*evidenzabgewandt*	*volatil*

Quelle: eigene Darstellung

Basierend auf einer breit angelegten Studie in öffentlichen Verwaltungen arbeiteten Stumm, Dormann und Ohlemann (2010) heraus, dass *organisationale Evidenzbasierung* ein mehrdimensionales Konstrukt ist, welches sich durch die drei Faktoren *externe Evidenzorientierung* (Berücksichtigung wissenschaftlicher Erkenntnisse), *interne Evidenzorientierung* (Berücksichtigung von Fakten zur eigenen Organisation) und *Ersatzwissensorientierung* (Berücksichtigung von Bauchgefühl und persönlicher Präferenz) beschreiben lässt.

3. Voraussetzungsvoller Transfer: Wissensdiffusion und kompetenter Umgang mit Evidenz

Im Laufe des letzten Jahrzehnts wurde in der Literatur vielfach der (mangelnde) Transfer von Evidenz in die Praxis thematisiert. So spricht z.B. Rynes (2007) in diesem Zusammenhang von „great divide" oder Latham (2007) von „research-practice gap". Mangelnder Transfer von Evidenz in die Praxis, etwa durch die „Entkopplung von Wissensproduktionsort und Wissensnutzungsort" (Eröss 2009, S. 378), lässt sich von zumindest zwei Seiten betrachten.

Zum einen ist auf Seiten der Sender, also bei jenen, die über Evidenzen verfügen, zu beklagen, dass sie diese unzureichend aufbereiten bzw. verbreiten. Diskussionswürdig ist daher die Frage der optimierten Verbreitung von evidenzbasiertem Wissen, z.B. über angepasste Kommunikationsformen und integrierte Interdisziplinarität (vgl. Bartunek 2007; Cummings 2007; Guest 2007). Dies beschreibt Latham (2007, S. 1028) als „*theory of diffusion*".

Auf der anderen Seite der Empfänger geht es darum, wie die *Rezeption* von Evidenz – z.B. durch den Kontaktaufbau zu Wissenschaftlerteams oder geeignete Aus-, Fort- und Weiterbildungsmaßnahmen – verbessert werden kann *(„theory of infusion"*; vgl. Stumm/Dormann/Ohlemann 2010). Diese Seite des Transfers wird in der Managementliteratur und entsprechender Forschung besonders problematisiert, da hier herausgearbeitet wurde, dass die Implementierung bereits vorhandenen Faktenwissens in der Praxis oft bewusst unterbunden wird. Hierbei kommt Entscheidungsträgern in Organisationen besondere Bedeutung zu. Wissenschaftliche Evidenz gilt bei vielen Führungskräften (hier: im außerschulischen Bereich) als nicht anwendbar, nicht glaubwürdig, zu kompliziert und als unbequem (vgl. z.B. Briner 2007). Zugleich gehen die meisten Autoren im Sinne des oben genannten Diffusionsansatzes davon aus, dass evidenzbasiertes Management wenig verbreitet ist.

3.1 Innerschulische Diffusion von Wissensbeständen

Für die Diffusion von Wissensbeständen, seien es – wie in zahlreichen Studien – Innovationen (für eine Übersicht vgl. Greenhalgh u.a. 2004) oder – wie im Rahmen dieses Beitrags – Evidenzen, werden Informationskanäle benötigt. Interaktions- und Kommunikationsbeziehungen zwischen Akteuren können diese Funktion erfüllen. Es ist weiterhin anzunehmen, dass die Verbindungen zwischen Akteuren nicht nur bei der Diffusion, sondern auch bei der Selektion und Akzeptanz von Wissensbeständen eine tragende Rolle spielen. Übertragen auf den schulischen Kontext bedeutet dies, dass z.B. Steuer- und Fachgruppen, aber auch informelle soziale Netzwerke innerhalb einer Schule und über die Schulgrenzen hinweg einen zentralen Bestimmungsfaktor für die Rezeption und Nutzung evidenzbasierten Wissens darstellen. Insofern kann man davon ausgehen, dass sowohl den formellen als auch den informellen Beziehungen der schulischen Akteure eine besondere Bedeutung für den Erhalt, die Akzeptanz und die Umsetzung neuer Informationen zukommt (vgl. Borgatti/Cross 2003).

Für die Untersuchung der Beziehungen zwischen Akteuren steht der sozialwissenschaftlichen Forschung der netzwerkanalytische Ansatz zur Verfügung (vgl. Ullmann/Stepancik 2009). Zum einen kann dabei ein Fokus auf die persönlichen bzw. egozentrierten Netzwerke gelegt werden, und zum anderen kann das gesamte Netzwerk einer Schule in den Blickpunkt der Betrachtung rücken (vgl. Jansen 2006; Wasserman/Faust 2009; Hennig u.a. 2012). Die beiden Perspektiven unterscheiden sich insofern, als bei der Betrachtung eines innerschulischen Gesamtnetzwerks a priori festgelegte Netzwerkgrenzen bestehen. Hier werden ausschließlich Beziehungen innerhalb einer Schule in Betracht gezogen. Bei der Analyse persönlicher Netzwerke sind Personen in der Wahl ihrer Netzwerkpartner frei, so dass eine größere Reichweite von Netzwerkpartnern erfasst werden kann. So können außerschulische Kontakte, wie z.B. zu Lehrpersonen anderer Schulen, erhoben werden. Die Betrachtung eines

Gesamtnetzwerks eröffnet dagegen die Möglichkeit, die Struktur eines Netzwerks zu analysieren und so Rückschlüsse über Informationsflüsse und Wissensaustausch innerhalb einer Schule zu ziehen (vgl. Rehrl/Gruber 2007). Kulin u.a. (2012) beschreiben die Analyse egozentrierter Netzwerke als Desiderat der Bildungs- und Schulforschung.

3.2 Kompetenter Umgang mit Evidenz

Neben dem Zugang zu evidenzbasierten Wissensbeständen, der Art ihrer Aufbereitung und Kommunikation (hier nicht vertieft; vgl. z.B. Peek 2006) in die Schule sowie dem Austausch darüber innerhalb der Schule besteht in der Literatur Einigkeit darüber, dass die handlungsleitende Funktion und somit die Wirksamkeit outputorientierter Steuerungsinstrumente ganz entscheidend von der spezifischen Kompetenz der Schulakteure abhängt, mit evidenzbasiertem Wissen professionell umzugehen. Hierbei wird in der Regel differenziert zwischen den Fragen, ob Lehrende für die Umsetzung evidenzbasierten Wissens kompetent sind und ob sie bereit sind, professionell zu handeln (vgl. z.B. die Befunde in Zlatkin-Troitschanskaia/Buske 2009; Zlatkin-Troitschanskaia/Kadach 2008). Damit richtet sich der Fokus auf die individuellen Voraussetzungen der Schulakteure in Form von kognitiven Dispositionen (wie Wissen), emotionalen Aspekten (z.B. geringe Ängstlichkeit) sowie nicht-kognitiven Bereitschaften (wie Interesse, Motivation).[3]

Das Verständnis von Evidenz stellt eine zentrale Bestimmungsgröße dar, welches in der entwicklungspsychologischen Literatur unter dem Begriff *wissenschaftliches Verständnis* oder *epistemologische Überzeugung* diskutiert wird (vgl. z.B. Kuhn 1989). Der Aufbau angemessener Vorstellungen über Evidenz wird unter Bezugnahme auf Theorien zur Genese und Funktion epistemologischer Überzeugungen betrachtet (vgl. Hofer/Pintrich 1997). Diese Arbeiten belegen die Bedeutsamkeit epistemologischer Überzeugungen im Kontext von Wissensrezeption und -nutzung (vgl. z.B. King/Kitchener 1994).

Zum Wissenschaftsverständnis von Lehrkräften und seiner Entwicklung liegen nur vereinzelt Studien vor (vgl. Günter u.a. 2004; Carey u.a. 1989). Jedoch konnte gezeigt werden (vgl. Hofer 2001), dass professionelles Handeln von Lehrkräften durch ihre epistemologischen Überzeugungen nennenswert beeinflusst wird (vgl. auch Pepin 1999).

3 Selbst wenn diese Voraussetzungen gegeben sind, muss noch eine weitere, notwendige Bedingung erfüllt sein, nämlich ein Arbeitsumfeld, welches überhaupt die Möglichkeit bietet, evidenzbasiert zu handeln. Vorgesetzte müssen dabei als Vorbild für evidenzbasiertes Handeln dienen und entsprechende Gelegenheiten bieten, dies zu ermöglichen oder sogar einzufordern (vgl. Dormann u.a. 2009).

4. Empirische Fundierung: Erste deskriptive Befunde

Ziel des EviS-Projektes zur evidenzbasierten Schulentwicklung ist die Gewinnung belastbarer Erkenntnisse zu Wirkungen und Wirksamkeit einer evidenzbasierten Steuerungspraxis im deutschen Schulwesen. Hierbei liegt ein besonderes Augenmerk auf der Operationalisierung und Deskription evidenzbasierten Wissens und Handelns in Schulen. Darüber hinaus geht es um die Identifizierung personeller und organisationaler Faktoren bzw. Faktorenkonstellationen, die ein evidenzbasiertes professionelles Handeln schulischer Akteure begünstigen oder behindern.

Das Forschungsdesign umfasst drei aufeinander bezogene Projekte: ein Zentral- und zwei Teilprojekte. Das Zentralprojekt ist als standardisierte Fragebogenerhebung zur Rezeption und Nutzung von Evidenzen mit rheinland-pfälzischen Lehrkräften (N=2640) und Schulleitungen (N=297) durchgeführt worden.

Ein Teilprojekt vertieft und ergänzt das Zentralprojekt, indem ausgewählte, bereits im Zentralprojekt untersuchte Schulen durch leitfadengestützte Interviews und die Erhebung der schulischen Gesamtnetzwerke einer tiefergehenden Betrachtung zu ihrem Umgang mit Evidenzquellen bzw. Evidenzen unterzogen werden. Im Rahmen des anderen Teilprojekts wird das Zentralprojekt ergänzt durch die Betrachtung der Entwicklung des professionellen Umgangs mit Evidenzen bei (angehenden) Lehrkräften. In einer längsschnittlich angelegten Studie wurden zum Zeitpunkt t1 1.272 Studierende der Bildungswissenschaften und Wirtschaftspädagogik sowie 328 Referendare und Referendarinnen zur wahrgenommenen Evidenzorientierung in der Ausbildung sowie zu ihren Einstellungen und Kenntnissen in diesem Bereich befragt. Zum Zeitpunkt t2 werden diese Probanden und Probandinnen erneut befragt, wobei sie sich in der Regel auf der nächsten Professionalisierungsstufe befinden (z.B. Vorbereitungs- oder Schuldienst).

Vor dem Hintergrund der dargestellten theoretisch-konzeptionellen Grundlagen werden im Weiteren auf der Basis der bislang ausgewerteten Daten erste deskriptive Befunde insbesondere zu folgenden beiden Fragstellungen präsentiert:

- Inwieweit können evidenzorientiertes und eher evidenzabgewandtes (substitutorientiertes, s.o.) Handeln in Schulen – aus Kollegiums- bzw. Schulleitungsperspektive – differenziert werden? (Kapitel 4.1)
- Welchen Stellenwert haben dabei einzelne Evidenzquellen im Vergleich miteinander? Inwiefern gibt es Unterschiede zwischen Schulleitung und Lehrkollegium? (Kapitel 4.2)

4.1 Evidenzorientiertes Handeln?

Mittlerweile ist es eine weit verbreitete Auffassung, dass Schulen sich zu einem substanziellen Anteil selbst steuern müssen, um in ihrem spezifischen Kontext und den sich wandelnden Rahmenbedingungen erfolgreich zu sein. Dabei können – so die Annahme – evidenzinformierte Handlungsmechanismen und -entscheidungen förderlich sein.

4.1.1 Methodisches Vorgehen

Im Rahmen des EviS-Zentralprojekts wurde die Ebene der Schule als Organisation fokussiert. Methodisch geschah dies in zweifacher Hinsicht: zum einen durch die Einschätzung der Schulleitungen und zum anderen durch das gesamte Kollegium, um der Frage nachzugehen, ob die Schule u.a. ein evidenzorientiertes Klima aufweist. Dies sollte sich einerseits darin zeigen, dass die Schulleitung evidenzorientiert vorgeht. Zum Beispiel sollte eine evidenzorientiert handelnde Schulleitung angeben, dass sie neue Methoden vor ihrer Einführung auf verfügbare Befunde zu deren Wirksamkeit prüft. Andererseits sollte sich ein evidenzorientiertes Klima im Sinne geteilter Wahrnehmungen der Lehrkräfte hinsichtlich des Ausmaßes zeigen, in dem an ihrer Schule evidenzbasiert bzw. -abgewandt gehandelt wird. Die geteilte Wahrnehmung von Prozessen, Prozeduren, Vorgehensweisen etc. ist in der Organisationsforschung als *Organisationsklima* definiert. Herrscht ein ausgeprägtes evidenzorientiertes Klima an einer Schule, sollten die Lehrkräfte zum Beispiel die Meinung teilen, dass es an der Schule erwünscht ist, sich mit Fachliteratur auf dem Laufenden zu halten.

Die Analyse orientierte sich sowohl konzeptionell als auch methodisch am Modell von Dormann u.a. (2009). Im Einklang mit der Studie von Stumm, Dormann und Ohlemann (2010) wurde die organisationale Evidenzbasierung mit einem an den Schulkontext angepassten Instrument erhoben, das sowohl von Schulleitungen als auch von Lehrpersonen beantwortet wurde.

4.1.2 Befunde

Exploratorische Faktorenanalysen sowohl der Schulleitungsdaten als auch der Daten der Lehrkräfte zeigen eine sehr gute Übereinstimmung mit der Faktorenstruktur von Stumm, Dormann und Ohlemann (2010; vgl. auch Kapitel 2.3). Insgesamt zeigen sich drei Faktoren: Der erste Faktor entspricht der *internen Evidenzorientierung*. Items, die hohe Ladungen auf dem Faktor interne Evidenzorientierung haben, sind in Tabelle 3 exemplarisch dargestellt, ebenso für den zweiten Faktor der *externen Evidenzorientierung* sowie den dritten Faktor der *Ersatzwissensorientierung*.

Tab. 3: Faktoren der Evidenz- bzw. Ersatzwissensorientierung: Bedeutung und Beispielitems im Kontext des EviS-Projektes

Faktor	Bedeutung	Beispielitem Lehrkräfte	Beispielitem Schulleiter/in
Interne Evidenzorientierung	Berücksichtigung von Fakten zur eigenen Schule bei Entscheidungen	Bevor unsere Schule ein Verfahren von einer anderen Schule übernimmt, prüfen wir, inwiefern sich die Rahmenbedingungen in den Schulen ähneln.	Bevor ich ein Verfahren von einer anderen Schule übernehme, prüfe ich, inwiefern sich die Rahmenbedingungen in den Schulen ähneln.
Externe Evidenzorientierung	Berücksichtigung wissenschaftlicher Erkenntnisse bei Entscheidungen	Unsere Schulleitung sucht den direkten Kontakt zu Wissenschaftler/inne/n, um ihre Entscheidungsqualität zu verbessern.	Ich suche den direkten Kontakt zu Wissenschaftler/inne/n, um meine Entscheidungsqualität zu verbessern.
Ersatzwissensorientierung (i.S.v. Wissen aus anderen Quellen)	Berücksichtigung von Intuition und persönlichen Präferenzen bei Entscheidungen	An unserer Schule geht man davon aus, dass man dem Bauchgefühl bei wichtigen Entscheidungen immer vertrauen kann.	Meinem Bauchgefühl kann ich bei wichtigen Entscheidungen immer vertrauen.

Quelle: eigene Darstellung

Die Befragten konnten die Aussagen zur Erfassung dieser Orientierungen auf einer Skala von 1 (= *stimme überhaupt nicht zu*) bis 5 (= *stimme voll zu*) bewerten. Es zeigte sich insgesamt, dass die *Mitglieder des Lehrerkollegiums* an ihren Schulen die externe Evidenzorientierung als am geringsten ausgeprägt sahen (Mittelwert M=2,51). Die Ersatzwissensorientierung stellte sich als signifikant stärker ausgeprägt dar (M=2,84), die ihrerseits wiederum signifikant geringer war als die interne Evidenzorientierung (M=3,43). Die *Schulleitungsmitglieder* sahen ihre externe Evidenzorientierung höher als die Mitglieder des Kollegiums (M=2,81) und etwas besser als ihre eigene Ersatzwissensorientierung (M=2,76). Aber auch die Schulleitungen stuften ihre interne Evidenzorientierung am höchsten ein (M=4,15).

In einem nächsten Schritt wurden die beiden Evidenzorientierungsfaktoren angesichts ihrer substanziellen Korrelation zusammengefasst (gleich gewichtet) und Schulen mit hoher Evidenzorientierung identifiziert (jene, die sich im Durchschnitt als „eher" (≥3) evidenzorientiert eingeschätzt haben). Ebenso wurden Schulen mit hoher Ersatzwissensorientierung bestimmt (jene, die sich im Durchschnitt als „eher" (≥3) ersatzwissensorientiert eingeschätzt haben) und dann die prozentualen Anteile der Schulen bestimmt, deren Leitung – gemäß Tabelle 2 – „typisierend" als sklerotisch, evidenzabgewandt, volatil oder evidenzbasiert bezeichnet werden kann.

Die Ergebnisse sind in Tabelle 4 zusammengefasst. Über die Hälfte (53,4 Prozent) der Schulleitungen schätzen ihre Führungsweise demnach als evidenzbasiert ein, und weiteren 31,7 Prozent kann ein volatiles Leitungshandeln attestiert werden. Auf der anderen Seite zeigt sich, dass sowohl das sklerotische Schulmanagement mit 6,9 Prozent als auch das evidenzabgewandte Schulmanagement mit 7,9 Prozent aus Sicht der Schulleitungsmitglieder nur selten wiederzufinden sind.

Zwar attestieren auch die meisten Lehrkräfte ihren Schulen, dass sie evidenzbasiert geleitet werden (32,5 Prozent), aber die „klimatische" Sicht bleibt doch weit hinter der Einschätzung der Schulleitung zurück. Darüber hinaus attestieren die Lehrkräfte ihren Schulen weitaus häufiger ein sklerotisches (22,9 Prozent) oder evidenzabgewandtes (25,7 Prozent) Leitungshandeln. Ein volatiles Schulmanagement wird aus Sicht der Lehrer nur in 18,9 Prozent der Fälle praktiziert.

Insgesamt beschreiben Schulleitungsmitglieder in 14,8 Prozent der Fälle eine geringe Evidenzbasierung. Dies tun hingegen 48,6 Prozent der Kollegiumsmitglieder.

Tab. 4: Empirische Ausprägungen der unterschiedlichen Formen des eingeschätzten Schulleitungshandelns

		Evidenzorientierung (nachweisorientiertes Wissen)	
		gering	hoch
		Schulleitungsperspektive	
Ersatzwissens-orientierung (Wissen aus anderen Quellen)	gering	sklerotisch 6,9%	evidenzbasiert 53,4%
	hoch	evidenzabgewandt 7,9%	volatil 31,7%
		Lehrerperspektive	
Ersatzwissens-orientierung (Wissen aus anderen Quellen)	gering	sklerotisch 22,9%	evidenzbasiert 32,5%
	hoch	evidenzabgewandt 25,7%	volatil 18,9%

Quelle: eigene Berechnungen

4.2 Welche Evidenzquellen werden tatsächlich genutzt?

Für einen erfolgreichen Transfer wissenschaftlicher Befunde in die Praxis ist es von besonderem Interesse, welche Informationsquellen und Instrumente von Lehrkräften und Schulleitungsmitgliedern als grundsätzlich nützlich für die Schul- und Unterrichtsentwicklung angesehen und welche Quellen vergleichsweise intensiv für

die Ausgestaltung der eigenen schulischen Arbeit tatsächlich genutzt werden. Um nicht a priori wirksame Wege des Transfers verobjektivierter Befunde in die Schul- und Unterrichtspraxis auszuschließen, folgt das Projekt EviS explizit einem weit ausgelegten Verständnis möglicher Evidenzquellen und evidenzbasierten Handelns (vgl. Kapitel 2 sowie Demski u.a. 2012, S. 140f.).

4.2.1 Methodisches Vorgehen

Auf der Mikro-Ebene der einzelnen Schulleitungsmitglieder und Lehrkräfte wurden mit Hilfe von fünfstufigen Likert-Skalen (1 = *gar nicht nützlich* bis 5 = *sehr nützlich*) unterschiedliche Evidenzquellen differenziert nach folgenden Aspekten erfasst:

- die prinzipielle Einschätzung der Nützlichkeit der Quelle für die Schul- und Unterrichtsentwicklung,
- die Intensität der Auseinandersetzung mit dem Instrument bzw. der Quelle sowie
- die tatsächliche Nutzung der Evidenzquelle für die eigene schulische Arbeit.

Ausgehend von Ansätzen des Informationsmanagements und der Wissensdiffusion wurden weiterhin erfasst:

- die Verfügbarkeit der Evidenzquelle in der Schule und
- die Weiterempfehlung dieser an andere Kollegiumsmitglieder.

4.2.2 Befunde

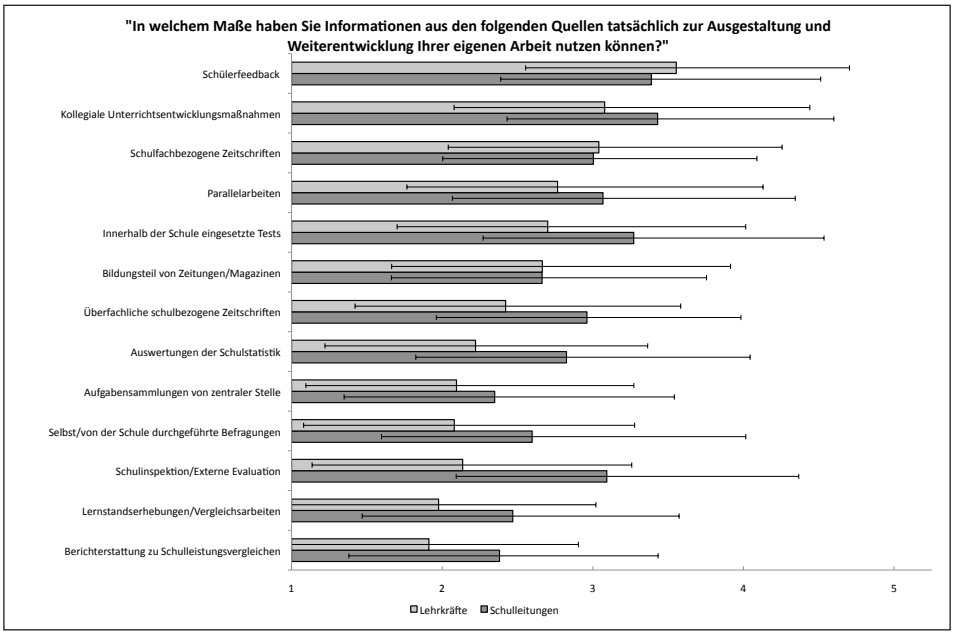
Abbildung 1 stellt die *Nutzung* der verschiedenen Evidenzquellen durch Lehrkräfte und Schulleitungsmitglieder dar. Diese detaillierte Erfassung konnte für alle 297 Schulleitungsmitglieder in der Stichprobe sowie für eine Teilstichprobe von 1.230 Lehrkräften realisiert werden.

Die Nutzung für die eigene Arbeit ist bei den *Lehrkräften* vor allem beim Schülerfeedback (M=3,55), bei kollegialen Unterrichtsentwicklungsmaßnahmen (M=3,08) und schulfachbezogenen Zeitschriften (M=3,04) vergleichsweise hoch. Es folgen Parallelarbeiten (M=2,77), innerhalb der Schule eingesetzte Tests (M=2,70) und der Bildungsteil von Zeitungen oder Magazinen (M=2,67). Mit einem Mittelwert von 2,42 spielen überfachliche schulbezogene Zeitschriften eine deutlich geringere Rolle als Zeitschriften mit Schulfachbezug. In vergleichsweise geringem Maße werden von den befragten Lehrkräften die Schulstatistik (M=2,22), Befunde der Schulinspektion (M=2,14[4]), Aufgabensammlungen von zentraler Stelle (M=2,10), selbst oder von der Schule durchgeführte Befragungen (M=2,08) sowie die Ergebnisse aus landesweiten Lernstandserhebungen (M=1,98) oder Schulleistungsvergleichen (M=1,91) genutzt.

4 Hier wurden nur Personen berücksichtigt, die angeben, dass bereits eine Rückmeldung durch die Schulinspektion erfolgte.

Insgesamt zeichnen die befragten *Schulleitungen* ein etwas positiveres Bild der Evidenznutzung. Auch die Schulleitungsmitglieder geben an, für die Ausgestaltung ihrer eigenen Arbeit am intensivsten auf kollegiale Unterrichtsentwicklungsmaßnahmen (M=3,43) und das Schülerfeedback (M=3,39) zurückzugreifen. Anders als bei den Lehrkräften liegt die durchschnittliche Nutzung von Tests innerhalb der Schule oder des Unterrichts (M=3,27), der Ergebnisse der Schulinspektion (M=3,09) sowie von Parallelarbeiten (M=3,07) über dem Skalenmittelpunkt von 3.

Abb. 1: Nutzung verschiedener Evidenzquellen durch Lehrkräfte (N=1.230) und Schulleitungen (N=297)



Quelle: eigene Berechnungen

Auffällige Diskrepanzen zwischen der Einschätzung der Lehrkräfte und der Schulleitungsmitglieder zeigen sich insbesondere bei der Ergebnisrückmeldung der Schulinspektion, die von den Schulleitungen im Mittel bedeutend intensiver in den Blick genommen wird; in geringerem Maße gilt dies auch für schulstatistische Auswertungen (M=2,83). Während sich bei den Lehrkräften ausgeprägte Unterschiede feststellen lassen, spielen schulfachbezogene (M=3,00) und überfachliche schulbezogene Zeitschriften (M=2,96) für die Schulleitungsmitglieder eine fast identische Rolle auf mittlerem Niveau für die Ausgestaltung und Weiterentwicklung der eigenen Arbeit. Die Nutzung des Bildungsteils von Magazinen oder Zeitungen ist bei Lehrkräften und Schulleitungen (M=2,66) gleich intensiv. Selbst oder von der Schule durchgeführte Befragungen sind wiederum Instrumente, die für Schulleiterinnen und Schulleiter eine größere Relevanz zu haben scheinen.

Zwar nutzen die befragten Schulleitungsmitglieder Ergebnisse aus Lernstandserhebungen (M=2,47) und Schulleistungsvergleichen (M=2,38) in stärkerem Maße als die Lehrkräfte, jedoch scheinen sie auch für die Schulleitungen von nachrangiger Bedeutung im Hinblick auf die Schul- und Unterrichtsentwicklung zu sein. Am wenigsten nutzen die Schulleiterinnen und Schulleiter nach eigenen Angaben Aufgabensammlungen, wie sie von zentraler Stelle, beispielsweise durch das Institut zur Qualitätsentwicklung im Bildungswesen (IQB), zur Verfügung gestellt werden (M=2,35).

Ausgehend von den deskriptiven Befunden erscheint es sinnvoll, die Evidenzquellen nicht nur auf Basis inhaltlicher Überlegungen zu systematisieren, sondern auch statistisch begründet zu gruppieren. Über alle Evidenzquellen hinweg wurden die Items des Lehrerfragebogens zur eingeschätzten Nützlichkeit sowie zur Intensität der Auseinandersetzung und der Nutzung dieser Quellen für die eigene Arbeit einer explorativen Faktorenanalyse unterzogen. Auf Basis der Hauptkomponentenanalyse lassen sich drei Faktoren extrahieren, wobei nicht alle Evidenzquellen ausschließlich auf einem Faktor laden (vgl. Tabelle 5).

Auf einem *ersten Faktor* laden mit der Berichterstattung zu Schulleistungsvergleichen, Lernstandserhebungen, Auswertungen der Schulstatistik, der Schulinspektion und Aufgabensammlungen von zentraler Stelle allesamt Evidenzquellen, die von außen an die Schulen herangetragen werden. Gemeinsam ist weiterhin eine eher als outputorientiert anzusehende Ausrichtung mit vorwiegend formalisiertem Charakter und explizitem Steuerungsanspruch. Auf einem *zweiten Faktor* laden schulfachbezogene und überfachliche schulbezogene Zeitschriften sowie Bildungsteile von Zeitungen oder Magazinen. Hier liegen die Befunde in aufbereiteter Form vor, sodass diese Informationsquellen als vermittelte Evidenzen (vgl. Kapitel 2.1) bezeichnet werden können. Der *dritte Faktor* schließlich vereint Evidenzquellen, denen eine interne Generierung sowie ein starker Bezug zur eigenen Unterrichtspraxis gemein sind. Hierzu zählen selbst oder von der Schule durchgeführte Befragungen oder Tests innerhalb der Schule oder des Unterrichts, das Schülerfeedback, kollegiale Unterrichtsentwicklungsmaßnahmen sowie schulbezogene Parallelarbeiten innerhalb von Jahrgangsstufen.

Die nicht durchgängig eindeutige Differenzierung insbesondere zwischen den Faktoren 1 und 3 könnte ein Hinweis auf unterschiedliche Deutungen der Lehrkräfte sowie eine noch zu ungenaue Operationalisierung der Konstrukte sein. So können beispielsweise „innerhalb der Schule oder des Unterrichts durchgeführte Tests" sowohl in starkem Maße formalisierte, von externer Stelle zur Verfügung gestellte Instrumente als auch selbst entwickelte und ad hoc verwendete Verfahren umfassen. Wie sich in vertiefenden qualitativen Interviews bestätigte, verwenden zumindest einige der befragten Lehrkräfte die Begriffe *Vergleichsarbeiten* und *Parallelarbeiten* sy-

nonym, was die Ladungen letzterer sowohl auf dem ersten als auch auf dem dritten Faktor erklären könnte.

Tab. 5: Ergebnisse der Faktorenanalyse – Zuordnung der Evidenzquellen

	Nützlichkeit	Auseinander-setzung	Nutzung	Zuordnung
Berichterstattung zu Schulleistungsvergleichen	F1	F1	F1	F1
Lernstandserhebungen / Vergleichsarbeiten	F1	F1	F1	F1
Schulinspektion	F1	F1	F1	F1
Auswertung der Schulstatistik	F1	F1/F2	F1	F1
Aufgabensammlungen von zentraler Stelle	F1/F2	F1/F2	F1	F1
Schulfachbezogene Zeitschriften	F2	F2	F2	F2
Überfachliche schulbezogene Zeitschriften	F2	F2	F2	F2
Bildungsteil von Zeitungen / Magazinen	F2	F2	F2	F2
Selbst / von der Schule durchgeführte Befragungen	F1/F3	F3	F1	F3
Schülerfeedback zum Unterricht	F3	F3	F2/F3	F3
Innerhalb der Schule / des Unterrichts eingesetzte Tests	F1/F3	F1/F3	F3	F3
Kollegiale Unterrichts-entwicklungsmaßnahmen	F3	F3	F3	F3
Schulbezogene Parallelarbeiten	F1/F3	F1	F1/F3	F3

Quelle: eigene Berechnungen

5. Resümee und Desiderate

Der bisherige Stand der Arbeiten im EviS-Projekt macht deutlich, dass man in der Breite nicht von einer mangelnden Evidenzbasierung in Schulen sprechen kann. Aber es zeigt sich, dass in dieser Hinsicht ein substanzielles bis hohes Entwicklungspotenzial besteht. Nicht zuletzt der Qualität der Aufbereitung und Kommunikation von Daten und Befunden sollte in der Forschung mehr Aufmerksamkeit geschenkt werden. Auch die deskriptiven Befunde zur Differenzierung der tatsächlichen Nutzung unterschiedlicher potenzieller Evidenzquellen belegen, dass die Nutzbarmachung intensiviert werden könnte und ‚viel Luft nach oben' gegeben ist,

zumal davon ausgegangen werden muss, dass die Selbstauskünfte aufgrund sozialer Erwünschtheit und Selbstrekrutierung evidenzbasierter Schulen ein tendenziell zu positives Bild zeichnen.

Instrumente eines outputorientierten Steuerungsparadigmas scheinen sowohl für Lehrkräfte als auch für Schulleitungen von vergleichsweise geringer Relevanz für die eigene Arbeit zu sein. Bedingt sein könnte dies z.B. durch eine als unzureichend wahrgenommene Passung der Rückmeldungen, da Schulinspektionsverfahren und Lernstandserhebungen universell in allen Schulen des Landes mit standardisierter Praxis durchgeführt werden. Im Fokus der Nutzung stehen Informationsquellen mit einem starken Bezug zur Einzelschule bzw. zur eigenen Unterrichtspraxis. Diese internen Informationsquellen wurden bisher in der Forschung eher wenig in den Blick genommen. Zu vermuten ist hier jedoch, dass die Akteure beispielsweise den Begriff des Schülerfeedbacks sehr unterschiedlich operationalisieren. Neben dem Grad der Formalisierung und Institutionalisierung bleibt deshalb auch zu untersuchen, in welchem Maße diese internen Informationsquellen gewissen Standards und einer Systematik folgen.

Literatur und Internetquellen

Ackeren, I. van/Zlatkin-Troitschanskaia, O./Binnewies, C./Clausen, M./Dormann, C./Preisendörfer, P./Rosenbusch, C./Schmidt, U. (2011): Evidenzbasierte Schulentwicklung. Ein Forschungsüberblick aus interdisziplinärer Perspektive. In: Die Deutsche Schule 103, H. 2, S. 170-184.

Bartunek, J. (2007): Academic-Practitioner Collaboration Need Not Require Joint or Relevant Research: Toward a Relational Scholarship of Integration. In: Academy of Management Journal 50, H. 6, S. 1323-1333.

Bellmann, J./Müller, T. (2011): Evidenzbasierte Pädagogik – ein Déjà-vu? Einleitende Bemerkungen zur Kritik eines Paradigmas. In: Bellmann, J./Müller, T. (Hrsg.): Wissen was wirkt. Kritik evidenzbasierter Pädagogik. Wiesbaden: VS, S. 9-32.

Berkemeyer, N./Bos, W. (2009): Professionalisierung im Spannungsfeld externer und interner Evaluation. In: Zlatkin-Troitschanskaia, O./Beck, K./Sembill, D./Nickolaus, R./Mulder, R. (Hrsg.): Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung. Weinheim/Basel: Beltz, S. 529-542.

Borgatti, S.P./Cross, R. (2003): A Relational View of Information Seeking and Learning in Social Networks. In: Management Science 49, H. 4, S. 432-445.

Briner, R. (2000): Evidence-based Human Resource Management. In: Trinder, L./Reynolds, S. (Hrsg.): Evidence-based Practice: A Critical Appraisal. Oxford: Blackwell Science, S. 184-211.

Briner, R. (2007): Is HRM Evidence-based and Does it Matter? Institute of Employment Studies Opinion Paper OP6. URL: www.employment-studies.co.uk/pdflibrary/op6.pdf; Zugriffsdatum: 13.01.2013.

Carey, S./Evans, R./Honda, M./Jay, E./Unger, C. (1989): An Experiment Is when You Try It and See if It Works. A Study of Junior High School Students' Understanding of the Construction of Scientific Knowledge. In: International Journal of Science Education 11, S. 514-529.

Cummings, T. (2007): Quest for an Engaged Academy. In: Academy of Management Review, H. 32, S. 355-350.

Demski, D./Rosenbusch, C./Ackeren, I. van/Clausen, M./Schmidt, U. (2012): Steuerung von Schule durch evidenzbasierte Einsicht? Konzeption und erste Befunde des Forschungsverbundes EviS. In: Hornberg, S./Parreira do Amaral, M. (Hrsg.): Deregulierung im Bildungswesen. Waxmann u.a.: Münster, S. 131-150.

Dormann, C./Dormann, T. (2010): Evidenzbasiertes Human Resource Management: Schlüssel zum Erfolg zukunftsgerichteter Personalarbeit. In: Zeitschrift für Personalführung 6, S. 74-81.

Dormann, C./Kleinschmidt, H./Urbich, A./Mohr, N./Stumm, S. (2009): Evidenzbasiertes organisationales Handeln. Personaltage Berlin, 18.-19.06.2009, Berlin.

Doyle McCarthy, E. (1996): Knowledge as Culture. The New Sociology of Knowledge. New York, NY/London: Routledge.

Eröss, G. (2009): Ignoranz, die Wissen erzeugt. „Spezifisches Nichtwissen" und Wissensersatzmechanismen im Spiegel des ungarischen Gesundheits- und Schulwesens. In: Soziale Welt 60, S. 371-387.

Greenhalgh, T./Robert, G./MacFarlane, F./Bate, P./Kyriakidou, O. (2004): Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations. In: Milbank Quarterly 82, H. 4, S. 581-629.

Gruber, H. (2006): Erfahrung als Grundlage von Handlungskompetenz. In: Bildung und Erziehung 59, H. 2, S. 193-197.

Günther, J./Grygier, P./Kircher, E./Sodian, B./Thoermer, C. (2004): Studien zum Wissenschaftsverständnis von Grundschullehrkräften. In: Doll, J./Prenzel, M. (Hrsg.): Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsentwicklung. Münster u.a.: Waxmann, S. 93-113.

Guest, D. (2007): Don't Shoot the Messenger: A Wake-up Call for Academics. In: Academy of Management Journal 50, H. 5, S. 1020-1026.

Hennig, M./Brandes, U./Pfeffer, J./Mergel, I. (2012): Studying Social Networks: A Guide to Empirical Research. Frankfurt a.M.: Campus.

Hofer, B.K. (2001): Personal Epistemology Research: Implications for Learning and Teaching. In: Journal of Educational Psychology Review 13, H. 4, S. 353-383.

Hofer, B.K./Pintrich, P.B. (1997): The Development of Epistemological Theories: Beliefs About Knowledge and Knowing and Their Relation to Learning. In: Review of Educational Research 67, H. 1, S. 88-140.

Jansen, D. (2006): Einführung in die Netzwerkanalyse: Grundlagen, Methoden, Forschungsbeispiele. Wiesbaden: VS.

Jornitz, S. (2008): Was bedeutet eigentlich „evidenzbasierte Bildungsforschung"? Über den Versuch, Wissenschaft für Praxis verfügbar zu machen, am Beispiel der Review-Erstellung. In: Die Deutsche Schule 99, H. 2, S. 206-216.

King, P.M./Kitchener, K.S. (1994): Developing Reflective Judgment: Understanding and Promoting Intellectual Growth and Critical Thinking in Adolescents and Adults. San Francisco, CA: Jossey Bass.

Kuhn, D. (1989): Children and Adults as Intuitive Scientists. In: Psychological Review 96, H. 4, S. 674-689.

Kulin, S./Frank, K./Fickermann, D./Schwippert, K. (2012): Egozentrierte Netzwerkanalysen: Resümee und Perspektiven für Forschungsansätze im Bildungsbereich. In: Kulin, S./Frank, K./Fickermann, D./Schwippert, K. (Hrsg.): Soziale Netzwerkanalyse. Münster u.a.: Waxmann. S. 293-295.

Latham, G. (2007): A Speculative Perspective on the Transfer of Behavioral Science Findings to the Workplace: The Times They Are A-Changin. In: Academy of Management Journal 50, H. 5, S. 1027-1032.

Maasen, S. ([2]2009): Wissenssoziologie. Komplett überarb. Auflage. Bielefeld: Transcript.

Maier, U./Kuper, H. (2012): Vergleichsarbeiten als Instrumente der Qualitätsentwicklung an Schulen – Überblick zum Forschungsstand. In: Die Deutsche Schule 104, H. 1, S. 88-99.

Peek, R. (2006): Dateninduzierte Schulentwicklung. In: Buchen, H./Rolff, H.-G.: Professionswissen Schulleitung. Weinheim: Beltz, S. 1343-1366.

Pepin, B. (1999): Epistemologies, Beliefs and Conceptions of Mathematics Teaching and Learning: The Story and What is Manifested in Mathematics Teachers' Work in England, France and Germany. In: TNTEE Publications 2, H. 1, S. 127-146.

Pfeffer, J./Sutton, R. (2006): Hard Facts, Dangerous Half-Truths, and Total Nonsense. Boston, MA: Harvard Business School Press.

Posch, P. (2009): Zur schulpraktischen Nutzung von Daten: Konzepte, Strategien, Erfahrungen. In: Die Deutsche Schule 101, H. 2, S. 119-135.

Rehrl, M./Gruber, H. (2007): Netzwerkanalysen in der Pädagogik: Ein Überblick über Methode und Anwendung. In: Zeitschrift für Pädagogik 53, H. 2, S. 243-265.

Rousseau, D. (2006): Is There Such Thing as "Evidence-based Management"? In: Academy of Management Review 31, H. 2, S. 256-269.

Rynes, S. (2007): Tackling the 'Great Divide' Between Research Production and Dissemination in Human Resource Management. In: Academy of Management Journal 50, S. 985-986.

Stumm, S./Dormann, C./Ohlemann, S. (2010): Evidenzbasiertes Management als Schlüssel für den Transfer zwischen Wissenschaft und unternehmerischer Praxis. 47. Kongress der Deutschen Gesellschaft für Psychologie (DGPs), 26.-30.09.2010, Bremen.

Ullmann, M./Stepancik, E. (2009): Das österreichische e-Learning Netzwerk eLSA. In: Berkemeyer, N./Kuper, H./Manitius, V./Müthing, K. (Hrsg.): Schulische Vernetzung. Eine Übersicht zu aktuellen Netzwerkprojekten. Münster u.a.: Waxmann, S. 93-106.

Wasserman, S./Faust, K. (2009): Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press.

Zlatkin-Troitschanskaia, O./Buske, R. (2009): Schulreform und Innovation in der Einzelschule – theoretische und empirische Annäherungen. In: Münk, D./Weiß, R. (Hrsg.): Qualität in der beruflichen Bildung. Forschungsergebnisse und Desiderata. Bielefeld: Bertelsmann, S. 179-190.

Zlatkin-Troitschanskaia, O./Kadach, A. (2008): Erfolgreiche Implementierung politischer Reformstrategien im öffentlichen (Berufs-)Schulwesen im Land Rheinland-Pfalz. In: Münk, D./Deißinger, T./Breuer, K. (Hrsg.): Berufs- und Wirtschaftspädagogik – Probleme und Perspektiven aus nationaler und internationaler Sicht. (Schriftenreihe der Sektion der Berufs- und Wirtschaftspädagogik der DGfE.) Opladen: Leske + Budrich, S. 101-111.

Isabell van Ackeren, Prof. Dr. phil., geb. 1974, Professorin für Bildungssystem- und Schulentwicklungsforschung, Universität Duisburg-Essen.
E-Mail: isabell.van-ackeren@uni-due.de

Marten Clausen, Prof. Dr. phil., geb. 1968, Professor für Unterrichtsforschung, Universität Duisburg-Essen.
E-Mail: marten.clausen@uni-due.de

Denise Demski, Dipl.-Soz.-Wiss, geb. 1980, wissenschaftliche Mitarbeiterin, AG Bildungsforschung, Universität Duisburg-Essen.
E-Mail: denise.demski@uni-due.de

Anschrift: Universität Duisburg-Essen, Fakultät für Bildungswissenschaften, Campus Essen/Weststadttürme, 45117 Essen

Carmen Binnewies, Prof. Dr. rer. nat., geb. 1979, Professorin für Arbeitspsychologie, Westfälische Wilhelms-Universität Münster.
E-Mail: carmen.binnewies@uni-muenster.de

Anna Rosa Koch, geb. 1985, wissenschaftliche Mitarbeiterin in der Arbeitseinheit Arbeitspsychologie, Westfälische Wilhelms-Universität Münster.
E-Mail: anna.r.koch@uni-muenster.de

Anschrift: Westfälische Wilhelms-Universität Münster, Institut für Psychologie, Fliednerstr. 21, 48149 Münster

Christian Dormann, Prof. Dr. phil. nat., geb. 1966, Professor für Arbeits-, Organisations- und Wirtschaftspsychologie, Ruhr-Universität Bochum.
E-Mail: christian.dormann@ruhr-uni-bochum.de

Anschrift: Ruhr-Universität Bochum, Fakultät für Psychologie, Universitätsstraße 150, 44780 Bochum

Bastian Laier, M.A., geb. 1981, wissenschaftlicher Mitarbeiter im Arbeitsbereich Organisation von Arbeit und Betrieb, Johannes Gutenberg-Universität Mainz.
E-Mail: laier@uni-mainz.de

Peter Preisendörfer, Prof. Dr. rer. pol., geb. 1953, Professor für Soziologie, Johannes Gutenberg-Universität Mainz.
E-Mail: preisendoerfer@uni-mainz.de

Anschrift: Johannes Gutenberg-Universität Mainz, Institut für Soziologie, Arbeitsbereich Organisation von Arbeit und Betrieb, Jakob-Welder-Weg 12, 55128 Mainz

Daja Preuße, Dipl.-Hdl., geb. 1979, wissenschaftliche Mitarbeiterin am Lehrstuhl Wirtschaftspädagogik der Johannes Gutenberg-Universität Mainz.
E-Mail: preusse@uni-mainz.de

Martin Stump, Dipl.-Hdl., geb. 1983, wissenschaftlicher Mitarbeiter am Lehrstuhl Wirtschaftspädagogik, Johannes Gutenberg-Universität Mainz.
E-Mail: stump@uni-mainz.de

Olga Zlatkin-Troitschanskaia, Prof. Dr. phil., geb. 1976, Professorin für Wirtschaftspädagogik, Johannes Gutenberg-Universität Mainz.
E-Mail: troitschanskaia@uni-mainz.de

Anschrift: Johannes Gutenberg-Universität Mainz, Fakultät für Wirtschafts- und Rechtswissenschaften, Jakob-Welder-Weg 9, 55128 Mainz

Christoph Rosenbusch, M.A. Soziologie und Publizistik, geb. 1979, wissenschaftlicher Mitarbeiter am Zentrum für Qualitätssicherung und -entwicklung der Johannes Gutenberg-Universität Mainz.
E-Mail: christoph.rosenbusch@zq.uni-mainz.de

Uwe Schmidt, Dr. phil., geb. 1960, Leiter des Zentrums für Qualitätssicherung und -entwicklung, Johannes Gutenberg-Universität Mainz.
E-Mail: uwe.schmidt@zq.uni-mainz.de

Anschrift: Johannes Gutenberg-Universität Mainz, Zentrum für Qualitätssicherung und -entwicklung (ZQ), Gresemundweg 4, 55128 Mainz

VERGLEICHSARBEITEN / LERNSTANDSERHEBUNGEN

DDS – Die Deutsche Schule
Beiheft 12, S. 74-96
© 2013 Waxmann

Uwe Maier/Carolin Ramsteck/Annette Frühwacht

Lehr-lerntheoretische Argumentationsmuster bei der Interpretation und Nutzung von Vergleichsarbeitsrückmeldungen durch Gymnasiallehrkräfte

Zusammenfassung
Vergleichsarbeiten sollen prüfen, welche zentralen Kompetenzen Schülerinnen und Schüler erworben haben. Lehrkräfte sollen daraus Schlussfolgerungen für die Unterstützung des weiteren Kompetenzerwerbs ziehen. In einer qualitativen Interviewstudie an Gymnasien (10 Schulen; 53 Deutsch-, Mathematik-, Englisch- und Französischlehrkräfte) in den Bundesländern Baden-Württemberg und Thüringen wurde der Frage nachgegangen, was im Anschluss an die zentralen Tests im Unterricht geändert wurde und wie dies begründet wird. Die lehr-lerntheoretischen Argumentationsmuster variieren sehr stark im Hinblick auf Argumentationstiefe, Anbindung an fachdidaktische Prinzipien und inhaltliche Fokussierung. Lehrerfortbildungen zur VERA-Nutzung könnten hier ansetzen und verstärkt den Zusammenhang zwischen zentralen Tests, Kompetenzmodellen und aktuellen fachdidaktischen Ansätzen aufzeigen.
Schlüsselwörter: Vergleichsarbeit, Lernstandserhebung, Feedback, Rezeptionsforschung, Kompetenzmessung, Unterrichtsentwicklung, Schulentwicklung

Instrumental Feedback Data Usage by German Grammar School Teachers

Summary
Mandatory tests in Germany aim at assessing core student competencies, and teachers are required to draw consequences for instructional improvement. This paper investigates what types of instrumental feedback data usage teachers report and how they connect them to their knowledge about teaching and learning. This qualitative interview

study sampled 53 teachers (German language, math, English, French) in 10 grammar schools in Baden-Württemberg and Thuringia. We found a broad variety of data usage and arguments on how to connect test data, instructional improvement and learning theory. The arguments highly differed in terms of elaboration, quality and content. Professional development programs might help teachers to make better sense of test data by drawing on competency models and highly effective instructional approaches.
Keywords: mandatory testing, feedback data usage, washback effect, instructional improvement, school improvement

1. Einleitung

Leistungstests gehören seit jeher zum Schulsystem moderner Industriegesellschaften. Während historisch gesehen zunächst die individuelle Zertifizierung von Schülerinnen und Schülern im Vordergrund stand, kann man seit einigen Jahrzehnten eine internationale Entwicklung beobachten, die zentrales Testen auf Einzelschul- und Systemebene als Werkzeug der Qualitätssicherung und -entwicklung versteht. Scheerens, Glas und Thomas (2003) unterscheiden diesbezüglich verschiedene Formen und Funktionen von Leistungstests im schulischen Mehrebenensystem:

a) internationale Leistungsvergleichsstudien mit dem Ziel des Systemmonitorings;
b) nationale Leistungsstudien mit dem Ziel, Mindeststandards in den Kernfächern zu prüfen und zu sichern;
c) zentrale, flächendeckende Leistungstests mit dem Ziel der öffentlich nachvollziehbaren Rechenschaftslegung und Qualitätsentwicklung auf Einzelschulebene;
d) schülerbezogene Leistungstests mit dem Ziel der individuellen Bewertung einerseits und der formativen Unterrichtsevaluation andererseits.

Die in Deutschland eingeführten Vergleichsarbeiten (VERA) sind Teil der KMK-Gesamtstrategie zur Sicherung und Entwicklung der Qualität im staatlichen Schulsystem (a). Aggregierte VERA-Daten werden herangezogen, um auf Systemebene Mindeststandards in den Kernfächern zu prüfen (b). Die Rückmeldedaten der jährlich durchzuführenden Tests sollen allerdings vorwiegend auf Ebene der Einzelschule als Anstoß für Qualitätsentwicklungsmaßnahmen dienen (c). Im Vergleich zu angloamerikanischen Schulsystemen wurde in Deutschland eine Form der testbasierten Schulreform eingeführt, die nicht auf eine Stärkung der Marktsteuerung setzt (z.B. durch Stärkung des elterlichen Schulwahlrechts), sondern externe Leistungsmessungen auf Schulebene als Instrument der internen Evaluation betrachtet (vgl. z.B. Baumert 2001; Altrichter/Maag Merki 2010). Zudem erwartet man von den Vergleichsarbeiten eine formative Unterrichtsevaluation (d), d.h. aufgrund der Rückmeldungen sollen Lehrkräfte zu einer fachdidaktischen Reflexion ihres bisherigen Unterrichts angeregt werden (vgl. KMK 2010, 2012a). Die Entwicklung und

Implementation von Vergleichsarbeiten stand damit von Anfang an unter dem Druck, den Anforderungen unterschiedlicher Akteure zu genügen.

1.1 Stand der empirischen Forschung zu Vergleichsarbeiten

Es gibt eine Reihe von deutschsprachigen Forschungsprojekten, die aufzeigen, dass zentrale Tests von der Lehrerschaft durchaus akzeptiert werden und Durchführung sowie Auswertung in der Regel kaum Probleme bereiten (vgl. z.B. Nachtigall/Jantowski 2007; Groß Ophoff/Hosenfeld/Koch 2007; Maier 2008). In Online-Befragungen zur Akzeptanz, Rezeption und Nutzung der Lernstandserhebungen in NRW gab eine große Mehrheit der Lehrkräfte an, sich mit den Ergebnissen auseinanderzusetzen (vgl. Kühle 2010). Ca. 20 Prozent der befragten Lehrkräfte weigerten sich dagegen hartnäckig, die Rückmeldungen zur Kenntnis zu nehmen. Ebenso äußerte in dieser Befragung ein nicht unerheblicher Teil der Lehrkräfte, vor allem in sprachlichen Fächern, Schwierigkeiten beim Verständnis der Testdaten.

Einige Studien berichten, dass der soziale Vergleich (Klasse im Landesmittelwert) und die Einschätzung einzelner Schülerinnen und Schüler bzw. Vergleiche mit der eigenen Notengebung im Vordergrund der Interpretation stehen (vgl. z.B. Nachtigall/Jantowski 2007; Schneewind/Kuper 2009; Diemer/Kuper 2011). Nur in einem geringeren Maße wird dagegen die Bedeutung der Tests für die Reflexion des eigenen Unterrichts gesehen (vgl. z.B. Nachtigall/Jantowski 2007; Sill/Sikora 2007; Maier 2008; Kühle 2010). Vergleichsarbeiten beeinflussen den weiteren Unterricht der Lehrkräfte vor allem über die Aufgabenformate. Testaufgabensammlungen werden als Pool für Übungsstunden genutzt.

Die Bedeutung der Ergebnisrückmeldungen für die unterrichtspraktische Arbeit wird von Schulleitungen höher eingeschätzt als von Lehrkräften (vgl. Bonsen/Büchter/Peek 2006; Maier 2009; Kühle 2010). Zudem gibt es große fachspezifische und interindividuelle Unterschiede im Rezeptions- und Nutzungsverhalten (vgl. z.B. Kuper/Hartung 2007; Diemer/Kuper 2011; Maier 2009; Kühle 2010; Schulze 2012). Lehrkräfte aus den sprachlichen Fächern schätzen den Nutzen von VERA für den eigenen Unterricht am geringsten ein. Eine sehr wichtige Bedingung für die Nutzungsbereitschaft ist die Akzeptanz der Tests, nicht jedoch das erzielte Testergebnis der eigenen Klasse und auch nicht Verständnisschwierigkeiten bei der Interpretation der Testergebnisse. Eine schulinterne Diskussion über die Leistungsdaten und Konsequenzen in kollegialen Gremien scheint eher die Ausnahme zu sein. Vor allem in qualitativen Studien wird deutlich, dass die Nutzung von Vergleichsarbeiten von der Fähigkeit des Kollegiums zur Kooperation abhängt. Die kollegiale Kooperation in Fachkonferenzen spielt bei Positivbeispielen eine große Rolle (vgl. z.B. Maier u.a. 2012; Asbrand/Heller/Zeitler 2012). Kühle (2010) fand zudem heraus, dass sich Fachkonferenzen eher an den kriterialen Aspekten der Testrückmeldungen (z.B. Kompetenzstufen) orientieren.

Auch wenn es große Unterschiede hinsichtlich Funktion und Ausgestaltung zentraler Tests zwischen Deutschland und anderen Ländern gibt, ist doch zu fragen, welche Aspekte der internationalen Forschung zu *test based accountability* bzw. *high-stakes testing* für Deutschland relevant sind. Eine Zusammenfassung des internationalen Forschungsstandes sowie eine Diskussion der Implikationen für die Entwicklungen im deutschsprachigen Raum wurden beispielsweise von Maag Merki (2010), Maier (2010a, 2010b) oder Maier und Kuper (2012) vorgelegt. Recht unbestritten ist der Umstand, dass *high-stakes testing* in den USA und anderen angloamerikanischen Ländern aufgrund der enormen Bedeutung für die Existenz von Schulen zu einer Fokussierung auf Testformate, Testinhalte und getestete Fächer führt. Die Datennutzung für die Schul- und Unterrichtsentwicklung wird gefordert, stößt aber auf ähnliche Hindernisse wie in Deutschland. Zentrale Leistungstests stehen dabei ebenfalls im Spannungsfeld zwischen bürokratischer Rechenschaftslegung und Professionalität von Lehrkräften (vgl. O'Day 2002). Über die Effekte eines hohen Rechenschaftsdrucks auf die Leistungsverbesserung wird in den USA heftig debattiert. Die empirische Befundlage hierzu ist allerdings sehr widersprüchlich (vgl. Maier 2010b).

Näher am deutschen Kontext sind Forschungsbefunde aus den Niederlanden. Dort wird von den Schulen gefordert, dass sie externe Testdaten für die interne Evaluation und Qualitätssicherung nutzen. Schildkamp und Visscher (2010) berichten beispielsweise über Schulen, die in ihren schriftlich formulierten Schulentwicklungsplänen auf zentrale Testrückmeldedaten zurückgreifen. Allerdings zeigt sich auch hier eine große Varianz zwischen den Schulen sowohl hinsichtlich der Schulentwicklungsaktivitäten als auch hinsichtlich der Wahrnehmung des Nutzens von externen Leistungsdaten. In einer weiteren Studie aus den Niederlanden stellten Vanhoof u.a. (2011) fest, dass Schulleiterfortbildungen die Dateninterpretationskompetenz und das Interesse an den Feedback-Berichten zwar verbessern können, jedoch wenig an einer eher geringen, instrumentellen Nutzung der Rückmeldungen ändern.

Es gibt eine Reihe von Versuchen, die zahlreichen und oft nur schwer miteinander vergleichbaren, internationalen Befunde zur Nutzung externer Testdaten theoretisch zu modellieren. Ein sehr breit angelegtes Modell, das explizit auf die internationalen Studien zu dieser Thematik rekurriert, ist das *School-Performance-Feedback-System*-Modell (kurz: SPFS-Modell), das erstmals von Visscher und Coe (2003) publiziert und von anderen Autoren genutzt bzw. weiterentwickelt wurde (vgl. z.B. Schildkamp/Visscher 2009; Verhaeghe u.a. 2010). Das Modell basiert teilweise auf empirischen Studien zur Effektivität externer Testsysteme, aber auch auf grundsätzlichen Überlegungen zu Effekten bzw. Kontextfaktoren dieser Systeme. Es ist weniger ein exaktes Modell für die Modellierung von Zusammenhängen, sondern eher eine Strukturierungshilfe für relevante Einflussfaktoren. Die deutschsprachigen Befunde lassen sich sehr gut in diese Kontextfaktorenmodelle einordnen. Dabei wird deutlich, dass eine instrumentelle Nutzung von externen Testdaten sowohl von der Qualität

der Tests und Testrückmeldungen und von den Bedingungen der Implementation als auch von schulinternen Faktoren abhängt.

1.2 Lehr-lerntheoretische Rahmung der Nutzung von Vergleichsarbeiten

Das IQB sieht in „der Begründung und Planung pädagogischer Interventionen und Fördermaßnahmen" eine Facette des pädagogischen Potenzials von VERA (vgl. URL: http://www.iqb.hu-berlin.de/vera; Zugriffsdatum: 12.05.2013). Ein Nachteil vieler Studien und Modelle ist allerdings eine ungenaue oder zu breite Definition der Nutzung von VERA für pädagogische Interventionen. Dies hängt mit den heterogenen und oft widersprüchlichen Zielsetzungen externer Tests zusammen. Im oben erwähnten SPFS-Modell wird die idealtypische Nutzung als „instrumentell" bezeichnet, d.h. Evaluationsdaten führen direkt zu einer Handlung oder stehen zumindest in einem direkten Zusammenhang mit einer konkreten Handlung eines Akteurs (vgl. Schildkamp/Kuiper 2010). Die instrumentelle Nutzung wird von konzeptueller Nutzung abgegrenzt, bei der Evaluationsdaten allenfalls zu einem erweiterten Verständnis eines Sachverhalts führen. Als nicht-intendiert wird dagegen eine strategische Nutzung bezeichnet, z.B. wenn Testdaten fälschlicherweise herangezogen werden, um sachfremde Ziele zu erreichen.

Bei der Komplexität von Handlungen und Entscheidungen im Schul- und Unterrichtsalltag ist eine Vielzahl von Formen der instrumentellen Testdatennutzung denkbar. In einigen Studien wurde gefragt, wie viele Maßnahmen im Anschluss an VERA beschlossen und umgesetzt wurden. Selbst wenn ungefähr bekannt ist, um welche Maßnahmen es sich handelt, sagen diese Daten noch nichts über Nachhaltigkeit, Qualität und Effekt der Maßnahme auf die Lernleistung der Schülerinnen und Schüler aus. Wo müsste man aber ansetzen, um beurteilen zu können, welche Formen der instrumentellen Nutzung prinzipiell möglich und auch erwünscht sind? Einen idealtypischen Ansatzpunkt für diese Frage lieferte Pellegrino (2003) mit dem *assessment triangle*. Für die Analyse der didaktischen Nutzbarkeit von Leistungsmessungen müssen drei Komponenten betrachtet werden:

1) Es muss ein kognitionspsychologisches Modell über Repräsentation und Entwicklung von Wissen in einer spezifischen Fachdomäne vorliegen.
2) Ein Test eröffnet aufgrund von Aufgabenstellungen oder Beobachtungen in Leistungssituationen Einblicke in die aktuelle, kognitive Repräsentation dieses Wissens.
3) Dem Testanwender muss es möglich sein, die einzelnen Beobachtungen bzw. die aggregierten Testwerte vor dem Hintergrund des kognitiven Modells zu interpretieren, d.h. er muss die diagnostischen Informationen mit seinen Vorstellungen über den domänenspezifischen Wissenserwerb verknüpfen und dann mögliche Maßnahmen begründen können.

Dieses *assessment triangle* ist Grundlage für den von Pellegrino, Chudowsky und Glaser (2001) herausgegebenen Bericht „Knowing What Students Know", der sich aus einer kognitionspsychologischen Perspektive sehr gründlich und kritisch mit den Auswüchsen des zentralen Testens in den USA auseinandersetzt. Für Pellegrino (2003, S. 49) lassen sich viele problematische Entwicklungen im Bereich der schulischen Leistungsdiagnostik vor allem auf mangelnde oder falsche kognitive Modelle zurückführen: „Much of what we've been doing in assessment has been based on impoverished models of cognition, which has led us to highly limited modes of observation that can yield only extremely limited interpretations of what students know." Es genügt nach Ansicht von Pellegrino (2003) auch nicht, Defizite einer Komponente im *assessment triangle* mit besonders ausgefeilter Methodik bei anderen Komponenten auszugleichen.

Wie lässt sich nun die Konzeption und Nutzung von Vergleichsarbeiten vor dem Hintergrund des *assessment triangle* analysieren? Hinweise hierfür erhält man in den didaktischen Handreichungen zu VERA 3 und 8. Das Kompetenzstufenmodell Orthographie (vgl. KMK 2012b) geht beispielsweise von den KMK-Standards zur Rechtschreibkompetenz und fachdidaktischen Erkenntnissen zur Rechtschreibentwicklung aus (kognitives Modell im *assessment triangle*) und begründet die Entwicklung der VERA-Testaufgaben im Bereich Orthografie (Testsituation im *assessment triangle*). Bei genauem Lesen erkennt man jedoch, dass aufgrund der Anforderungen an Testitems nicht alle fachdidaktisch relevanten Aspekte berücksichtigt werden können. Bei der Herleitung eines Kompetenzstufenmodells für die Orthografiekompetenz wird nicht zwischen verschiedenen Rechtschreibstrategien, orthografischen Regeln oder Fehlertypen differenziert, wie es in der deutschdidaktischen Forschung und Diagnostik sonst üblich ist. Das Kompetenzstufenmodell wird rein induktiv über die Lösung von Testaufgaben entwickelt. Dies hat zur Folge, dass auf einer Niveaustufe sehr viele Rechtschreibphänomene bzw. Fehlertypen zusammengefasst werden. D.h., die Vergleichsarbeiten basieren auf kognitiven Modellen, die sich zwar an fachdidaktischen Erkenntnissen orientieren, diese aber in ihrer Komplexität deutlich reduzieren, um einfache Vergleichsmöglichkeiten zu garantieren.

Wie muss man sich in einem weiteren Schritt die Verknüpfung von Vergleichsarbeit und Unterrichtsentwicklung vorstellen? Welche lehr-lerntheoretischen Argumentationsmuster werden von den Testentwicklern und -entwicklerinnen und den Fachdidaktikern und -didaktikerinnen angeboten? Die an der Entwicklung von Vergleichsarbeiten beteiligten Fachdidaktiker und Fachdidaktikerinnen gingen diesen Fragen in einer Reihe von Publikationen nach (vgl. z.B. Bremerich-Vos/Granzer/Köller 2008; Tesch/Leupold/Köller 2008). Die Antworten bestehen aus einem Konglomerat aus alten und neuen fachdidaktischen Prinzipien, aktuellen Erkenntnissen und Vorstellungen zur stärkeren Kompetenzorientierung von Unterricht. Mathematikdidaktiker und -didaktikerinnen verstehen Vergleichsarbeiten beispielsweise als Chance, dass Lehrkräfte einen „Kompetenzblick" auf die Schülerinnen und Schüler

entwickeln können (vgl. z.B. Büchter/Leuders 2005; Blum u.a. 2005; Lorenz 2005; Drüke-Noe/Keller/Blum 2008; Blum u.a. 2010). Man fordert die Nutzung kompetenzorientierter Aufgaben in Lern-, Leistungs- und Übungssituationen. Mathematiklehrkräfte sollten ihr Augenmerk eher auf den langfristigen Kompetenzaufbau legen (z.B. Kommunizieren, Modellieren oder Argumentieren) und verstärkt das diagnostische Potenzial von (Test-)Aufgaben nutzen, z.B. durch gezielte Fehleranalysen und fachdidaktische Aufgabenkommentierungen. Allerdings erfordert dies von den Lehrkräften eine zusätzliche Analyseleistung auf Itemebene und ein fachdidaktisches Lehr-Lernmodell darüber, wie sich das domänenspezifische Wissen aufbaut. Aber auch traditionelle mathematikdidaktische Prinzipien sollen durch die Einführung von Bildungsstandards und Vergleichsarbeiten zur Geltung gebracht werden, z.B. die gezielte Variation von Aufgaben oder die Nutzung von Realitätsbezügen für mathematische Probleme (komplexere Sachaufgaben). Zudem wird auch „inhaltlich" argumentiert. Die in den Bildungsstandards und VERA prominent vertretene Leitidee „Daten und Zufall" soll verstärkt im Mathematikunterricht Beachtung finden.

Auch in den sprachlichen Fächern werden eine Reihe von Bezügen zwischen Bildungsstandards, Vergleichsarbeiten, fachdidaktischer Forschung und Unterrichtsentwicklung hergestellt. Hu und Leupold (2008) beispielsweise interpretieren die Bildungsstandards und Vergleichsarbeiten in Französisch vor dem Hintergrund des Europäischen Referenzrahmens für Sprachen und leiten daraus eine Reihe von Forderungen für die Unterrichtsentwicklung ab:

- Lesen sollte als kreativer Sinnbildungsprozess verstanden werden. Das Abfragen von Informationen oder Zusammenhängen aus dem Text, wie es überwiegend in VERA praktiziert wird, ist allerdings nur ein Aspekt dieses Prozesses. Darüber hinaus kommt es auf die Einbettung des Textes, den Entstehungszusammenhang und die Anschlusskommunikation an.
- In den Fremdsprachvergleichsarbeiten spielen Hörverstehensaufgaben eine zentrale Rolle. Hörverstehen gilt als eine hoch komplexe Kompetenz. Schülerinnen und Schüler müssen im Bruchteil von Sekunden relevante Informationen verarbeiten und einordnen können. Ein Problem des Fremdsprachunterrichts ist vor allem die Variantenarmut an Hörbeispielen (ganz im Gegensatz zum natürlichen Spracherwerb). Die Hörbeispiele der Vergleichsarbeiten und entsprechende Zusatzmaterialien können hier den Unterricht ergänzen.
- Schreiben muss als kreative, kommunikative Aktivität verstanden werden. Schreibanlässe müssen motivierend und alltagsbezogen sein. Die Rechtschreibung sollte erst am Ende eines Schreibprozesses eingefordert werden. Bisher wurde der Französischunterricht von einer zu rigiden Fehlerkultur beim Schreibprozess dominiert.

In der Rezeptions- und Nutzungsforschung wird an verschiedenen Stellen diese Passung zwischen Test und lehr-lerntheoretischer Expertise der Lehrkräfte bzw. Unterrichtsentwicklung diskutiert. Für Peek und Dobbelstein (2006) beispielsweise

ist die Kompatibilität der Lernstandserhebungen in NRW mit den fachdidaktischen Vorstellungen der Lehrkräfte entscheidend. Halbheer und Reusser (2008) weisen zudem auf die Notwendigkeit der Einordnung von Vergleichsarbeiten in die bestehende Diagnosepraxis von Lehrkräften hin. Auch Fend (2011) kritisiert, dass die Ebene des Handelns (Lehren und Lernen) in der Diskussion über neue Steuerung weitgehend ausgeblendet wird. Die Befunde von Kühle (2010) deuten zudem darauf hin, dass zwischen wahrgenommener Kompatibilität und tatsächlicher Kompatibilität zwischen Tests und Nutzung differenziert werden sollte.

Ein Problem besteht natürlich auch darin, dass mit den Vergleichsarbeiten nur ein bestimmter Ausschnitt relevanter Kompetenzen geprüft werden kann und damit der Impuls für die Unterrichtsentwicklung abgeschwächt wird (z.B. Schreibkompetenz wird weitgehend ausgeblendet). Die Kritik an der mangelhaften instrumentellen bzw. fachdidaktischen Nutzung von Vergleichsarbeitsrückmeldungen wurde zudem von der Bildungspolitik aufgegriffen. Die KMK (2010) forderte in ihrer Konzeption zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung, dass kompetenzorientierte Standards und Tests beispielsweise die Frage der Anwendbarkeit von Wissen in alltagsnahen Situationen in den Vordergrund der Unterrichtsentwicklung rücken sollen. Für die Weiterentwicklung der Vergleichsarbeiten schlägt die KMK (2012a) zudem vor, dass Landesinstitute verstärkt geeignete Fortbildungsangebote machen sollen.

2. Fragestellung

Die Begleitliteratur zu Vergleichsarbeiten weist auf eine Reihe von fachdidaktischen Optionen im Anschluss an VERA hin. Die instrumentelle Nutzung auf Unterrichtsebene (Ableitung von konkreten Maßnahmen der Optimierung von Lehr-Lernprozessen) wird in vielen Studien allerdings als kritisch beschrieben. Ein Problem ist zudem, dass viele Befunde auf quantitativen Lehrerbefragungen basieren. Damit können zwar durchaus komplexe Zusammenhänge zwischen Akzeptanz, Rezeption und Nutzungsverhalten aufgezeigt werden. Die Qualität der instrumentellen Nutzung und die domänenspezifische Reflexionstiefe der Lehrkräfte bei der Begründung von Maßnahmen können allerdings nicht hinreichend erfasst werden. Quantitative Fragebogenitems sind in der Regel nicht fachdidaktisch spezifiziert. Aber auch qualitative Studien, die Nutzungsvarianten erfassen und auswerten, hatten bisher einen eher breiten Fokus auf alle möglichen, von den Lehrkräften genannten Verwendungsmöglichkeiten der Testergebnisse (vgl. z.B. Kuper/Hartung 2007; Maier 2009; Diemer/Kuper 2011; Asbrand/Heller/Zeitler 2012).

Eine qualitative, dichte Beschreibung domänenspezifischer Argumentationszusammenhänge bei der Rezeption und Nutzung von zentralen Testdaten bleibt ein

Forschungsdesiderat. Diese Studie möchte deshalb genauer ergründen, welche lehr-lerntheoretischen Argumentationsmuster Lehrkräfte für die Reflexion über Testdaten und Schlussfolgerungen nutzen:

1) Orientiert sich die Reflexion der testinduzierten Maßnahmen an den Kompetenzmodellen der Bildungsstandards bzw. den jeweiligen VERA-Handreichungen?
2) Auf welche fachdidaktischen Erkenntnisse über Lehr- und Lernmethoden greifen Lehrerinnen und Lehrer zurück, um testinduzierte Maßnahmen zu begründen?
3) Welche Rahmenbedingungen ermöglichen bzw. verhindern eine auf fachdidaktischen oder lehr-lerntheoretischen Erkenntnissen basierende Reflexion testinduzierter Maßnahmen?

Antworten auf diese Fragen könnten in Zukunft verstärkt dazu beitragen, sowohl die Tests selbst als auch didaktische Handreichungen bzw. Lehrerfortbildungen besser an die Expertise von Lehrkräften anzupassen.

3. Methodisches Vorgehen

Um lehr-lerntheoretische Argumentationsmuster bei der Rezeption und Nutzung von Vergleichsarbeitsrückmeldungen beschreiben, analysieren und ordnen zu können, wurden qualitative Lehrerbefragungen inhaltsanalytisch ausgewertet. Die Datenerhebung fand im Rahmen des BMBF-Projekts „Testbasierte Schulreformen in der Mehrebenenstruktur des Schulsystems“ statt. In vier Bundesländern mit unterschiedlichen Regelungskontexten wurden Akteure auf vier Ebenen des Schulsystems (Lehrkräfte, Fachleiter und Fachleiterinnen, Schulleiter und Schulleiterinnen, Schulaufsicht) zum Umgang mit Vergleichsarbeiten befragt. Für diesen Beitrag wurden die Interviews mit Lehrkräften in den Bundesländern Thüringen und Baden-Württemberg ausgewertet.

Die Lehrerinterviews wurden an ausgewählten Gymnasien geführt. Die Auswahl wurde im Rahmen der Genehmigung zusammen mit der Schulaufsichtsbehörde getroffen. Leitendes Kriterium für die Stichprobengewinnung war die zu erwartende Informationsdichte (vgl. Merkens 1997). Die Schulaufsicht wurde gebeten, die Erhebung an fünf Gymnasien, an denen ein konstruktiver Umgang mit Vergleichsarbeiten und eine gewisse Auskunftsbereitschaft erwartet werden können, zu ermöglichen. Die Schulen sollten zudem mehrzügig sein, um eine größere Zahl von Fachlehrkräften in den relevanten Kernfächern befragen zu können. Diese Vorgehensweise führte einerseits zu einer hohen Teilnahmebereitschaft, andererseits zu einer Positivselektion der befragten Lehrkräfte. Ziel der Studie ist jedoch nicht die repräsentative, sondern die möglichst dichte Beschreibung lehr-lerntheoretischer Argumentationsmuster.

In beiden Bundesländern wurden an zehn Gymnasien insgesamt 115 Interviews geführt, davon 74 Interviews mit Lehrkräften, 14 mit Fachleitern und Fachleiterinnen, 19 mit Vertretern und Vertreterinnen der Schulleitung und acht auf Schulaufsichtsebene. Für diesen Beitrag wurden die Daten der 74 Lehrerinterviews herangezogen. Von den Interviews wurden 54 mit Lehrerinnen und 20 mit Lehrern durchgeführt. Die untersuchte Stichprobe auf Lehrerebene weist ein Durchschnittsalter von rund 44 Jahren auf und verfügt über durchschnittlich 18 Jahre Berufserfahrung. 27 der befragten Lehrkräfte unterrichten Deutsch (D), 17 Englisch (E), 14 Mathematik (M) und sechs das Fach Französisch (Fr). 21 Lehrkräfte standen sowohl am ersten als auch am zweiten Messzeitpunkt für Interviews zur Verfügung, d.h. insgesamt wurden 53 unterschiedliche Lehrkräfte befragt.

Es wurden problemzentrierte, leitfadengestützte Interviews (vgl. Hopf 2000) zu zwei Messzeitpunkten (1. Quartal Schuljahr 2010/11; 2. Quartal Schuljahr 2011/12) durchgeführt. In der Regel waren es Einzelinterviews. Nur in Ausnahmefällen wurden zwei Personen in einem Interview befragt. Die für diesen Beitrag ausgewerteten Interviews fokussierten auf die Durchführung, Interpretation und Nutzung der Kompetenztests 6 sowie VERA 8 (jeweils M, D, E) in Thüringen und der Diagnose- und Vergleichsarbeiten (DVA) in den Klassen 7 sowie 9 (jeweils M, D, E, Fr) in Baden-Württemberg. Ein Spezifikum ist, dass Baden-Württemberg nicht an VERA 8 teilnimmt. Die DVA sind jedoch konzeptionell mit VERA 8 vergleichbar. Ebenso wurden die Lehrkräfte zu den organisatorischen Rahmenbedingungen, die möglicherweise den Umgang mit Vergleichsarbeitsergebnissen beeinflussen, befragt. Darüber hinaus wurden Kontextinformationen zu den Schulen erhoben (z.B. Schulprofil, Sozialstruktur, Führungsstil der Schulleitung etc.).

Nach der transkriptorischen Datenaufbereitung wurden gemäß der strukturierenden Inhaltsanalyse nach Mayring (2010) für jede befragte Akteursebene und jeden Messzeitraum umfangreiche Kategoriensysteme erstellt. Allerdings machte nicht jede befragte Lehrkraft für die hier interessierende Fragestellung relevante Aussagen. Es gab eine Reihe von Lehrkräften, die sich aufgrund einer grundsätzlich ablehnenden Haltung weigerten, aus den Testrückmeldungen eine Schlussfolgerung für ihren eigenen Unterricht zu ziehen. Andere wiederum entzogen sich einer weiterführenden, fachdidaktischen Reflexion aufgrund zum Teil nachvollziehbarer Argumente, z.B.: Test prüfte Themen, die im Unterricht noch nicht „behandelt" wurden; Klasse hat gut abgeschnitten, und deshalb wurden die Ergebnisse nicht weiter analysiert. Insgesamt konnten die Codings von 42 unterschiedlichen Lehrkräften aus 60 Interviews (d.h. bei 18 Lehrkräften aus Interviews zum 1. und 2. Messzeitpunkt) in der Oberkategorie „Verwendung der Ergebnisse" für die hier vorliegende Fragestellung herangezogen werden.

4. Ergebnisse

Um die inhaltsanalytische Auswertung des Interviewmaterials besser nachvollziehen zu können, wird zunächst an einem Fallbeispiel aufgezeigt, welche Schlussfolgerungen eine Lehrerin aus den zentralen Tests gezogen hat und wie sie diese lehr-lerntheoretisch begründet. In einem zweiten Schritt werden dann die im Material gefundenen lehr-lerntheoretischen Argumentationsmuster zusammenfassend dargestellt.

4.1 Fallbeispiel

Das Fallbeispiel ist einem Interview mit einer Mathematiklehrerin aus Thüringen entnommen und bezieht sich auf den Kompetenztest 6 in Mathematik. Die Lehrkraft äußert sich im Laufe des Interviews immer wieder positiv über die Kompetenztests. Sie betont vor allem die evaluative Funktion:

> *L: „Ja, das ist also wirklich wichtig, dass man, naja, dass es ein Anlass ist in Diskussion zu kommen, in der Fachschaft über gewisse Dinge zu reden. Und jede Diskussion braucht einen Anlass und das wäre ein möglicher von vielen."* #00:00:38-6#

Allerdings sähen dies nicht alle Kollegen so. Die Fachschaft Mathematik sei sehr groß, und es werde kaum über Inhalte geredet, obwohl die befragte Lehrerin dies begrüßen würde. Ihre Vorstellung wäre, dass man nicht nur auf zentrale Vorgaben warte, sondern sich innerhalb der Fachschaft auf weitere, schulintern gestellte Tests einige. Ein Problem sei allerdings, dass der Fachschaftsvorsitzende nur in der Oberstufe arbeite und sich weniger um die Belange der Unterstufe kümmere.

Die Schulleitung verlangt eine Auswertung der Kompetenztests, was von der Lehrkraft allerdings als „deplatziert" empfunden wird. Auch die Vorstellung der Auswertung in der Gesamtlehrerkonferenz wird nicht als sinnvoll erachtet, weil die anderen Fachlehrkräfte wenig Interesse daran haben. In der Schule gibt es allerdings folgendes System: Auf Klassenebene wird ein Aushang erstellt, wer unter (rot), über (grün) oder im Landesdurchschnitt lag. Im Interview gab es keine Hinweise, ob dieses System die Nutzung positiv oder negativ beeinflusst.

Die Lehrerin wurde durch die Kompetenztests angeregt, ihre Aufgaben anschaulicher zu gestalten, sofern dies der Platz auf Arbeitsblättern und die Kopierkapazität zulassen:

> *L: „Ich gebe mir immer mehr Mühe, Mathe-Arbeiten auch von der Anschaulichkeit her abwechslungsreich / Also dass ich Textaufgaben, die mit Text ver-*

bunden sind / Aufgaben, die nackige Zahlen haben / Aufgaben, die mit Tabellen verbunden sind / Aufgaben, die Bildchen noch haben zur Erklärung oder sowas. Wobei ich sicherlich diese Breite dieser Kompetenztests ja gar nicht nutzen kann. Dafür hab' ich nicht genug Papier und auch nicht genug Kopierkapazität. Aber was ich innerhalb auf einer Seite so kriegen kann, versuche ich zu gestalten, ja." #00:13:21-4#

Die Interviewerin fragt nun genauer nach, ob sich diese Maßnahme auf das Lernverhalten der Schülerinnen und Schüler auswirke. Dies wird von der Lehrkraft bejaht. Die Bedeutung des Arbeitens mit Tabellen wird als wichtige Kompetenz bezeichnet. Im gleichen Atemzug wird allerdings wieder auf die Ressource Zeit angespielt:

I: „Haben Sie auch irgendwie gemerkt, dass die Schüler da positiver darauf reagiert haben?" #00:13:24-5#

L: „Auf Bildchen reagieren sie positiv, ja. Bildchen ist ganz beliebt. Mit Tabellen zu arbeiten, gut, da stelle ich fest / Ich muss viel, viel üben mit Tabellen zu arbeiten, weil / Ja, ich weiß auch nicht warum. Ich bin vielleicht eine der wenigen, die mit Tabellen arbeitet und ich hab' irgendwo kein Vorfeld. Und wenn ich dann also immer irgendeine Sortierung in Tabellen haben möchte, ja, bedarf es einer längeren Übungsphase. Ja manchmal gebe ich auf und manchmal mache ich weiter, aber meistens mache ich weiter. Also es kommt immer d'rauf an, wie man gehetzt wird durch die Mathematik, durch die Vorgaben des Lehrplans. Ob man sich gerade in dieser Phase gehetzt fühlt und weitermachen muss und nun doch nicht diese Kompetenz, die ich eigentlich wichtig finde, Einordnung in Tabellen, gerade momentan ausleben kann oder nicht." #00:14:19-6#

Die Interviewerin hakt noch einmal nach und erhält eine Art neurowissenschaftliche Begründung mit einer erneuten Betonung der begrenzten Zeitressourcen:

L: „Das Gehirn braucht mehrere Ebenen, in denen es bedient wird. Und wenn man mehr Zeit hätte in Mathematik, könnte man auch viel mehr Ebenen bedienen. Und das ist ja überhaupt das Dilemma. Aber wie gesagt, da jammern alle. Es kann nicht sein, aber das sag' ich Ihnen auch noch mal, dass wir immer noch fast den gleichen Lehrplan haben wie zu DDR-Zeiten, nur mit zwei Stunden weniger pro Woche. Und ich hetze mich durch die Gegend." #00:15:07-7#

L: „Die Schüler fühlen sich gehetzt und ich fühle mich gehetzt. Und ich finde es einfach traurig, dass das so ist. Man könnte naturwissenschaftlich bei den Schülern viel mehr erreichen, wenn ihnen Zeit gelassen wird. Das Gehirn

> *braucht Zeit, das Gehirn hat sich nicht geändert in zwanzig Jahren, braucht Zeit um etwas setzen zu lassen, um etwas zu verarbeiten, braucht genügend Übungsphasen, braucht verschiedene Möglichkeiten der Veranschaulichung. Das kann ich aber nur gestalten, wenn ich genügend Zeit habe.*" #00:16:20-2#

Aus diesen Interviewpassagen lässt sich folgende lehr-lerntheoretische Argumentationsstruktur der Lehrkraft rekonstruieren: Wenn mehr Wert auf unterschiedliche Darstellungsformen mathematischer Inhalte gelegt wird, dann ist die Behaltensleistung der Schülerinnen und Schüler höher. Die Vergleichsarbeiten scheinen zudem die Vorstellung zu untermauern, dass das Arbeiten mit Tabellen wichtiges Ziel des Mathematikunterrichts ist. Dieses Ziel erreicht man allerdings nur durch langfristig angelegte Übungen. Begründet werden beide subjektiven Theorien mit dem Verweis auf die Funktionsweise des Gehirns, das auf Veranschaulichungen angewiesen sei und sehr viel Zeit brauche. Die Lehrerin spricht ebenfalls sehr deutlich die Grenzen einer Umsetzung ihrer subjektiven Vorstellungen vom Wissenserwerb in Mathematik an: begrenzte Kopierkapazität, Zeitprobleme, überfrachteter Lehrplan. Aber auch die Diskussionen in der Fachkonferenz und die schulinterne Berichtslegung werden nicht als Unterstützung wahrgenommen.

4.2 Lehr-lerntheoretische Argumentationsmuster

Die aus den Einzelinterviews rekonstruierten lehr-lerntheoretischen Argumentationsmuster bei der Reflexion über die Nutzung von Vergleichsarbeiten wurden induktiv zu Kategorien zusammengefasst. Die Kategorien werden kurz umschrieben und mit paraphrasierten Aussagen von Lehrkräften beispielhaft erläutert. Aus Platzgründen können keine weiteren Interviewpassagen zitiert werden.

Mehr Zeit für Inhalte, bei denen Schülerinnen und Schüler Schwächen haben

Das mit Abstand häufigste Argumentationsmuster zur Begründung einer Schlussfolgerung aus den Testrückmeldungen lässt sich wie folgt zusammenfassen: Die Schülerinnen und Schüler benötigen mehr Unterrichtszeit für Inhalte, die in der Vergleichsarbeit „nicht gekonnt" wurden. Interessant ist, dass die Lehrkräfte ausschließlich entlang von Fachinhalten („Stoff") argumentieren. Sie identifizieren den fachlichen Inhalt, der dafür verantwortlich ist, dass die Lernenden bestimmte Aufgaben in den Tests lösen oder eben nicht lösen (z.B. Thema Adjektiv/Adverb; Rechtschreibung; Zeichensetzung; wörtliche Rede; Präpositionen; Vergangenheitsformen in Französisch). Daraufhin werden unterschiedlichste Überlegungen angestellt, wie diese Inhalte verstärkt geübt oder wiederholt werden können. Eine Lehrkraft nutzt die Testrückmeldungen beispielsweise sehr systematisch zur Planung der Stoffverteilung. Bereiche mit Defiziten werden im Unterricht zuerst angegangen (z.B. der Bereich

Textverstehen). Dadurch, dass sie mit diesen Bereichen anfängt, sieht sie über den Schuljahresverlauf, wie viel sie noch üben lassen muss. Zudem hat die Lehrkraft im Vorfeld der DVA Grammatik und Gedichte (dies konnte die Vorgängerklasse in der DVA nur unzureichend beantworten) wiederholt und die Art der Aufgabenstellung anhand von Übungsaufgaben verdeutlicht.

Mehr Tests und Notendruck

Ein weiteres Argumentationsmuster, das wir überwiegend bei Mathematiklehrkräften fanden, besteht darin, die aufgrund der Testrückmeldungen als defizitär erkannten Inhalte verstärkt zu wiederholen und in weiteren Klassenarbeiten oder Kurztests abzufragen. Eine Lehrkraft stellt z.B. einfache Aufgaben zu schwierigen Themenbereichen in jeder Klassenarbeit. Eine weitere Lehrkraft differenziert sehr genau zwischen relevanten und weniger relevanten Themen. Aus den DVA zum Befragungszeitpunkt zieht die Lehrkraft keine Konsequenzen, da die abgefragten Themen (Thaleskreis, Kongruenzsätze und Wahrscheinlichkeitsrechnung) keine wesentlichen Lerninhalte darstellen und für nicht so wichtig gehalten werden, um sie nachhaltig zu verankern. Aus den vorherigen DVA, die mit quadratischen Funktionen und quadratischen Gleichungen relevante Schwerpunkte abgefragt haben, hat die Lehrkraft die Konsequenz gezogen, dass die Schülerinnen und Schüler diese beiden Themen verstärkt üben müssen. Daher hat die Lehrkraft vier Tests zu diesen Themen angesetzt, die als eine Klassenarbeit gewertet wurden.

Kleinschrittiges Üben von grundlegendem Wissen und Grundfertigkeiten

Eine Reihe von Lehrkräften führt schwächere Leistungen in den zentralen Tests vor allem auf mangelndes Grundwissen bzw. nicht vorhandene Grundfertigkeiten zurück. In der Konsequenz achten diese Lehrkräfte verstärkt auf Übungsmöglichkeiten für diese Grundfertigkeiten. An erster Stelle steht dabei das genaue Lesen von Texten und Aufgabenstellungen. Eine Lehrkraft wiederholt beispielsweise intensiv das kleinschrittige Arbeiten mit Texten anhand der Fünf-Schritt-Lesemethode. Eine andere Lehrkraft löst sich vom „veralteten" Schulbuch und baut verstärkt DVA-Aufgaben in den Unterricht ein, um das genaue Lesen der Aufgabenstellungen zu üben. Eine Gymnasiallehrerin in Thüringen legt nach dem schlechten Abschneiden der Klasse mehr Wert auf Regelkenntnisse und gestaltet ihren Unterricht „kleinschrittiger". Am Ende einer „Stoffeinheit" sollen die Schülerinnen und Schüler fünf Regeln nennen können. Die Lehrerin lässt in den Heftern zudem einen Merkteil mit Regelkenntnissen anlegen, was ihrer Meinung nach zu guten Erfolgen führt. Eine Mathematiklehrkraft achtet verstärkt auf die Einhaltung von Rechenregeln (z.B. Punktrechnung vor Strichrechnung).

Eine Deutschlehrkraft übt mit ihren Schülerinnen und Schülern die Beantwortung von Fragen im „Telegrammstil". In ihren eigenen Klassenarbeiten legt sie Wert auf eine gewählte Ausdrucksweise. Dies wurde einigen Schülerinnen und Schülern in der DVA allerdings zum Verhängnis. Sie konnten nicht schnell und kurz genug eine Frage in Stichworten beantworten. Auch eine von uns befragte Englischlehrerin in Thüringen hat sich genau mit den Grundfertigkeiten beschäftigt, die Lernende zur Bearbeitung von Hörverstehensaufgaben benötigen. Dazu hält sie sich an die Vorgaben des neuen Schulbuchs, das auf die Kompetenztests abgestimmt ist. Sie achtet z.B. bei der Aufgabenstellung von Hörtexten darauf, ob der Text einmal oder zweimal angehört werden kann. Auch beim Lesen achtet sie darauf, dass die Schülerinnen und Schüler allgemeine Kompetenzen wie Unterstreichen und Vergleichen beherrschen, was ihrer Meinung nach zu besseren Ergebnissen beim Kompetenztest geführt hat. Eine weitere Englischlehrkraft machte sich Gedanken zu zwei Grammatikthemen (*reported speech* und *adjective* vs. *adverb*), die in den DVA abgefragt und von den Schülern und Schülerinnen nur ungenügend verstanden wurden. Sie wiederholte beide Themen im Anschluss an die DVA, jedoch isoliert voneinander. Aus ihrer Sicht führte dies zu deutlich besseren Leistungen in einer anschließenden Klassenarbeit.

An die Lernvoraussetzungen angepasstes, differenziertes Üben

Einige Lehrkräfte berichten von einer gezielten Binnendifferenzierung im Anschluss an die zentralen Tests bzw. sehen bereits laufende Differenzierungsmaßnahmen durch die Testergebnisse bestätigt. Zum Teil greifen Lehrkräfte direkt auf Testaufgaben oder ähnliche Aufgabenformate zurück, um differenzierte Übungsphasen zu gestalten. Eine Mathematiklehrkraft weist darauf hin, dass es im Mathematikbuch zwar „tausende" von Übungsaufgaben gibt, die Schülerinnen und Schüler aber überfordert sind, sich die geeigneten Aufgaben auszuwählen. Aus diesem Grund entwickelte die Lehrkraft eigene Übungsaufgabenblätter mit Lösungen. Einige der in Baden-Württemberg interviewten Lehrkräfte berichten von sog. „WADI"-Aufgaben („Wiederholung – Analyse – Diagnose"). WADI ist eine zentral erstellte Aufgabensammlung zu verschiedensten Themenbereichen der Mathematik. Eine Lehrkraft gibt den Schülerinnen und Schülern für die in der DVA nur schwach bearbeiteten Themenbereiche die entsprechenden WADI-Aufgaben für Mathematik. Diese Differenzierungsaufgaben werden zunächst gemeinsam über die Dauer von ein bis zwei Wochen im Unterricht bearbeitet. Anschließend können die Schülerinnen und Schüler das Übungsmaterial selbständig bearbeiten und nach der Bearbeitung die Lösungen bei der Lehrkraft erfragen.

Vernetzung von Inhalten bzw. Fertigkeiten

Im Gegensatz zu der Argumentation, Grundwissen kleinschrittig zu wiederholen und zu üben, gibt es eine Reihe von Schlussfolgerungen, die auf eine stärkere Vernetzung von Inhalten bzw. Fertigkeiten abzielen. Eine baden-württembergische Deutschlehrerin greift den Bereich „Wortarten" auf und bindet ihn in das aktuell zu behandelnde Thema „Kurzgeschichten" ein. Ganz ähnlich argumentiert eine weitere Gymnasiallehrerin. Sie wird im Fach Deutsch das Thema „Novelle", mit dem die Schülerinnen und Schüler in der DVA Probleme hatten, im Unterricht aufgreifen und Kriterien wiederholen, die allen Erzähltexten zu Grunde liegen. Einer Deutschlehrkraft in Thüringen ist durch die Kompetenztests bewusst geworden, dass sie die Deutschdisziplinen bisher stark getrennt unterrichtet hatte. Mittlerweile verknüpft sie Gebiete wie „Wortarten" und „Ausdrucksaufgaben" und orientiert sich damit verstärkt am Prinzip „Sprache". In einem vierten Fall findet sogar eine fächerübergreifende Vernetzung statt. Die in der Mathematik-DVA geforderte Interpretation von Balkendiagrammen wird im Englischunterricht aufgegriffen.

Komplexere Aufgaben

Eine Reihe von Lehrkräften wurde durch die zentralen Tests zum Nachdenken über die Komplexität der eigenen Lernaufgaben angeregt. Es wird einerseits argumentiert, dass man die Schülerinnen und Schüler auf die „ungewöhnlichen" Aufgaben der Vergleichsarbeiten vorbereiten muss. Eine Lehrkraft möchte den „Erwartungshorizont" ausweiten. Andererseits finden sich aber auch Äußerungen, die auf die Bedeutung von anspruchsvollen Aufgabenstellungen für ein tieferes Verständnis des zu erlernenden Wissens hindeuten. Dieses lehr-lerntheoretische Argumentationsmuster findet sich überwiegend bei Mathematiklehrkräften. Eine Lehrkraft möchte verstärkt Aufgaben zur Wahrscheinlichkeitsrechnung von den Schülerinnen und Schülern selbst formulieren lassen. In einem anderen Fall werden vermehrt längere Textaufgaben in den Unterricht eingebaut.

Zusammenspiel verschiedener Repräsentationsformen von Wissen

Die zentralen Tests regen einige Lehrerinnen und Lehrer auch zum Nachdenken über das Zusammenspiel verschiedener Wissensrepräsentationen an. Für Mathematiklehrkräfte geht es dabei vor allem um die optimale Kombination von Text und grafischer Darstellung in Aufgaben. Das Fallbeispiel in Abschnitt 4.1 wurde dieser Kategorie zugeordnet. Die Lehrerin kommt zu der Erkenntnis, dass ihre eigenen Mathematikaufgaben immer sehr „abstrakt" sind und die „Bilder" in den Kompetenztests eine wertvolle Veranschaulichung des mathematischen Sachverhalts leisten können. Begründet wird dieser lehr-lerntheoretische Zusammenhang mit der

„Funktionsweise des Gehirns", das Veranschaulichungen und Zeit zur Verarbeitung benötige. Eine weitere Lehrkraft achtet im Anschluss an die Testauswertung verstärkt darauf, zwischen Textaufgaben mit viel Text, reinen Zahlenaufgaben, Aufgaben mit Tabellen und Aufgaben mit Bildern abzuwechseln. Sie verspricht sich davon eine „größere Bandbreite" und eine höhere Übungsmotivation.

Aber auch zwei Englischlehrkräfte machten sich im Anschluss an die zentralen Tests Gedanken über die sinnvolle Integration verschiedener Modi der Repräsentation von Sprache. Eine Englischlehrerin nutzt verstärkt die auf CD veröffentlichten Hörverstehensaufgaben. Die Lehrerin findet diese Kombination von Hören und schriftlicher Verarbeitung des Wissens sinnvoll und möchte auch in anderen Fächern (Latein, Geschichte) so vorgehen. In einem zweiten Fall wird dagegen die Problematik zu vieler verschiedener Hörbeispiele reflektiert. Die Englischlehrkraft lehnt es ab, sich weiteres „Zusatzmaterial zu besorgen". An die Hörbeispiele aus dem Schulbuch haben sich die Schülerinnen und Schüler gewöhnt. Bei buchfremden Hörbeispielen (z.B. VERA-Zusatzmaterial) sieht sie die Gefahr von „Sprachbarrieren", d.h. sie fürchtet eher eine Verwirrung der Lernenden durch andere Sprecher oder neuartige Aufmachungen der Hörbeispiele.

5. Zusammenfassung und Diskussion

In diesem Beitrag wurde der Versuch unternommen, mit qualitativen Lehrerbefragungen die lehr-lerntheoretischen Argumentationsmuster bei der Rezeption und Nutzung von Vergleichsarbeitsrückmeldungen zu erfassen und anschließend zu rekonstruieren. Dies ist bei einem Teil der von uns befragten Gymnasiallehrkräfte ansatzweise gelungen. In einer Reihe von Interviews war gut erkennbar, mit welcher fachdidaktischen Expertise und welchen Vorstellungen über Lehren und Lernen die Lehrkräfte bestimmte Schlussfolgerungen aus den Testrückmeldungen begründen. Allerdings setzten Erinnerungsvermögen und Auskunftsbereitschaft einiger Lehrkräfte einer validen Rekonstruktion von Argumentationsmustern auch deutliche Grenzen. Es gab eine Reihe von Interviews mit sehr spärlichen und kaum verwertbaren Begründungen oder mit einer generellen Ablehnung jeglicher Art von Nutzung der Testdaten. Zudem sind alle in einem Interview gemachten Aussagen grundsätzlich anfällig für nachträgliche Glättungen von Erinnerungslücken. Die Aussagen beziehen sich zudem auf eine Schulart und auf zwei Bundesländer.

Welche Schlussfolgerungen können nun aus den vorliegenden Daten gezogen werden? Trotz der möglichen Verzerrungen durch die Interviewsituation oder die Stichprobenauswahl konnte eine erstaunliche Vielfalt von lehr-lerntheoretischen Argumentationsmustern rekonstruiert werden. Große interindividuelle Differenzen bei der Rezeption und Nutzung zentraler Tests wurden in bisherigen Studien auch beschrie-

ben (vgl. z.B. Kühle 2010; Schulze 2012). Die qualitative Analyse zeigt darüber hinaus, dass Lehrkräfte vorwiegend auf ihre Praxiserfahrungen zurückgreifen, um testinduzierte Maßnahmen zu begründen. In vielen Fällen wird beispielsweise mit dem Verhältnis von Unterrichtszeit und Lernerfolg argumentiert. „Mehr Unterrichtszeit führt zu besseren Lernergebnissen", scheint eine sehr einfache, aber wirkmächtige subjektive Lehr-Lerntheorie zu sein, die ohne Rückgriff auf Studien oder fachdidaktische Ansätze auskommt. Auch das Fallbeispiel zeigt, dass einfache, „neurowissenschaftliche" Begründungen schneller zur Hand sind als fachdidaktische Theorien oder Prinzipien über Lernen im Mathematikunterricht.

Interessant ist auch, dass die Lehrkräfte aus lehr-lerntheoretischer Sicht teilweise in ganz unterschiedliche Richtungen argumentieren. Ein Teil der Lehrkräfte sieht eine kleinschrittige Erarbeitung bzw. Übung von Grundfertigkeiten geboten (behavioristisch gedacht), während ein anderer Teil der Lehrkräfte die Vernetzung von Lerninhalten und die Komplexität von Lernaufgaben erhöhen möchte (konstruktivistisch gedacht). Weiter fällt auf, dass Lehrkräfte bei der Reflexion über Maßnahmen nicht in Kompetenzkategorien, sondern in Stoff- bzw. Inhaltskategorien denken. Sie gehen von Stoffgebieten, die durch Testaufgaben repräsentiert werden, aus und reflektieren daraufhin ihre Vorgehensweise in der entsprechenden Unterrichtseinheit. Wenn von Kompetenzen die Rede ist, dann sind eher grundlegende Fertigkeiten oder Techniken (wie z.B. exaktes Rechnen oder genaues Lesen von Aufgaben) gemeint. Kompetenzstufen oder Kompetenzentwicklungsmodelle aus der VERA-Begleitliteratur (vgl. z.B. Blum u.a. 2010; KMK 2012b) wurden in den Interviews nicht erwähnt.

Auf einen Teil der Lehrkräfte machen die Testaufgaben, die überwiegend als interessant und innovativ wahrgenommen werden, einen gewissen Eindruck. Dies zeigt sich unter anderem in den Argumentationsmustern, die den Kategorien „Zusammenspiel verschiedener Repräsentationsformen" und „komplexe Aufgaben" zugeordnet werden konnten. Die Bedeutung der Testaufgaben für die VERA-Nutzung entspricht der bisherigen Befundlage (vgl. z.B. Nachtigall/Jantowski 2007; Maier 2009; Kühle 2010). Allerdings dominieren Oberflächenmerkmale von Testaufgaben die Reflexion. Beispielsweise wird von einigen Lehrkräften die Länge der Texte in mathematischen Modellierungsaufgaben entweder als Chance oder Problem beschrieben. Eine Schlussfolgerung ist dann die Einübung in das genaue Lesen von Texten. Die eigentlichen Anforderungen mathematischer Modellierungen werden in den Interviews dagegen nicht angesprochen (z.B. Erkennen eines mathematischen Modells in einer Sachsituation). Mögliche Hilfestellungen oder didaktische Ansätze zur Verbesserung der mathematischen Modellierungskompetenz werden nicht reflektiert (vgl. z.B. Greefrath 2007).

Die Englisch- und Französischvergleichsarbeiten regen einige Lehrkräfte auch zum Nachdenken über Hörverstehen als zentrale Fremdsprachkompetenz an. Die di-

daktischen Schlussfolgerungen sind jedoch unterschiedlich. Einige Lehrkräfte sehen in Hörverstehensaufgaben eine gute Ergänzung des bisherigen Unterrichts. Eine Lehrkraft argumentiert dagegen, dass verschiedene Hörbeispiele eher zu Verwirrung führen. Diese Begründung ist erfahrungsbasiert und durchaus plausibel, steht allerdings im Gegensatz zur fachdidaktischen Forderung, durch eine Variabilität der Hörbeispiele sprachliche Muster besser verankern zu können (vgl. Hu/Leupold 2008).

In den Interviewpassagen wird auch deutlich, welche Ressourcen eine lehr-lerntheoretische Nutzung des Feedbacks begünstigen bzw. behindern. Da sich die Nutzung sehr stark auf die Erhöhung des Zeitanteils für bestimmte Inhalte oder die Intensivierung der Bearbeitung entsprechender Aufgabenstellungen beschränkt, werden Lehrbücher, das Internet, Testaufgabensammlungen oder VERA-Korrekturanweisungen als wichtige Ressource erwähnt. Ergänzungsstunden werden als notwendige Ressource für binnendifferenzierende Maßnahmen genannt. In den Interviews wird aber auch deutlich, dass die übergeordneten Ebenen Fachkonferenz und Schulleitung eine wichtige Ressource für die Nutzung auf Unterrichtsebene darstellen. Wie in anderen Rezeptions- und Nutzungsstudien werden auch Zeitmangel, Geldmangel oder Desinteresse des Kollegiums als Haupthindernisse für die Umsetzung von Konsequenzen aus den Tests genannt (vgl. Maier u.a. 2012; Asbrand/Heller/Zeitler 2012).

Die Studie deutet insgesamt darauf hin, dass ein Teil der Lehrkräfte motiviert und in der Lage ist, zentrale Testrückmeldungen und mögliche Schlussfolgerungen vor dem Hintergrund der eigenen fachdidaktischen Expertise zu reflektieren. Allerdings zeigen die Beispiele auch, dass in dieser Hinsicht noch großer Klärungsbedarf besteht. Sowohl die Grundlagen für die Entwicklung von Vergleichsarbeiten als auch fachdidaktische Ansätze zur Begründung eines kompetenzorientierten Unterrichts scheinen für die meisten der von uns befragten Lehrkräfte entweder unklar, nicht artikulierbar oder schlichtweg unbekannt zu sein. Genau an dieser Stelle müssten die von der KMK (2012a) geforderten Fortbildungsmaßnahmen ansetzen. Vergleichsarbeiten könnten ein geeigneter Anlass sein, um zentrale Probleme eines Unterrichtsfaches näher zu beleuchten. Allerdings müssten hierfür zunächst einmal die fachdidaktischen Grundlagen zentraler Tests, aktuelle Ansätze und Forschungsbefunde der Fachdidaktik und die stark von Erfahrungen geprägte Expertise von Lehrkräften so aufeinander abgestimmt sein, dass eine gemeinsame Sprache möglich ist.

Literatur und Internetquellen

Altrichter, H./Maag Merki, K. (Hrsg.) (2010): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS.

Asbrand, B./Heller, N./Zeitler, S. (2012): Die Arbeit mit Bildungsstandards in Fachkonferenzen – Ergebnisse aus der Evaluation des KMK-Projektes for.mat. In: Die Deutsche Schule 104, H. 1, S. 31-42.

Baumert, J. (2001): Vergleichende Leistungsmessung im Bildungsbereich. In: Zeitschrift für Pädagogik, 43. Beiheft, S. 13-36.

Blum, W./Drüke-Noe, C./Hartung, R./Köller, O. (Hrsg.) ([4]2010): Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen. Berlin: Cornelsen Scriptor.

Blum, W./Drüke-Noe, C./Leiß, D./Wiegand, B./Jordan, A. (2005): Zur Rolle von Bildungsstandards für die Qualitätsentwicklung im Mathematikunterricht. In: Zentralblatt für Didaktik der Mathematik (ZDM) 37, H. 4, S. 267-274.

Bonsen, M./Büchter, A./Peek, R. (2006): Datengestützte Schul- und Unterrichtsentwicklung. Bewertungen der Lernstandserhebungen in NRW durch Lehrerinnen und Lehrer. In: Bos, W./Holtappels, H.G./Pfeiffer, H./Rolff, H.-G./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung, Bd. 14. Weinheim: Beltz, S. 125-148.

Bremerich-Vos, A./Granzer, D./Köller, O. (2008): Lernstandsbestimmung im Fach Deutsch. Gute Aufgaben für den Unterricht. Weinheim: Beltz.

Büchter, A./Leuders, T. (2005): Quality Development in Mathematics Education by Focussing on the Outcome: New Answers or New Questions? In: Zentralblatt für Didaktik der Mathematik (ZDM) 37, H. 4, S. 263-266.

Diemer, T./Kuper, H. (2011): Formen innerschulischer Steuerung mittels zentraler Lernstandserhebungen. In: Zeitschrift für Pädagogik 57, H. 4, S. 554-571.

Drüke-Noe, C./Keller, K./Blum, W. (2008): Bildungsstandards – Motor für Unterrichtsentwicklung und Lehrerbildung? In: Beiträge zur Lehrerbildung 26, H. 3, S. 372-382.

Fend, H. (2011): Die Wirksamkeit der Neuen Steuerung – theoretische und methodische Probleme ihrer Evaluation. In: Zeitschrift für Bildungsforschung 1, H. 1, S. 5-24.

Greefrath, G. (2007): Das Validieren diagnostizieren – Ein genauer Blick auf eine wichtige Teilkompetenz beim Modellieren. In: Praxis der Mathematik in der Schule 49, H. 15, S. 42-44.

Groß Ophoff, J./Hosenfeld, I./Koch, U. (2007): Formen der Ergebnisrezeption und damit verbundene Schul- und Unterrichtsentwicklung. In: Empirische Pädagogik 21, H. 4, S. 411-427.

Halbheer, U./Reusser, K. (2008): Outputsteuerung, Accountability, Educational Governance – Einführung in Geschichte, Begrifflichkeiten und Funktionen von Bildungsstandards. In: Beiträge zur Lehrerbildung 26, H. 3, S. 253-266.

Hopf, C. (2000): Qualitative Interviews – ein Überblick. In: Flick, U./von Kardorff, E./Steinke, I. (Hrsg.): Qualitative Forschung. Ein Handbuch. Reinbek: Rowohlt, S. 349-360.

Hu, A./Leupold, E. (2008): Kompetenzorientierung und Französischunterricht. In: Tesch, B./Leupold, E./Köller, O. (Hrsg.), Bildungsstandards Französisch: konkret. Sekundarstufe I: Grundlagen, Aufgabenbeispiele und Unterrichtsanregungen. Berlin: Cornelsen Scriptor, S. 51-84.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2010): Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung. Bonn/Berlin: Carl Link.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2012a): Beschluss zur Weiterentwicklung von VERA. Beschluss der Kultusministerkonferenz vom 08.03.2012.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2012b): Kompetenzstufenmodell zu den Bildungsstandards im Kompetenzbereich Rechtschreiben. Stand 25.04.2012. URL: http://www.iqb.hu-berlin.de/vera; Zugriffsdatum: 18.04.2013.

Kühle, B. (2010): Zentrale Lernstandserhebungen – Ergebnisorientierte Unterrichtsentwicklung? Schulische Strategien beim Umgang mit Ergebnissen aus den Schulrückmeldungen im Kontext der ersten Lernstandserhebungen 2004/2005 in Nordrhein-Westfalen. Berlin: Köster.

Kuper, H./Hartung, V. (2007): Überzeugungen zur Verwendung des Wissens aus Lernstandserhebungen. In: Zeitschrift für Erziehungswissenschaft 10, H. 2, S. 214-229.

Lorenz, J.H. (2005): Zentrale Lernstandsmessung in der Primarstufe: Vergleichsarbeiten Klasse 4 (VERA) in sieben Bundesländern (Central Student Assessment in Primary Schools: Comparative Tests for Grade 4 in Seven Federal States of Germany). In: Zentralblatt für Didaktik der Mathematik (ZDM) 37, H. 4, S. 317-324.

Maag Merki, K. (2010): Theoretische und empirische Analysen der Effektivität von Bildungsstandards, standardbezogenen Lernstandserhebungen und zentralen Abschlussprüfungen. In: Altrichter, H./Maag Merki, K. (Hrsg.): Handbuch neue Steuerung im Schulsystem. Wiesbaden: VS, S. 145-169.

Maier, U. (2008): Rezeption und Nutzung von Vergleichsarbeiten. Ergebnisse einer Lehrerbefragung in Baden-Württemberg. In: Zeitschrift für Pädagogik 54, H. 1, S. 95-117.

Maier, U. (2009): Wie gehen Lehrerinnen und Lehrer mit Vergleichsarbeiten um? Eine Studie zu testbasierten Schulreformen in Baden-Württemberg und Thüringen. Baltmannsweiler: Schneider Verlag Hohengehren.

Maier, U. (2010a): Effekte testbasierter Rechenschaftslegung auf Schule und Unterricht: Ist die internationale Befundlage auf Vergleichsarbeiten im deutschsprachigen Raum übertragbar? In: Zeitschrift für Pädagogik 56, H. 1, S. 112-128.

Maier, U. (2010b): Effekte von testbasiertem Rechenschaftsdruck auf Schülerleistungen: Ein Literaturüberblick zu quasi-experimentellen Ländervergleichsstudien. In: Journal for Educational Research Online 2, H. 2, S. 125-152.

Maier, U./Kuper, H. (2012): Vergleichsarbeiten als Instrumente der Qualitätsentwicklung an Schulen – Überblick zum Forschungsstand. In: Die Deutsche Schule 104, H. 1, S. 88-99.

Maier, U./Metz, K./Bohl, T./Kleinknecht, M./Schymala, M. (2012): Vergleichsarbeiten als Instrument der datenbasierten Schul- und Unterrichtsentwicklung in Gymnasien. In: Wacker, A./Maier, U./Wissinger, J. (Hrsg.): Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung. Empirische Befunde und forschungsmethodische Implikationen. (Educational Governance, Bd. 9.) Wiesbaden: VS, S. 197-224.

Mayring, P. ([11]2010): Qualitative Inhaltsanalyse. Grundlagen und Techniken. Aktualisierte und überarb. Aufl. Weinheim: Beltz.

Merkens, H. (1997): Stichproben bei qualitativen Studien. In: Friebertshäuser, B./Prengel, A. (Hrsg.): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim/München: Juventa, S. 97-106.

Nachtigall, C./Jantowski, A. (2007): Die Thüringer Kompetenztests unter besonderer Berücksichtigung der Evaluationsergebnisse zum Rezeptionsverhalten. In: Empirische Pädagogik 21, H. 4, S. 401-410.

O'Day, J.A. (2002): Complexity, Accountability, and School Improvement. In: Harvard Educational Review 72, H. 3, S. 293-329.

Peek, R./Dobbelstein, P. (2006): Benchmarks als Input für die Schulentwicklung. In: Kuper, H./Schneewind, J. (Hrsg.): Rückmeldung und Rezeption von Forschungsergebnissen. Zur Verwendung wissenschaftlichen Wissens im Bildungssystem. Münster u.a.: Waxmann, S. 41-58.

Pellegrino, J.W. (2003): Knowing What Students Know. Issues in Science and Technology 19, H. 2, S. 48-52.

Pellegrino, J.W./Chudowsky, N./Glaser, R. (2001): Knowing What Students Know. The Science and Design of Educational Assessment. Washington, DC: National Academic Press.

Scheerens, J./Glas, C./Thomas, S.M. (2003): Educational Evaluation, Assessment, and Monitoring. A Systemic Approach (Contexts of Learning. Classrooms, Schools and Society). Lisse/Abingdon/Axton/Tokyo: Swets & Zeitlinger.

Schildkamp, K./Kuiper, W. (2010): Data-informed Curriculum Reform: Which Data, What Purposes, and Promoting and Hindering Factors. In: Teaching and Teacher Education 26, H. 3, S. 482-496.

Schildkamp, K./Visscher, A. (2009): Factors Influencing the Utilisation of a School Self-Evaluation Instrument. In: Studies in Educational Evaluation 35, H. 4, S. 150-159.

Schildkamp, K./Visscher, A. (2010): The Use of Performance Feedback in School Improvement in Louisiana. In: Teaching and Teacher Education 26, H. 7, S. 1389-1403.

Schneewind, J./Kuper, H. (2009): Rückmeldeformate und Verwendungsmöglichkeiten der Ergebnisse aus zentralen Lernstandserhebungen. In: Bohl, T./Kiper, H. (Hrsg.): Lernen aus Evaluationsergebnissen. Verbesserungen planen und implementieren. Bad Heilbrunn: Klinkhardt, S. 113-129.

Schulze, F. (2012): Folgen zentraler Lernstandserhebungen: Unterscheiden sich Lehrkräfte in der Auseinandersetzung mit Lernstandserhebungen? Empirische Befunde zu Formen unterrichtsbezogener Steuerungslogik. In: Die Deutsche Schule 104, H. 1, S. 43-56.

Sill, H.-D./Sikora, C. (2007): Leistungserhebungen im Mathematikunterricht. Theoretische und empirische Studien. Hildesheim: Franzbecker.

Tesch, B./Leupold, E./Köller, O. (2008): Bildungsstandards Französisch: konkret. Sekundarstufe I: Grundlagen, Aufgabenbeispiele und Unterrichtsanregungen. Berlin: Cornelsen Scriptor.

Vanhoof, J./Verhaeghe, G./Verhaeghe, J.P./Valcke, M./Van Petegem, P. (2011): The Influence of Competences and Support on School Performance Feedback Use. In: Educational Studies 37, H. 2, S. 141-154.

Verhaeghe, G./Vanhoof, J./Valcke, M./Van Petegem, P. (2010): Using School Performance Feedback: Perceptions of Primary School Principals. In: School Effectiveness and School Improvement 21, H. 2, S. 167-188.

Visscher, A.J./Coe, R. (2003): School Performance Feedback Systems: Conceptualisation, Analysis, and Reflection. In: School Effectiveness and School Improvement 14, H. 3, S. 321-349.

Uwe Maier, Prof. Dr., geb. 1971, Prof. für empirische Schulforschung an der Pädagogischen Hochschule Schwäbisch Gmünd.
E-Mail: uwe.maier@ph-gmuend.de

Carolin Ramsteck, M.A., geb. 1982, wissenschaftliche Mitarbeiterin am Institut für Erziehungswissenschaft, Abteilung Schulpädagogik/Empirische Schulforschung der Pädagogischen Hochschule Schwäbisch Gmünd.
E-Mail: carolin.ramsteck@ph-gmuend.de

Anschrift: Pädagogische Hochschule Schwäbisch Gmünd, Arbeitsbereich empirische Schulforschung, Oberbettringerstraße 200, 73525 Schwäbisch Gmünd

Annette Frühwacht, Dr., geb. 1986, wissenschaftliche Mitarbeiterin am Lehrstuhl für Schulpädagogik der Friedrich-Alexander-Universität Erlangen-Nürnberg.
E-Mail: annette.fruehwacht@ewf.uni-erlangen.de

Anschrift: Friedrich-Alexander-Universität Erlangen-Nürnberg, Lehrstuhl für Schulpädagogik, Regensburger Straße 160, 90478 Nürnberg

DDS – Die Deutsche Schule
Beiheft 12, S. 97-120
© 2013 Waxmann

Barbara Muslic/Carolin Ramsteck/Harm Kuper

Das Verhältnis von Schulleitung und Schulaufsicht im Kontext testbasierter Schulreform

Kontrastive Fallstudien zur Rezeption von Lernstandsergebnissen im Mehrebenensystem der Schule

Zusammenfassung

Die testbasierte Schulreform hat eine grundlegende Veränderung der Qualitätsentwicklung im Schulsystem und die Neustrukturierung der Koordination zwischen Schulleitung und Schulaufsicht zur Folge. Mittels kontrastiver Fallstudien konnten auf diesen beiden Akteursebenen unterschiedliche Muster der Organisation von Qualitätsentwicklung identifiziert werden, die die Nutzung von Vergleichsarbeiten entweder befördern oder behindern.

Schlüsselwörter: Schulleitung, Schulaufsicht, Schulentwicklung, testbasierte Schulreform, Vergleichsarbeiten

The Relation between Principals and School Supervising Authorities in the Context of Test-Based School Reform

Contrastive Case Studies on the Usage of Mandatory Proficiency Test Data in the Multi-Level Structure of Schools

Summary

Test-based school reform has lead to a fundamental change of quality development in schools. A new organization of coordination between principals and school supervising authorities was thereby induced. For analytical purposes, this structural change was resembled by contrastive and qualitative case studies. On both levels particular patterns of pace-setting respectively foot-dragging actions regarding the usage of mandatory tests could be identified.

Keywords: principal, school supervising authority, school development, test-based school reform, state-wide mandatory proficiency tests

1. Einleitung

Als Reaktion auf die PISA-Ergebnisse wurden in den Bundesländern Deutschlands zentrale Lernstanderhebungen[1] als ein Element testbasierter Schulreform eingeführt. Sie dienen dem empirischen Abgleich der Bildungsstandards und zielen über die Rückmeldung von Ergebnissen auf die Sicherung bzw. Steigerung der Leistungsfähigkeit des Bildungssystems (vgl. Klieme 2004). Auf Schulebene sollen zentrale Lernstandserhebungen empirisches Steuerungswissen bereitstellen, um die Unterrichtsqualität zu verbessern und die Kompetenzen sowie den Förderbedarf der Schüler und Schülerinnen transparent zu machen (vgl. Tillmann u.a. 2008). Lernstandserhebungen gelten somit als Instrument zur Schul- und Unterrichtsentwicklung.

Zu zentralen Lernstandserhebungen liegen mittlerweile einige Studien vor (zur Übersicht vgl. Hartung-Beck 2009; Maier/Kuper 2012). Diese behandeln vorwiegend die schulinterne Rezeption und Nutzung der zurückgemeldeten Ergebnisse aus den Lernstandserhebungen. In ihrem Fokus liegen fast ausschließlich Lehrkräfte (vgl. u.a. Helmke 2004; Groß Ophoff/Hosenfeld/Koch 2007). Wie mit zentralen Lernstandserhebungen auf übergeordneten Ebenen des Schulsystems, besonders auf den Ebenen der Schulleitung und der Schulaufsicht[2] umgegangen wird, wurde bislang noch nicht systematisch untersucht. Da zentrale Lernstanderhebungen zum Vergleich in und zwischen Schulen anregen sollen (vgl. EMSE 2008), sind gerade diese beiden hierarchischen Ebenen zentral für die Nutzung ihrer Ergebnisse (vgl. Campbell/Gross 2012). Empirische Arbeiten, die den Umgang mit VERA und die Verwendung der Daten durch die Schulleitung und Schulaufsicht fokussieren, sind ein Desiderat.

Die Verantwortungs- und Tätigkeitsbereiche von Schulleitung und Schulaufsicht haben sich durch diese bildungspolitischen Rahmenbedingungen, vor allem durch die Einführung von Vergleichsarbeiten, erweitert (vgl. Brüsemeister/Newiadomsky 2008). Ob und auf welche Weise sich diese Neustrukturierungen – insbesondere in Bezug auf die Zuständigkeit und die Auseinandersetzung mit den zurückgemeldeten Ergebnissen – im Kontext von Lernstandserhebungen darstellen, soll im Folgenden untersucht werden.

1 Lernstandserhebungen werden z.B. auch als Vergleichsarbeiten (Abkürzung VERA) oder Kompetenztests bezeichnet. Im Folgenden werden diese Begriffe synonym verwendet.

2 Innerhalb dieser Arbeit wird die Bezeichnung Schulaufsicht für die der Schule übergeordnete Akteursebene verwendet. Sie ist in den hier vorgestellten Bundesländern für Vergleichsarbeiten zuständig.

2. Forschungsstand

2.1 Schulleitung im Kontext zentraler Lernstandserhebungen

Schulleitung ist in der deutschsprachigen Bildungsforschung ein recht junges Thema (vgl. Terhart 1997), das im Zuge zweier miteinander verkoppelter Entwicklungen – der Stärkung der Autonomie einzelner Schulen (vgl. Schleicher 2009) und der Professionalisierung von Schulleitungen – an Bedeutung gewonnen hat. Die Forschungsbeiträge lassen sich folgenden Bereichen zuordnen:

- Betrachtungen mit handlungsorientierendem Charakter und teilweise normativ rückversichernder Argumentation, die sich mit den Anforderungen an die Tätigkeit der Schulleitung und deren Rolle auseinandersetzen (u.a. Wissinger 2011);
- Beschreibungen zum Wandel des Aufgabenbereichs und der Führungstätigkeit von Schulleitung von einem administrativen zu einem manageriellen Zuschnitt (u.a. Rosenbusch 2005; Harazd/Gieske/Rolff 2008);
- Überlegungen zur Qualifikation und Auswahl von Schulleitern und Schulleiterinnen (zur Übersicht vgl. Bonsen/Pfeiffer 1998);
- empirische Studien zu den Folgen veränderter Anforderungen an Schulleitungen unter den Aspekten der Koordination zwischen Schulleitung und Lehrerkollegium (Kuper 2002), der Führungsstile (Bonsen 2003), der zeitlichen Beanspruchung (Warwas 2009a), des beruflichen Selbstkonzeptes (Warwas 2009b) und der Delegation von Verantwortung an das Lehrerkollegium (Brauckmann 2012).

Das thematische Portfolio der Schulleitungsforschung schließt in vielfältiger Weise zur Forschung über die Rezeption von Lernstandserhebungen und zur Sicherung von Unterrichtsqualität auf. Diese Relation ist aber bislang nicht dezidiert ausgearbeitet worden. Insgesamt resultieren aus den Modellen der neuen Steuerung Veränderungen für Schulleitungen, die sich neben der Erweiterung des Funktionsbereichs auch auf neue Kompetenzanforderungen und den steigenden professionellen Anspruch beziehen (vgl. Böttcher 2002; Buchen/Rolff 2009). Mit der Bindung an Bildungsstandards und der Rückmeldung von Leistungsergebnissen an die Schulen beinhaltet das Konzept der Lernstandserhebungen zentrale und dezentrale Elemente der Steuerung des Bildungssystems. Beide im Sinne einer schulinternen Sicherung von Unterrichtsqualität zu verkoppeln, fällt in den Verantwortungsbereich der Schulleitung. Die auf die Schul- und Unterrichtsentwicklung ausgerichtete Verarbeitung der Ergebnisse zentraler Lernstandserhebungen wird als Managementaufgabe der Schulleitungen aufgefasst (vgl. Bonsen 2011; Brauckmann/Herrmann 2012). So erwartet die EMSE (2008, S. 5) von Schulleitungen die Schaffung angemessener Rahmenbedingungen für die Verarbeitung, bspw. in Form „einer professionell gestalteten Rückmeldekultur als Gelenkstelle zwischen Unterrichtsentwicklung und Evaluation sowie die Sicherstellung einer verbindlichen Zusammenarbeit von Lehrpersonen, vor allem in den Fachkonferenzen."

Im Gefolge der *School-Effectiveness-Research* (vgl. u.a. Mortimore u.a. 1988; Sammons/Hillman/Mortimore 1995) hat sich unter dem Stichwort S*chool Leadership Effects* (vgl. zusammenfassend Scheerens 2012) ein Strang der Schulleitungsforschung entwickelt. Diese reflektiert den Einfluss von Schulleitungen auf die Qualität des Unterrichts und das Lernen der Schüler und Schülerinnen (vgl. Leithwood u.a. 2004). Die Schulleitungen nehmen einen mittelbaren Einfluss auf die Leistungsentwicklung der Schülerinnen und Schüler (vgl. Supovitz/Sirinides/May 2010).

Das Argument einer indirekten Wirkung der Schulleitungen findet sich auch in den Modellen der *Distributed Leadership* (vgl. Spillane/Halverson/Diamond 2001; Gronn 2002). Es beschreibt Mechanismen der beteiligungsorientierten Entscheidung in Schulen und die Verankerung von Verantwortung in der organisatorischen Struktur der Einzelschule. Nach Fletcher und Käufer (2003, S. 22) sind Praktiken der *Distributed Leadership* „enacted by people at all levels rather than a set of personal characteristics and attributes located in people at the top". Damit verweist das Modell auf das Desiderat, die Aufmerksamkeit in der Schulleitungsforschung von der Person des Schulleiters bzw. der Schulleiterin abzuziehen und auf das in die Schulorganisation eingebettete Handeln zu richten. Im Sinne dieses Ansatzes ist eine enge Verkopplung der Entscheidungen von Schulleitungen mit den kollegialen und organisatorischen Strukturen der Schule ein erfolgskritisches Merkmal (vgl. Harris 2008, S. 172). Mit der Konzentration der Rezeptionsforschung auf Themen wie die Außenbeziehungen der Schule, die schulinterne Arbeit an Programmen und Curricula, die Evaluation des Unterrichts sowie fachliche Leistungen der Schüler und Schülerinnen und die professionelle Kompetenz von Lehrkräften werden genuine Verantwortungsbereiche der Schulleitungen im Verständnis der *Distributed Leadership* angesprochen.

Jüngere Befunde der Rezeptionsforschung schließen an die skizzierte Argumentation aus der Schulleitungsforschung an. Die Rückmeldung von Ergebnissen aus zentralen Lernstandserhebungen erfordert von den Schulen ein hohes Maß an aktiver Verarbeitung der Information und Verantwortung bei der Ableitung von Maßnahmen. Die Rezeption der Ergebnisse stellt auch eine Herausforderung für Schulleitungen dar. Folgende Befunde spezifizieren die Relation zwischen Lernstandserhebungen und Schulleitungen:

- Zur Erklärung der Auswirkungen von Lernstandserhebungen auf die Arbeit an Schulen wird auf die schulübergreifenden Regelungskontexte verwiesen (vgl. Maier 2010). Insbesondere die Differenz zwischen *low-stakes* und *high-stakes testing* strukturiert die Möglichkeiten vor, Schulen untereinander in kooperative oder konkurrierende Beziehungen zu bringen.
- Generell wird auf die hohe Varianz der Rezeptionspraktiken zwischen den Schulen verwiesen (vgl. Koch 2011), wodurch die Aufmerksamkeit auf die schulinternen organisatorischen Voraussetzungen für eine produktive Verwendung von Ergebnissen aus Lernstandserhebungen gelenkt wird.

2.2 Die Rolle der Schulleitung in der Schulorganisation

Wirkungsvolle Praktiken der Rezeption und Nutzung von Ergebnissen aus Lernstandserhebungen hängen von den jeweiligen internen Organisationsstrukturen sowie Rahmenbedingungen der Schulen ab (vgl. Visscher/Coe 2003). Aus einer organisationstheoretischen Perspektive variiert die Rolle der Schulleitungen zwischen der Sicherung eines Anspruchs auf professionelle Autonomie in der verwalteten Schule *(Autonomous Professional Organization)* und der Leitung einer effizienzorientierten Schule, die funktional differenziert und strategisch sowie operativ vernetzt ist *(Managed Professional Organization)* (vgl. Thiel 2008). Die *Autonomous Professional Organization* ist gekennzeichnet durch eine Janusköpfigkeit der Schulleitungen, die in den Rollen der Kollegen und Kolleginnen als Verwaltungschefs agieren. Ein weiteres typisches Merkmal dieser (Schul-)Organisation ist die Aufrechterhaltung der *„ceremonial facade"* autonom handelnder Professioneller nach außen.

Demgegenüber steht das Konzept der *Managed Professional Organization*. Es entspricht einer leistungsverantwortlichen Expertenorganisation, wobei insbesondere die Entscheidungskompetenzen der Schulleitung und deren strategische und operative Verantwortung im Vordergrund stehen. Der Strukturdifferenzierung innerhalb dieser Schulorganisation entsprechend „werden neue Formen der horizontalen und vertikalen Aufgabenteilung implementiert, von besonderem Interesse sind hier neben der Erweiterung von Funktionsrollen (z.B. Evaluationsbeauftragte) die Differenzierung von Entscheidung und Entscheidungsvorbereitung (Steuergruppe) und der Ausbau hierarchischer Entscheidungskompetenz durch die Neuprofilierung der Schulleiterrolle" (ebd., S. 37). Rückübertragen auf zentrale Lernstandserhebungen kann davon ausgegangen werden, dass sie Entwicklungen hin zu einer *Managed Professional Organization* begünstigen.

2.3 Schulaufsicht im Kontext zentraler Lernstandserhebungen

Das Instrument der zentralen Lernstandserhebungen hat nicht nur auf Ebene der Schulleitung Auswirkungen auf das Selbstverständnis und das Handeln der Akteure. Auch die Ebene der Schulaufsicht unterliegt einem Wandel ihrer Rolle und Funktion. Zentrale Lernstandserhebungen dienen auf übergeordneter Makroebene der Rechenschaftslegung, indem sie das an einer Schule erreichte Kompetenzniveau am Maßstab der Bildungsstandards abbilden. Von dort ausgehend soll die Schulaufsicht Qualitätssicherung auf der Mesoebene der Schulentwicklung und auf der Mikroebene der Unterrichtsentwicklung initiieren und unterstützen (vgl. Maier/Kuper 2012). Wie sich die Schulaufsicht in der schulischen Mehrebenenstruktur und im Rahmen von Veränderungen durch (testbasierte) Schulreformen gestaltet, wird bislang in Deutschland kaum untersucht (vgl. Laukart 2009).

Die beinahe antagonistischen Ziele von VERA (Kontrolle vs. Beratung/Verbesserung) spiegeln sich im Verantwortungsbereich der Schulaufsicht bereits seit einigen Jahrzehnten wider. Seit den 1970er-Jahren, als der Deutsche Bildungsrat eine Trennung der Aufgaben der Schulaufsicht im Rahmen der Autonomiedebatte forderte, steht die Schulaufsicht in Deutschland zwischen einer Beratungs- und einer Kontrollfunktion (vgl. Schnell 2006). Die empirische Schulleistungsforschung (*School-Effectiveness-Research*) kam in den 1980er- und frühen 1990er-Jahren jedoch zu dem Schluss, dass die Steuerungswirkung von administrativ gesteuerten Reformen überschätzt wird. Dementsprechend wurde die Einzelschule – und mit ihr die Schulleitungen – als treibende Kraft von Schulreformen gesehen. In diesem Kontext forderten Burkard und Rolff (1994, S. 206), dass ein neues Steuerungskonzept, das auf autonome Schulen setzt, von einem „korrespondierenden Steuerungsmodell durch staatliche Schulaufsicht" begleitet sein muss.

2.4 Interdependenzmanagement als neue Aufgabe der Schulaufsicht

Durch den neuen, testbasierten Blick auf die Schule wird die Beurteilung von Lehrkräften als bisheriges zentrales Tätigkeitsfeld der Schulaufsicht abgelöst. Es wird gefordert, dass sich die Schulaufsicht als dienstleistungsorientiertes Unterstützungssystem reorganisiert (vgl. Laukart 2009, S. 49), das sich nach Lange (2003) „auf die Schule als organisatorische Einheit beziehen" (S. 145) soll. Für das Instrument VERA bedeutet dies bspw., dass die Schulaufsicht die Testimplementation an Schulen unterstützt. Durch die aufkommende Orientierung der Verwaltung an Modellen wie dem *New Public Management (NPM)*, das auf Output-Orientierung, Dezentralisierung und Kundennähe setzt (vgl. Dubs 1996), kam die Frage auf, wie die schulische Entwicklung auch weiterhin staatlich gesichert werden kann.

Die *Governance*-Forschung beantwortet dies durch eine interdisziplinäre Analyse von Steuerungs- und Umstrukturierungsprozessen im Bildungswesen, „die […] als Probleme der Handlungskoordination zwischen Akteurskonstellationen in einem Mehrebenensystem" (Altrichter/Brüsemeister/Wissinger 2007, S. 10) gesehen werden. Hierbei treten unterschiedliche Typen schulaufsichtlichen Handelns auf (vgl. Brüsemeister/Newiadomsky 2008). Diese Typen variieren in ihrem Führungsstil (kommunikativ vs. formalistisch) und der Organisationsart (Management vs. Bürokratie) und befördern somit in Analogie zu Thiel (2008) entweder das schulische Organisationskonzept der *Autonomous Professional Organization* oder das der *Managed Professional Organization*. Der Typ des *Neuen Managements* entspricht den Charakteristika des NPM. Durch seine stark kommunikative Ausrichtung ist er auf das „Interdependenzmanagement" (Brüsemeister/Newiadomsky 2008, S. 87), die Handlungskoordination und die Kommunikation zwischen den unterschiedlichen Ebenen ausgerichtet. Die Schule wird als *gemanagte professionelle Organisation* wahrgenommen, indem sie effizient geleitet, funktional differenziert

sowie strategisch und operativ vernetzt ist. Der Typ des *Distanzierten Managements* ist eine hybride Form: Er umfasst zwar die Managementorientierung des NPM, das Interdependenzmanagement ist jedoch nicht ausgeprägt. So externalisieren Schulaufsichtsvertreter dieses Typs Erwartungen z.B. im Umgang mit VERA an die Schulen, ohne dabei Unterstützung anzubieten. Die Schule wird als *autonome professionelle Organisation (Autonomous Professional Organization)* verstanden. Daher kann angenommen werden, dass die Umsetzung von VERA begünstigt wird, wenn eine als *gemanagte professionelle Organisation (Managed Professional Organization)* geleitete Schule auf den Schulaufsichtstypen des *Neuen Managements* trifft.

2.5 Zum Verhältnis von Schulaufsicht und Schulleitung im Kontext zentraler Lernstandserhebungen

Im Rahmen der testbasierten Schulreform werden der Schulaufsicht und der Schulleitung wichtige komplementäre Rollen zugeschrieben. Die Schulaufsicht soll die Schule bei der Implementation testbasierter Schulentwicklung unterstützen und weniger als Kontrolleur tätig werden. Als der Schulleitung übergeordnete Ebene hat sie die Aufgabe, schulische Entwicklung aus einer übergreifenden Perspektive voranzutreiben (vgl. Weitzel 2002). Wie sie zu den Schulen Kontakt aufnimmt, ist bundeslandspezifisch gestaltet (vgl. Füssel/Leschinsky 2008). Zumeist beschränkt sich der Kontakt jedoch auf die Ebene der Schulleitung. Es ist erwartbar, dass die Art der Kontaktaufnahme abhängig vom Typus des Schulaufsichtshandelns ist. Im Zusammenhang mit der Forschung zum Schulleitungshandeln kann angenommen werden, dass die beiden managementorientierten Typen des Schulaufsichtshandelns die *Distributed Leadership* in einer gemanagten professionellen Organisation wie der Schule befördern. Der Typ des *Neuen Managements* legt den Schwerpunkt seines Handelns auf Kommunikation und wird dabei dem Konzept der *gemanagten professionellen Organisation* gerecht. Der formalistische Typ des *Distanzierten Managements* entspricht hingegen dem Konzept der *autonomen professionellen Organisation*.

Die Implementation von Lernstandserhebungen führt zu Veränderungen im Gefüge der Koordination von Schulleitung und Schulaufsicht. Wie sich dies gestaltet, kann mittels Modellen zur Rezeption von Evaluationsergebnissen nachvollzogen werden. So unterscheiden das *Evaluationszyklusmodell* von Helmke (2004) und das *School-Performance-Feedback-System*-Modell (kurz: SPFS-Modell) von Visscher und Coe (2003) Aspekte und (Prozess-)Phasen der Implementation, Verarbeitung und Verwertung von Vergleichsarbeitsdaten. Diese Modelle können auf alle schulischen Akteure übertragen werden.

3. Fragestellung

Ziel der im Folgenden vorgestellten empirischen Studie ist es, Verantwortungsbereiche und Nutzungspraktiken auf den Ebenen der Schulleitung und der Schulaufsicht im Kontext testbasierter Steuerung anhand von Fallstudien darzustellen. Hierbei soll die Annahme überprüft werden, dass das Zusammenspiel verschiedener schulischer Organisationskonzepte mit unterschiedlichen schulaufsichtlichen Handlungstypen die Nutzung von VERA entscheidend beeinflusst.

Es werden dabei folgende Fragestellungen analysiert:
1) Wie adressieren und koordinieren Schulleitungen weitere schulische Akteure im Hinblick auf die Entwicklung der Schul- und Unterrichtsqualität auf Grundlage zurückgemeldeter Lernstandsergebnisse?
2) Wie gestalten sich das Verhältnis und die Zusammenarbeit von Schulleitung und Schulaufsicht bei der Auseinandersetzung mit den zurückgemeldeten Lernstandsergebnissen?

3.1 Stichprobe und Datenbasis

Zur Beantwortung der Fragen wurden im qualitativen Forschungsprojekt „Die Realisierung testbasierter Schulreform in der Mehrebenenstruktur des Schulsystems“[3] 229 leitfadengestützte Interviews an 19 Gymnasien aus vier Bundesländern (Baden-Württemberg, Berlin, Brandenburg und Thüringen) geführt, an die Fallstudien angeschlossen wurden. In zwei Erhebungszeiträumen wurden Lehrkräfte der Sekundarstufe I, Fachbereichsleitungen, Schulleitungen (35 Interviews) sowie Vertreter und Vertreterinnen der Schulaufsicht (17 Interviews) befragt. Bei der Auswahl der Schulen war das theoretische Sampling nach Glaser/Strauss (2010) leitend. Die Schulen sollten einen aktiven Umgang mit Lernstandserhebungen erwarten lassen und stammen aus Bundesländern, in denen verschiedene Testverfahren zur Anwendung kommen.

3.2 Methodisches Vorgehen

Für die Interviewbefragung wurden Leitfäden entwickelt, deren theoretische Grundlage die bereits erwähnten Modelle zur Rezeption von Evaluationsergebnissen (Helmke 2004; Visscher/Coe 2003) sind. Es wurden Fragen zur Einstellung zu den Lernstandserhebungen, zum Wissen über deren Konzeption und zur Durchführung der Tests gestellt. Dies bildet die Bereiche *Design Process*, *SPSF Features* und

3 Das Projekt wurde vom Bundesministerium für Bildung und Forschung unter dem Forschungsschwerpunkt „Steuerung im Bildungswesen“ gefördert und im Verbund von Wissenschaftlern der Freien Universität Berlin, der Friedrich-Alexander-Universität Erlangen-Nürnberg und der Pädagogischen Hochschule Schwäbisch-Gmünd durchgeführt.

Implementation Process aus dem *SPSF*-Modell von Visscher und Coe ab. Fragen nach der nachhaltigen Verwendung der Ergebnisse sowie der Reflexion der Datennutzung ordnen sich bei Visscher und Coe dem Bereich *SPFS Usage* bzw. bei Helmke den Phasen Rezeption, Reflexion, Aktion und Evaluation zu.

Um den Spezifika der jeweiligen Akteursebene gerecht zu werden, wurden für alle Ebenen eigene Interviewleitfäden erarbeitet. Darüber hinaus fand eine Erhebung der Kontextinformationen zu den Schulen statt, um die Einflussnahme der von Visscher und Coe als *School Organisation Features* bezeichneten Rahmenbedingungen nachvollziehen zu können. Auch bei der Auswertung wurden die Spezifika der einzelnen Ebenen durch eigene Kodierschemata berücksichtigt.

Um die Güte der Kodierung und der Kodierschemata zu überprüfen, wurde die Interkodier-Reliabilität durch *Cohen's Kappa* (κ) bestimmt. Da innerhalb des Verbundprojekts bei den in der vorliegenden Arbeit betrachteten Ebenen der Schulaufsicht und Schulleitung insgesamt sieben Personen kodiert haben, war es unabdingbar, sich gegen Fehler der kategorialen Analyse abzusichern (vgl. Cohen 1960; Grouven u.a. 2007). Für die beiden hier analysierten Ebenen wurden jeweils zehn Prozent der Codings zweitkodiert. Die Beurteilung der Interrater-Reliabilität basiert auf Wirtz und Caspar (2002). Es konnten durchweg akzeptable bis sehr gute *Cohen's Kappa*-Werte ermittelt werden ($0.45<\kappa<0.80$). Im Nachgang der Bestimmung der Interkodier-Reliabilität wurde die Kodierung auf der Grundlage der geprüften Kodierschemata durchgeführt. Bei den Kodierschemata, bei denen lediglich akzeptable Reliabilitätswerte erzielt wurden, wurde im Forscherteam über die Ausprägung der jeweiligen Textstellen diskutiert, um Konsens über die Kodierungen herzustellen. Es fand eine Präzisierung der Kodierregeln für die anschließenden Kodiervorgänge statt.

Zudem kamen unterschiedliche Erhebungsmethoden zum Einsatz (Methodentriangulation durch Befragung, Beobachtung und Dokumente) (vgl. Flick u.a. 2012). Durch die Offenlegung der deskriptiven Befunde in der Grundauswertung[4] ist die Prüfung der Objektivität sichergestellt. Die externe Validität ist durch die Einhaltung der Kriterien der Inhalts- und Konstruktvalidität gewahrt.

Nach der Strukturierung des Datenmaterials durch die Inhaltsanalyse nach Mayring (2010) wurden Einzelfallanalysen (*Within-Case Analysis*) entwickelt, um besondere Handlungskoordinationen auf Ebenen der Schulleitung und der Schulaufsicht zu veranschaulichen und wissenschaftlich zu explorieren (vgl. Heimerl 2007). Die Durchführung einer Fallstudie nach Yin (2009) ermöglicht es, ein „zeitgenössisches Phänomen in einem realen Kontext" (ebd., S. 18) zu untersuchen. Als Einzelfall gilt

4 Die umfangreiche, ca. 300 Textseiten umfassende kategoriale Grundauswertung dient als Referenz für die fallbezogene und die kontrastierende Auswertung (vgl. URL: http://www.ewi-psy.fu-berlin.de/einrichtungen/arbeitsbereiche/weiterbildung_bildungsmanagement/media/downloads/kategoriale_grundauswertung.pdf?1364397875; Zugriffsdatum: 19.07.2013).

bei der vorliegenden Studie die organisationale Konstellation Einzelschule – Schulaufsicht unter Berücksichtigung der institutionellen Regelungsstrukturen.

4. Ergebnisse

4.1 *Within-Case Analysis* auf den Ebenen der Schulleitung und der Schulaufsicht

Die empirischen Befunde, die im Folgenden vorgestellt werden, beruhen auf Fallstudien an Schulen in den Bundesländern Berlin und Brandenburg und stammen aus beiden Erhebungszeiträumen[5]. Daraus werden vier ausgewählte Einzelfälle, jeweils zwei auf den Ebenen Einzelschule/Schulleitung (Fälle A und B) und der dazugehörigen Schulaufsicht (Fälle C und D), präsentiert, die die Zuständigkeit und Auseinandersetzung von Schulleitung und -aufsicht mit VERA vor dem Hintergrund veränderter Rahmenbedingungen fokussieren.

Für die Analyse der Fragestellung 1, die darauf zielt, wie Schulleitungen weitere schulische Akteure auf Grundlage zurückgemeldeter Lernstandsergebnisse adressieren, wurden auf Schulleitungsebene vier Kategorien innerhalb der Inhaltsanalyse berücksichtigt:

- Konzeptionelles Verständnis,
- Kommunikation,
- Rahmenbedingungen für Verarbeitung,
- Maßnahmen/Entscheidungen.

Die Kategorie „Konzeptionelles Verständnis“ umfasst Aussagen von Schulleitungen über ihr Verständnis der mit VERA erfassten Kompetenzen, der Testkonstruktion und der Ergebnisrückmeldung. Basierend auf diesen Vorstellungen der interviewten Schulleitungen können erste Hinweise zu ihrer Rolle und Handlungskoordination (bspw. in Bezug auf Zuständigkeit und Verantwortung für VERA) in der Schulorganisation entnommen werden. Die Kategorie „Kommunikation“ beinhaltet Angaben der Schulleitungen zu Kommunikationsstrukturen im Kontext der Reaktion auf VERA-Ergebnisse. Dabei wird der Austausch sowohl mit anderen Akteuren (Lehrkräfte, Schulaufsicht) als auch innerhalb von Gremien (Fach- und Gesamtlehrerkonferenz) erfasst. Unter der Kategorie „Rahmenbedingungen für Verarbeitung“ werden strukturelle, organisatorische und individuelle Gegebenheiten, die die Verarbeitung von VERA-Ergebnissen begünstigen oder hemmen, subsumiert. Die Kategorie „Maßnahmen/Entscheidungen“ setzt sich zusammen aus Aussagen über Maßnahmen auf Schul- und Unterrichtsebene, die basierend auf den Lernstandsergebnissen abgeleitet werden. Die Informationen aus diesen vorangegan-

5 Der erste Erhebungszeitraum ist in der Signatur mit „A“ gekennzeichnet, der zweite mit „B“.

genen vier Kategorien ermöglichen Rückschlüsse auf die Ausprägung des organisationsbezogenen Führungshandelns von Schulleitungen im Kontext von VERA.

Durch die Kategorisierung und Ordnung des Materials konnten für die Ebene der Schulleitung kontrastierende Fälle herausgearbeitet werden. Um diese Fälle in der schulischen Mehrebenenstruktur verorten zu können, wurden die beiden Interviews derjenigen Schulaufsichtsvertreter hinzugezogen, in deren Zuständigkeitsbereich die Schulen der Fälle A und B liegen. Bei dieser Untersuchung war die Forschungsfrage 2 maßgebend. Hierfür sind die folgenden drei Kategorien der Inhaltsanalyse relevant:

- Konzeptionelles Verständnis,
- Nutzung von VERA durch Schuladministration,
- Verständnis der eigenen Rolle.

Die Kategorie „Konzeptionelles Verständnis" erfasst analog zur Schulleitungsebene die handlungsleitende Auffassung von VERA. Dies lässt Rückschlüsse über die Ausprägung der Managementorientierung und der wahrgenommenen Zuständigkeit der Schulaufsicht für die Aufarbeitung von Vergleichsarbeitsdaten zu. In den Kategorien, die der „Nutzung von VERA durch Schuladministration" zugeordnet sind, sind diejenigen Fundstellen subsumiert, die eine Verwendung von VERA auf Schuladministrationsebene belegen und somit Hinweise auf die Ausgestaltung des Interdependenzmanagements geben. Zudem kann die Schuladministration, der auch die Dienstaufsicht obliegt, die Vergleichsarbeitsresultate zu Kontrollzwecken einsetzen. Dies lässt Rückschlüsse auf die Managementorientierung und die formalistische Bearbeitung der Ergebnisse aus Lernstandserhebungen zu. Die Kategorie „Verständnis der eigenen Rolle" beinhaltet Aussagen zum Selbstverständnis der Schuladministration im Hinblick auf die VERA-Ergebnisse (wahrgenommene Zuständigkeit und Externalisierung an andere Akteure).

4.2 Wie adressieren und koordinieren Schulleitungen weitere schulische Akteure im Hinblick auf die Entwicklung der Schul- und Unterrichtsqualität auf Grundlage zurückgemeldeter Lernstandsergebnisse?

Im Folgenden werden zwei Fallstudien[6] (A und B), deren Formen der Koordination in stärkstem Kontrast zueinander stehen, präsentiert. Gleichzeitig stellen die Fallstudien in Rekurs auf Thiel (2008) sowie Brüsemeister/Newiadomsky (2008) idealtypische Organisationskonzepte und differente Handlungstypen dar, die miteinander verknüpft werden.

6 Für die Explikation der Fallstudien werden Zitate in paraphrasierter Form verwendet, um den Lesefluss zu gewährleisten. Alle angegebenen Zitate sind mit Zeilenangaben versehen.

Während Fallbeispiel A der idealtypischen Schulleitung einer *gemanagten professionellen Organisation* entspricht, ist die Schulleitung im Fall B einer *autonomen professionellen Organisation* zuzuordnen.

Im Fallbeispiel A lassen verschiedene Hinweise darauf schließen, dass die Schulleitung eine *gemanagte professionelle Organisation* leitet. Innerhalb der Kategorien „Kommunikation" (6 BS[7]) und „Rahmenbedingungen für Verarbeitung" (3 BS) zeigt sich das vor allem in der Förderung der strategischen und operativen Vernetzung aller Ebenen. Damit sorgt die Schulleitung für den organisationalen Rahmen für die Durchführung und Rezeption von Lernstandsergebnissen und festigt die Vernetzung über alle Ebenen hinweg. Dabei agiert die Schulleitung – im Sinne der organisatorischen Verankerung der Verantwortung für die Qualität des Unterrichts – in besonderer Weise koordinierend. In diesem Falle werden die Lernstandsergebnisse zum Anlass genommen, um die Wahrnehmung professioneller Verantwortung durch die Lehrkräfte zu intensivieren, um Maßnahmen zur Sicherung der Unterrichts- und Schulqualität bzw. deren (programmatische) Verankerung über die Gremien zu forcieren und um die Leistungsfähigkeit der Schule mit den Erwartungen und Bedingungen des sozialen Umfeldes abzustimmen (vgl. 22SL1B, 44). Diese Schulleitung einer *gemanagten professionellen Organisation* korrespondiert mit dem Schulaufsichtstypen des *Neuen Managements*, der stark kommunikativ und managementorientiert agiert.

Funktionsdifferenzierung als Merkmal der *gemanagten professionellen Organisation* ist im Fall A unter anderem darin erkennbar, dass die Schulleitung ihre Lehrkräfte als Unterrichtsexperten und -expertinnen adressiert. Es wird die Variation des Unterrichts thematisiert, um die Lernbedingungen an eine heterogene Schülerschaft anzupassen (vgl. ebd., 60). Die Schulleitung bemüht sich um eine operative Vernetzung der Lehrkräfte, indem die über den Fachunterricht hinausgehende Verantwortung für die Lernvoraussetzungen der Schüler und Schülerinnen im Mittelpunkt steht. Das wird besonders in den Fundstellen der Kategorie „Maßnahmen/Entscheidungen" (3 BS) deutlich. Hierbei kann zur Veranschaulichung auf die Sprachförderung (Unterrichtssprache) als schulübergreifendes Ziel im Sinne der Schulentwicklung verwiesen werden: „*Also bei uns ist es sicher im Bewusstsein aller Lehrkräfte, ganz gleich welcher Fachrichtung, verankert, dass der Schwerpunkt der Schule auf Sprachförderung liegt*" (22SL1A, 58). Bezüglich schulübergreifender Zielsetzungen werden kollektiv verbindliche Entscheidungen getroffen.

Auch auf der Ebene der Gremien resp. Schulstrukturen entspricht die Form der Koordination von Schulleitung A einer *gemanagten professionellen Organisation*. Das wird in erster Linie durch die Interviewpassagen aus der Kategorie

7 Die Anzahl der Belegstellen (BS) innerhalb einer Kategorie soll die Relevanz der Aussage im Interview verdeutlichen und erlaubt Rückschlüsse auf die Ausprägung der jeweiligen typenspezifischen Charakteristika.

„Kommunikation" belegt (5 BS). Die Schulleitung delegiert die Verantwortung bei der Auseinandersetzung mit den Lernstandserhebungen an die Fachbereichsleiter und -leiterinnen in den Konferenzen (vgl. ebd., 18), weil sie auf diese Weise die Autonomie dieser Organisationseinheit stärken möchte: *„Eigentlich mische ich mich selber dann in diesen Bereich nicht ein. Da sollen sie auch autonomer sein"* (ebd., 32). Innerhalb der Schule sind ausdifferenzierte Gremienstrukturen in Form von Fachleiter- und Fachbereichskonferenzen sowie Jahrgangskonferenzen, an denen die Schulleitung in regelmäßigen Abständen teilnimmt, vorhanden (vgl. ebd., 20). Eine strategische und operative Vernetzung auf dieser Ebene zeigt sich darin, dass in den (Fach-)Konferenzen Verantwortung für fachdidaktische Entscheidungen getragen wird: *„Wir müssen uns in der Fachkonferenz anschauen, wie gehen wir als Fachgruppe mit diesen Ergebnissen bezogen auf einen Jahrgang um"* (22SL1B, 8). Sie sind verantwortlich für die Verkopplung fachunabhängiger Unterrichtsziele (z.B. Sprachförderung), die bspw. in den Jahrgangskonferenzen festgelegt werden (vgl. 22SL1A, 26), mit den fachspezifischen Unterrichtszielen. Die Koordination der Konferenzen liegt in der Verantwortung der Fachbereichsleitungen, wie das folgende Zitat verdeutlicht: *„Ich versuche, wenn es der Zeitkalender bei mir ermöglicht, auch von Zeit zu Zeit an Fachkonferenzen teilzunehmen. Aber ich denke, man kann das wirklich auch gut delegieren"* (ebd., 20). Durch das Delegieren von Aufgaben, die Partizipation des Kollegiums an verschiedenen Schulentwicklungsprozessen, gemeinsame Zielsetzungen und kollektive Verantwortung im Sinne einer *Distributed Leadership* wird die delegativ-kooperative Führung dieser Schulleitung in der Auseinandersetzung mit Lernstandserhebungen erkennbar.

Das Verhältnis von Schulleitung und Schülerschaft betreffend bieten die Lernstandsergebnisse Informationen (4 BS in Kategorie „Konzeptionelles Verständnis"), um zum einen den Unterricht an die Lernausgangslage der Schüler und Schülerinnen zu adaptieren (vgl. 22SL1B, 106). Zum anderen ermöglichen diese Ergebnisse, den Kompetenzerwerb der Schüler und Schülerinnen an einem unabhängigen Maß zu beurteilen (vgl. 22SL1A, 2), wie im Folgenden deutlich wird: *„Dass man einfach wirklich deutlicher sieht, Schüler A, Schüler B, Schüler C unterscheiden sich in folgenden Kompetenzen. Das wird eben durch solche VERA 8-Erhebungen viel schneller deutlich, als eben durch die regulären Klassenarbeiten"* (22SL1B, 32).

Das Fallbeispiel B verkörpert eine Schulleitung einer *autonomen professionellen Organisation*, d.h. einer verwalteten Schule, die sich insbesondere durch die Zuschreibung von individueller Verantwortung – vor allem auch in der Auseinandersetzung mit Lernstandsergebnissen – auszeichnet. Es findet weder eine Koordination noch Kopplung der Verantwortungsbereiche statt. Als Schulleitung dieses Organisationstyps überträgt sie ihren Lehrkräften nur begrenzt Verantwortung: Sie vermeidet, dass Auswertungen oder Deutungen der Ergebnisse aus Lernstandserhebungen von einzelnen Lehrkräften verantwortet werden. Stattdessen werden die Ergebnisse – im Sinne einer segmentären Differenzierung – verallgemeinert

auf Fachbereichsebene thematisiert. Das belegen die Aussagen aus den Kategorien „Kommunikation" (5 BS) und „Rahmenbedingungen für Verarbeitung" (3 BS).

In diesem Zusammenhang nimmt die Schulleitung gegenüber den Lehrkräften eine gewisse Schutzfunktion ein, indem sie Verständnis dafür zeigt, dass die Einführung der Lernstandserhebungen für die Lehrkräfte einen höheren Mehraufwand bedeutet und zu Überlastung führt (vgl. 31SL1A, 56; 64). Deshalb spricht sie sich auch dafür aus, dass Lernstandsergebnisse zur Schülerbeurteilung eingesetzt werden sollten: *„Insbesondere ist das natürlich ein bisschen ärgerlich, weil die Arbeiten nicht als Pflichtarbeiten gewertet werden dürfen. Das ist also eindeutig so festgelegt. Das ist natürlich ein bisschen ärgerlich, denn letzten Endes bezieht es sich ja auf den Unterricht. Warum eigentlich nun eine Arbeit schreiben, den Lehrern die Arbeit auch machen (BETONT), sie müssen ja korrigiert werden, und dann letzten Endes nicht nutzen dürfen für die Bewertung der Schüler"* (31SL1A, 6).

Auf der Ebene der Gremien resp. Schulstrukturen wird die fragmentarische Verantwortung innerhalb dieser Schule besonders sichtbar. Konferenzen sollen Gelegenheit zum Austausch der Fachlehrkräfte über Lernstandsergebnisse bieten: *„Und andere Klassen wo man dann sagen muss: ‚Also, da müsste in irgendeiner Weise dann etwas passieren'. Das bekommt die Fachkonferenz aufgetischt, dass sie sich über bestimmte Dinge dann noch einmal Gedanken machen müssen"* (ebd., 12). Hieraus resultieren individuelle, lediglich einzelne Fächer oder Klassen betreffende Entscheidungen (3 BS in Kategorie „Kommunikation" und 2 BS in Kategorie „Maßnahmen/Entscheidungen").

Die Auseinandersetzung mit den Lernstandserhebungen in den Fachkonferenzen wird von der Schulleitung zwar erwartet, aber nicht begleitet (2 BS in Kategorie „Kommunikation"). Das bedeutet, dass diese in formaler Separation und Verantwortung der Fachbereichsleiter stattfinden, wie das folgende Zitat veranschaulicht: *„Also, den festgesetzten Termin gibt es dann höchstens in Form der Fachkonferenz. Das plant aber entsprechend der Fachkonferenzleiter. Da wird natürlich auch die Vergleichsarbeit thematisiert. Ich gehe zumindest davon aus, dass das auch wirklich immer im Anschluss geschieht"* (ebd., 28).

Auf Ebene der Schüler und Schülerinnen bilden die Ergebnisse aus Lernstandserhebungen dieser Schulleitung zufolge die Heterogenität zwischen den Klassen ab und unterliegen dabei aufgrund der Klassenzusammensetzung starken Schwankungen, wie im Folgenden deutlich wird: *„Es ist, denke ich mal, auch durchaus gefährlich, weil man Jahrgänge miteinander vergleicht, die nicht unbedingt gleich sind. Und da gibt es sicherlich auch sehr unterschiedlich entwickelte Jahrgänge. Also, man kann daraus dann nicht Rückschlüsse auf die Qualität der eigenen Arbeit schließen. Das glaube ich nicht"* (31SL1B, 51). Aus diesem Grund werden die Lernstandserhebungen in dieser Schule auch nicht als förderdiagnostisches Instrument zur individuel-

len Förderung der Schüler und Schülerinnen eingesetzt. Dies wird besonders in den Fundstellen der Kategorien „Konzeptionelles Verständnis“ (4 BS) und „Rahmenbedingung für Verarbeitung“ (2 BS) deutlich.

4.3 Wie gestalten sich das Verhältnis und die Zusammenarbeit von Schulleitung und Schulaufsicht bei der Auseinandersetzung mit den zurückgemeldeten Lernstandsergebnissen?

Analog zu den vorangegangenen Fallstudien wurden auch auf Ebene der zuständigen Schulaufsicht die Handlungskoordinationen bei VERA analysiert. Im Fall C finden sich Hinweise auf das Konzept des *Distanzierten Managements*. Fall D repräsentiert den Typus des *Neuen Managements*.

VERA und vor allem die daraus resultierenden Rückmeldungen versteht der Typ des *Distanzierten Managements* im Fall C als Möglichkeit der öffentlichen Rechenschaftslegung. Er erwartet in der Folge förderliche Prozesse der Unterrichts- und der Schulentwicklung, wobei die operative Auseinandersetzung in der Verantwortung der Schule liegt (5 BS in Kategorie „Externalisierung“): *„Wir hatten vorher eine ganz starke Bindung an den individuellen Unterricht. Und durch viele Vorgänge – sei es durch VERA 3 oder VERA 8 – wird von außen etwas herangetragen an die Schule und ich denke, das ist förderlich für die Schulentwicklung. Man muss auch die Anforderung von außen bewältigen können. Und diesen Prozess finde ich gut“* (2SA3B, 38).

Als Instrument zur Diagnose und zur Unterrichtsentwicklung sollen aus den standardisierten Vergleichsarbeiten Rückschlüsse auf die Heterogenität in der Leistungsfähigkeit der Schüler und Schülerinnen gezogen und der Unterricht individualisiert und differenziert gestaltet werden. Bei dieser tiefgreifenden Veränderung des Unterrichtsgeschehens verweist der Typ des *Distanzierten Managements* auf die Schulleitung: *„VERA machen ja jetzt die Schulleiter. Und auch dadurch wird das unterstützt, dass wir über die konkreten Konsequenzen nur von Ferne und nur auf Anfrage Bescheid wissen“* (ebd., 26). Die Schulaufsicht selbst sieht sich nicht als zuständig an. Sie beschreibt sich für diesen Prozess als zu fern (vgl. 2SA3B, 48) und überlässt das Instrument den Schulen bzw. den Schulleitungen. Im Fall C externalisiert sie Erwartungen bezüglich der Rezeption und Nutzung von Vergleichsarbeitsdaten deutlich an die Schule bzw. an die Schulleitung. Durch diese Externalisierung handelt die Schulaufsicht gemäß dem *Distanzierten Management*.

Dies korrespondiert mit dem Verständnis von Steuerung und der Rolle der Schulaufsicht im *Distanzierten Management*. Die Schulaufsicht sieht sich im Fall C als Unterstützer bei der von ihr geforderten schulischen Autonomie (3 BS in Kategorie „Unterstützung“). Sie gibt Hilfe zur Selbsthilfe: *„Das passt eigentlich ganz gut zu*

dem, was mein Ideal von Zusammenarbeit ist. Den anderen zu helfen, selbständiger zu werden" (ebd., 8). Durch die bewusste Zurücknahme der eigenen Rolle soll die schulische Autonomie gestärkt werden. So bleibt es der Schulleitung überlassen, die eigenen schulinternen Fortbildungsbedarfe zum Themenkomplex „VERA und deren Rezeption" festzustellen und an die Schulaufsicht heranzutragen. Die Steuerung der Schule durch VERA ist im Fall C nicht von Bedeutung. VERA-Ergebnisse werden von der Schule bürokratisch und formalistisch abgearbeitet (11 BS in Kategorie „Formalistischer Umgang"). Mit der Schulleitung existiert kaum aktiver Austausch über die Ergebnisse, da ein solcher nicht verpflichtend ist (vgl. ebd., 11-12) und aufgrund positiver Lernstandsergebnisse im Zuständigkeitsbereich von Fall C keine Veranlassung dazu gesehen wird (vgl. ebd., 10). Die Schulaufsicht erfährt von den Ergebnissen bzw. den daraus abgeleiteten Maßnahmen mittelbar über Protokolle, Berichte und Zielvereinbarungsgespräche: „*Nur über Berichte bin ich im Bilde über die Konsequenzen an den Schulen. Entweder über Berichte oder Zielvereinbarungen*" (ebd., 17-20).

Aus Sicht der Leitung der Schule, die im Zuständigkeitsbereich des befragten Schulaufsichtsvertreters aus Fallbeispiel C liegt, wird die zurückhaltend angebotene Unterstützung (Fortbildungskoordination, Schulentwicklungsberatung) nicht wahrgenommen und zudem auch nicht erwartet: „*Ich wüsste jetzt nicht, ob die Schulaufsicht in VERA 8 irgendwo steuern oder stützend eingreifen könnte*" (22SL1A, 86). Dennoch besteht auf Seiten der Schulleitung der Wunsch nach einer Koordinationsstelle in der Schulaufsicht, die für VERA zuständig ist und als Ansprechpartner zur Verfügung steht (vgl. ebd., 98). Dies zeigt, dass die Schulleitung trotz der scheinbar souverän ausgeführten, eigenständigen Steuerung der Schule im Kontext zentraler Lernstandserhebungen die Möglichkeit auf Unterstützung befürwortet. Der Schulaufsicht gelingt es nicht, entsprechend den Bedürfnissen der Schulleitung zu handeln und zu kommunizieren. Sie nimmt zwar wahr, dass das Unterstützungsangebot bei VERA von den Schulen nicht angenommen wird, leitet aber keine gegensteuernden Maßnahmen ein.

Das Interdependenzmanagement seitens der Schulaufsicht, welches das Verhältnis zur Ebene der Schulleitung nachhaltig beeinflusst, ist im Fall C somit schwach ausgeprägt. Die Schulaufsicht hat keinen Ansatz entwickelt, um an die Vergleichsarbeiten gestalterisch und schulentwicklungsfördernd anzuschließen. Zudem verwendet die Schulaufsicht die Ergebnisse aus den Lernstandserhebungen nicht strategisch im Sinne von Kontrolle, sondern sie externalisiert Erwartungen an die Schule, sich eigenständig operativ und strategisch mit dem Instrument der VERA auseinanderzusetzen. Die Zusammenarbeit von Schulaufsicht und Schulleitung bei der Auseinandersetzung mit den zurückgemeldeten Lernstandsergebnissen besteht aus Unterstützungsleistungen. Die *gemanagte professionelle Organisation*, die vom befragten Schulleiter vertreten wird, wird durch das *Distanzierte Management* der Schulaufsicht und die Betonung der Eigenverantwortung der Schule befördert. Der Typ des

Distanzierten Managements vermag die bestehenden Unterstützungswünsche (und möglicherweise auch -bedarfe) seitens der Schulleitung aus Fall A bei der strategischen und operativen Handhabung der Testergebnisse nicht abzudecken.

Der im Fall D befragte Schulaufsichtsvertreter zählt VERA als neue Steuerungsform zu seinem originären Aufgabengebiet: „*Und darin hat sich schon die Rolle der Schulaufsicht geändert. Denn ich finde schon wichtig, dass es so eine Draufsicht von Unabhängigen, von außen, bei der Standardvergleiche angesetzt werden, gibt. Und das ist, schon eine veränderte Rolle*" (3SA2B, 12). Er sieht sie als wichtige Stütze der schulaufsichtlichen Arbeit an (4 BS in Kategorie „Zuständigkeit"). Dieses geäußerte Rollenverständnis entspricht dem Organisationskonzept des *Neuen Managements* und ist konträr zu Fall C. Die Schulaufsicht sieht ihre Rolle durch VERA gestärkt, da eine externe Analyse der schulischen Leistungen mittels VERA keine subjektive Beeinflussung zulässt (vgl. ebd., 12) und eine unabhängige Informationsbasis bietet. Bei der Rezeption und Reflexion ist die Schulaufsicht Impulsgeber für diejenigen Schulen, die nicht selbständig Aktionen ableiten können, und steht als Ansprechpartner zur Verfügung. In der Phase der Aktion trifft die Schulaufsicht mit der Schulleitung Abmachungen, welche Maßnahmen einzuleiten sind, um sich in den durch VERA als problematisch erwiesenen Bereichen zu verbessern (vgl. ebd., 31). Dieser Handlungsplan wird zeitlich fixiert und seine Umsetzung und Wirksamkeit entsprechend überprüft. Durch diese Überprüfung führt die Schulaufsicht in Fall D die Prozessphase der Evaluation aus und nimmt die Rolle als übergeordnete Kontrollinstanz wahr: „*Vergleichsarbeiten sind Thema bei Visitationen. Wenn große Auffälligkeiten sind, bin ich natürlich auch aufgefordert, nachzufragen und natürlich auch anzuregen*" (ebd., 4).

Trotz der eigenen Zuständigkeit bei VERA schätzt die Schulaufsicht aus Fall D die Rolle der Schulleitung bei Lernstandserhebungen als zentral ein: „*Also, ich denke, dass der Schulleiter schon eine tragende Rolle in diesem Bereich hat. Also, meiner Auffassung nach, gehört das auch zu seinen Aufgaben, dass er das ganz einfach mit vorantreibt und aus der Analyse heraus dann natürlich auch guckt, wie kann man damit umgehen*" (3SA2B, 33). Für die Schulaufsicht sind Lernstandserhebungen ein Ansatz, mittels datenbasierter Informationen schulische Veränderungen zu initiieren (vgl. ebd., 4). In dieser managementorientierten Perspektive (4 BS in Kategorie „Managementorientierung") geht die Schulaufsicht davon aus, dass die konkrete operative Aufarbeitung der Testdaten an den Schulen per se durchgeführt wird, und verlässt sich dabei auf die Schulleitung: „*Die andere Geschichte, selbstständige Schule, ist da schon wichtiger. Und wenn ich weiß, dass also diese Impulse zur Nutzung der Testdaten durch den Schulleiter gegeben werden, dann ist es natürlich so, dass diese Kompetenz auch abgegeben wird*" (3SA2B, 35). Sie adressiert Erwartungen bezüglich der Aufarbeitung von Vergleichsarbeitsdaten direkt an die Schulleitung (4 BS in Kategorie „Externalisierung"). Diese soll als Impulsgeber fungieren und schulinterne Diskussionen in Gang bringen (vgl. ebd., 33).

Dieser scheinbare Gegensatz zwischen Zuständigkeit und Externalisierung löst sich im Fall D durch das Interdependenzmanagement auf (10 BS in Kategorie „Interdependenzmanagement"), was kennzeichnend für den Typ des *Neuen Managements* ist. Die Ebene der Schulaufsicht führt intensive Kommunikationsprozesse und tauscht sich bei der Aufarbeitung der Vergleichsarbeitsdaten aktiv mit der Ebene der Schulleitung aus: *„Und dass man da in ständigem Kontakt auf Augenhöhe mit den Schulen ist. Man ist jetzt nicht der Vorgesetzte, der nun irgendwas anweist, sondern dass man da einfach im kollektiven Stil eine bestimmte Beratung findet und da auch guckt, wie man also Lösungen herbeiführt"* (ebd., 2). Die Schulaufsicht in Fall D begleitet den Prozess der Rezeption und Reflexion. Sie erarbeitet gemeinsam mit der Schulleitung Verbesserungsmaßnahmen und kontrolliert deren Umsetzung mittels datengestützter Qualitätsgespräche. In allen Phasen steht das kooperative und kommunikative Vorgehen im Mittelpunkt: *„Ich denke schon – so ist zumindest mein Ansatz –, dass man bei schlechten Ergebnissen nicht nur sagt: ‚So, nun berichtet mal, woran das liegen könnte'. Und dann ist das abgeschlossen. Sondern man muss dazu in einen Kommunikationsprozess kommen. Die Schulaufsicht darf nicht gleich sagen, dass irgendetwas falsch ist und dass das irgendwie geändert werden muss. Sondern dass man gemeinsam guckt, was die Gründe dafür sind"* (ebd., 14).

Auch bei der Ausübung von Kontrolle ist im Fall D das Interdependenzmanagement deutlich ausgeprägt. Die Schulaufsicht geht sensibel mit schulischen Einzelfällen um und legt den Schwerpunkt der Überprüfung auf einen bilateralen Austausch mit der Ebene der Schulleitung: *„Und für uns bieten datengestützte Qualitätsgespräche die Möglichkeit, innerhalb der Schule in kleiner Runde Ergebnisse u.a. aus VERA anzusprechen"* (ebd., 25). Dabei prüft sie mögliche Ursachen für das schulische Abschneiden sehr genau und berücksichtigt unterschiedlichste Einflussfaktoren: *„Viele Sonderfälle, die so einfach von den Gesetzmäßigkeiten her nicht zu regeln sind, kläre ich, indem ich auf Austausch und die Beratung und Empfehlung, wie man in den verschiedensten Bereichen dann dort auch unterstützen kann, setze"* (ebd., 2).

Durch das umfassende Interdependenzmanagement steht im Fall D die Schulaufsicht trotz der Externalisierung der Testdaten-Aufarbeitung in engem und stetigem Kontakt mit der Schulleitung, so dass sich das Führungshandeln insgesamt dem Typus des *Neuen Managements* zuordnen lässt. Die Schulaufsicht bietet den Schulen offensiv Unterstützung in der Reflexion der Ergebnisse an. Sie erwartet zudem eine Aufarbeitung der Ergebnisse, die in Zielvereinbarungen – auch zu Kontrollzwecken – festgehalten wird. Die Schule bleibt in der operativen Verantwortung. Da die Schulleitung aus Fall B das Konzept der *autonomen professionellen Organisation* verfolgt, welches eigenständige Entscheidungen betont und eine ebenenübergreifende Kommunikation gering gewichtet, werden Kommunikationsprozesse auf Schulebene hinter einer *„ceremonial facade"* ausgeführt. Handlungen im Sinne von Schulentwicklungsprozessen, die als Ziel der Gespräche erarbeitet worden sind, bleiben weitgehend aus. Dies vermag die Schulaufsicht – möglicherweise auch bedingt

durch die langen zeitlichen Abstände der datengestützten Qualitätsgespräche – nicht zu erfassen.

5. Zusammenfassung und Diskussion

Zusammenfassend ist mit den vorliegenden Befunden zum einen die zentrale Rolle und Funktion der Schulleitung als maßgeblicher Akteur evidenzbasierter Steuerung für die Auseinandersetzung mit Lernstandsergebnissen deutlich geworden (Forschungsfrage 1). Zum anderen konnten divergierende Ansatzpunkte und Einstellungen zur Neustrukturierung der Koordination zwischen Schulleitung und Schulaufsicht aufgezeigt werden, die teilweise mit den vorgestellten, idealtypischen Organisationsmodellen korrespondieren (Forschungsfrage 2).

Theoretisch lässt sich ableiten, dass die Nutzung von Lernstandserhebungen auf Meso- und Mikroebene in einer *gemanagten professionellen Organisation* produktiv und wirkmächtig erfolgen kann. Dies wird auf Makroebene, der Ebene der Schulaufsicht, durch den Typus des *Neuen Managements* unterstützt. Gleichzeitig verstärkt dieser Typus die *Distributed-Leadership*-Praxis in der schulischen Organisation. Im Kontrast dazu zeigt sich, dass die Rezeption und Nutzung von Vergleichsarbeitsdaten behindert wird, wenn die Makroebene nach dem Typ des *Distanzierten Managements* handelt und dabei auf Meso- und Mikroebene auf eine *autonome professionelle Organisation* trifft.

Die empirisch dargestellten Ausgangspunkte der Koordination lassen unterschiedliche Muster der Selbst- und Fremdzuschreibung von Verantwortung für Folgeprozesse der Ergebnisrückmeldung erkennen. Sie betreffen maßgeblich die Positionierung der Schulleitungen in der Mehrebenenorganisation des Schulsystems und ihre Handlungsoptionen für die Durchsetzung der testbasierten Schulreform. Lehrkräfte werden von den Schulleitungen durchgängig als Unterrichtsexperten und -expertinnen adressiert; die damit verbundene Verantwortung für das operative Erreichen schulischer Ziele findet jedoch unterschiedliche Zuschreibungen.

Diese Differenz spiegelt sich in der Inanspruchnahme von Gremien bzw. Schulstrukturen, die als Infrastruktur der Wahrnehmung kollektiver Verantwortung fungieren, durch die Schulleitungen wider. Es werden zwei konträre Handlungsmuster deutlich: Die Gremienarbeit wird für die Formulierung gemeinsamer Zielsetzungen (bspw. gesamtschulisches Querschnittsthema „Sprachförderung“) genutzt, auf die – im Sinne einer *Distributed Leadership* – arbeitsteilig hingearbeitet wird. Auf diese Weise vernetzt die Schulleitung die verschiedenen Ebenen und Verantwortungszentren innerhalb der Schule; die Verwendung von Lernstandsergebnissen zur Verbesserung des Unterrichts erfolgt in diesem Netzwerk. Demgegenüber stehen Schulleitungen,

die Gremien als in sich abgeschlossene Expertenzirkel wahrnehmen. Sie dienen der Delegation individueller, auf Schulebene nicht vernetzter Verantwortung.

Schulleitungen thematisieren Schülerleistungen sowohl als Voraussetzung als auch als Produkt effizienten Unterrichtens, wobei folgende Handlungskonzepte in Kontrast zueinander stehen: Schülerleistungen werden kausal auf einzelne Unterrichtsmerkmale zurückgeführt, Varianzen im Kompetenzzuwachs der Ausgestaltung von Unterrichtsprozessen zugeschrieben. Dieses Handlungskonzept ist mit einer ausgeprägten Bereitschaft zur Reflexion des Unterrichts verbunden. Im Gegensatz dazu existiert auch ein eindimensionales Konzept des „guten" Unterrichts. Hier führt die Schulleitung die Kompetenzunterschiede auf die Heterogenität der Schülergruppen zurück. Die Kausalattributionen werden auf Schülermerkmale bezogen und die Lernstandsergebnisse als kollektive Schülerleistung, die auf Klassenebene aggregiert wurde, aufgefasst – eine Reflektion des Unterrichts bleibt in der Folge aus.

Die Handlungskoordination in der Verantwortung von Schulleitungen spiegelt institutionelle Normen wider, die sich im Verhältnis zwischen Schulleitungen und Schulaufsichten etablieren. Das Führungshandeln der Schulleitungen variiert zwischen einem proaktiv gestaltenden Muster, bei dem die Schule Verantwortung für schulübergreifende Kriterien (bspw. bei Konkurrenz zu anderen Schulen) trägt, und einem reaktiv isolierenden Muster. Beim letzteren erfüllt die Schule formale Erfordernisse und schottet sich in der Verantwortung vor internen Kriterien im Sinne einer *„ceremonial facade"* ab. Dabei wird die innerschulische Aktivitätsstruktur nicht verändert. Stattdessen adressiert die Schulleitung die Schulaufsicht mit einer Unterstützungserwartung. Das entspricht nach dem neo-institutionalistischen Ansatz dem Entkopplungsmechanismus.

In den Fallbeispielen zeigt sich jedoch, dass die Schulaufsicht diesen Erwartungen nicht nachkommt bzw. nicht in genügendem Ausmaß mit den Schulen über Beratungs- und Unterstützungsangebote kommuniziert. Wenn die Schulaufsicht gemäß dem *Distanzierten Management* handelt, wird deutlich, dass das Interdependenzmanagement nicht ausgeprägt ist. In Kombination von *Distanziertem Management* und *gemanagter professioneller Organisation*, wie bei Fall A und C, bedeutet dies, dass die schulischen Unterstützungserwartungen unerfüllt bleiben. Durch die Betonung der schulischen Autonomie seitens der Schulaufsicht erfährt die Schulleitung jedoch eine Stärkung ihrer Rolle. Die Überlassung der Verantwortung bei VERA ermöglicht ihr, hierarchisch und vernetzt agieren zu können.

Handelt die Schulaufsicht nach dem Konzept des *Neuen Managements* und die Schulleitung gemäß einer *autonomen professionellen Organisation* (Fallbeispiele B und D), dann versanden die Kommunikationsbemühungen seitens der Schulaufsicht. Die Schulleitung handelt eigenständig und führt lediglich die notwendigen Gespräche mit der Schulaufsicht. Die Formalstruktur wird angepasst, um Legitimität zu wahren. Die

tatsächlichen Handlungen, d.h. die Aktivitätsstruktur, bleiben unverändert. Auch in diesem Fall tritt der bereits erwähnte Entkopplungsprozess ein.

Die Fallstudien zeigen einen erheblichen Entwicklungsbedarf für einen sachkundigen Umgang mit den Lernstandserhebungen in der Koordination zwischen den Ebenen des Schulsystems. Angesichts der noch recht jungen Erfahrungen wäre es verfrüht, von elaborierten Konzepten auszugehen, wie sie etwa in gezielten Organisations- und Personalentwicklungsmaßnahmen im Schulsystem zu leisten wären. Bemerkenswert ist jedoch, dass bereits die hier geschilderten Ansätze zur Koordination der Verarbeitung von Lernstandserhebungen sehr unterschiedliche Gelingensbedingungen kenntlich werden lassen. Maßgebliche Differenzen bestehen in den geschilderten Fällen hinsichtlich der Frage, welche Instanz in der Mehrebenenstruktur des Schulsystems Verantwortung für die Initiation von Entwicklungsprozessen und ihre Unterstützung trägt. Ebenso werden variierende Interpretationen der Grenzziehungen zwischen operativer und strategischer Verantwortung im Schulsystem deutlich. Die Implementation der Lernstandserhebungen ruft nicht nur Fragen einer inhaltlich (fach- und allgemeindidaktisch) und methodisch (Verarbeitung der aufgearbeiteten statistischen Informationen) tragfähigen Rezeption bzw. Verwendung der Ergebnisse hervor; darüber hinaus macht sie auch auf offene Aspekte der Entwicklung des Schulsystems in einem umfassenderen Sinne aufmerksam. Zu ihnen zählen die institutionelle Einbettung der einzelnen Schule, die Verteilung operativer und strategischer Verantwortung und – im Knotenpunkt dieser Fragen – die Leitung der Schule.

Literatur und Internetquellen

Altrichter, H./Brüsemeister, T./Wissinger, J. (2007): Einführung. In: Altrichter, H./Brüsemeister, T./Wissinger, J. (Hrsg.): Educational Governance. Handlungskoordination und Steuerung im Bildungssystem, Bd. 1. Wiesbaden: VS, S. 9-13.

Böttcher, W. (2002): Kann eine ökonomische Schule auch eine pädagogische sein? Schulentwicklung zwischen Neuer Steuerung, Organisation, Leistungsevaluation und Bildung. Weinheim/München: Juventa.

Bonsen, M. (2003): Schule, Führung, Organisation. Eine empirische Studie zum Organisations- und Führungsverständnis von Schulleiterinnen und Schulleitern. Münster u.a.: Waxmann.

Bonsen, M. (2011): Von der „failing" zur „turnaround" school. Was Schulleitungen aus der Forschung zur erfolgreichen Umsteuerung von Organisationen lernen können. In: Lernende Schule 14, H. 56, S. 12-15.

Bonsen, M./Pfeiffer, H. (1998): Schulleitungsforschung in Deutschland: Forschung und Ergebnisse im Überblick. In: Buchen, H./Horster, L./Rolff, H.-G (Hrsg.): Schulleitung und Schulentwicklung. Erfahrungen, Konzepte und Strategien. Berlin: Raabe, S. 2-18.

Brauckmann, S. (2012): Schulleitungshandeln zwischen deconcentration, devolution und delegation (3D) – empirische Annäherungen aus internationaler Perspektive. In: Empirische Pädagogik 26, H. 1, S. 78-102.

Brauckmann, S./Herrmann, C. (2012): Schulleitungshandeln im Rahmen neuer Steuerung – Belastung oder Chance?! In: Zeitschrift für Bildungsverwaltung 28, H. 1, S. 87-98.

Brüsemeister, T./Newiadomsky, M. (2008): Schulverwaltung – Ein unbekannter Akteur? In: Langer, R. (Hrsg.): „Warum tun die das?" Governanceanalysen zum Steuerungshandeln in der Schulentwicklung. Wiesbaden: VS, S. 73-93.

Buchen, H./Rolff, H.-G. (Hrsg.) ([2]2009): Professionswissen Schulleitung. Weinheim: Beltz.

Burkard, C./Rolff, H.-G. (1994): Steuerleute auf neuem Kurs? Funktionen und Perspektiven der Schulaufsicht für die Schulentwicklung. In: Rolff, H.-G./Bauer, K.-O./Klemm, K./Pfeiffer, H./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung, Bd. 8. Daten, Beispiele und Perspektiven. Weinheim: Beltz, S. 205-265.

Campbell, C./Gross, B. (2012): Principal Concerns: Leadership Data and Strategies for States. Washington, DC. URL: http://www.crpe.org/sites/default/files/pub_principal_concerns_sept12.pdf; Zugriffsdatum: 22.05.2013.

Cohen, J. (1960): A Coefficient of Agreement for Nominal Scales. In: Educational and Psychological Measurement 20, H. 1, S. 37-46.

Dubs, Rolf (1996): Schule, Schulentwicklung und New Public Management. St. Gallen: Institut für Wirtschaftspädagogik.

EMSE – Netzwerk Empiriegestützte Schulentwicklung (2008): Positionspapier Zentrale standardisierte Lernstandserhebungen. Berlin. URL: http://www.emse-netzwerk.de/uploads/Main/EMSE_Positionsp1_Lernstandserh.pdf; Zugriffsdatum: 05.03.2013.

Fletcher, J.K./Käufer, K. (2003): Shared Leadership: Paradox and Possibility. In: Pearce, C.L./Conger, J.A. (Hrsg.): Shared Leadership: Reframing the Hows and Whys of Leadership. Thousand Oaks, CA: Sage, S. 21-47.

Flick, U./Garms-Homolová, V./Herrmann, W.J./Kuck, J./Röhnsch, G. (2012): "I Can't Prescribe Something Just Because Someone Asks for It …": Using Mixed Methods in the Framework of Triangulation. In: Journal of Mixed Methods Research 6, H. 2, S. 97-110.

Füssel, H.-P./Leschinsky, A. (2008): Der institutionelle Rahmen des Bildungswesens. In: Cortina, K.S. (Hrsg.): Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick. Reinbek: Rowohlt, S. 130-203.

Glaser, B./Strauss, A. ([3]2010): Grounded theory: Strategien qualitativer Forschung. Bern: Verlag Hans Huber.

Gronn, P. (2002): Distributed Leadership. In: Leithwood, K./Hallinger, P. (Hrsg.): Second International Handbook of Educational Leadership and Administration. Dordrecht u.a.: Kluwer, S. 653-696.

Groß Ophoff, J./Hosenfeld, I./Koch, U. (2007): Formen der Ergebnisrezeption und damit verbundene Schul- und Unterrichtsentwicklung. In: Empirische Pädagogik 21, H. 4, S. 411-427.

Grouven, U./Bender, R./Ziegler, A./Lange, S. (2007): Der Kappa-Koeffizient. In: Deutsche medizinische Wochenschrift (DMW) 132, H. 1, S. 65-68.

Harazd, B./Gieske, M./Rolff, H.-G. (2008): Herausforderungen an Schulleitung: Verteilung von Verantwortung und Aufgaben. In: Bos, W./Holtappels, H.G./Pfeiffer, H./Rolff, H.-G./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung, Bd. 15. Weinheim/München: Juventa, S. 225-256.

Harris, A. (2008): Distributed Leadership: According to the Evidence. In: Journal of Educational Administration 46, H. 2, S. 172-188.

Hartung-Beck, V. (2009): Schulische Organisationsentwicklung und Professionalisierung. Folgen von Lernstandserhebungen an Gesamtschulen. Wiesbaden: VS.

Heimerl, P. (2007): Fallstudien als forschungsstrategische Entscheidung. In: Buber, R./Holzmüller, H. (Hrsg.): Qualitative Marktforschung. Konzepte – Methoden – Analysen. Wiesbaden: Gabler, S. 381-400.

Helmke, A. (2004): Von der Evaluation zur Innovation: Pädagogische Nutzbarmachung von Vergleichsarbeiten in der Grundschule. In: Seminar 10, H. 2, S. 90-112.

Klieme, E. (2004): Begründung, Implementation und Wirkung von Bildungsstandards: Aktuelle Diskussionslinien und empirische Befunde. In: Zeitschrift für Pädagogik 50, H. 5, S. 625-634.

Koch, U. (2011): Verstehen Lehrkräfte Rückmeldungen aus Vergleichsarbeiten? Datenkompetenz von Lehrkräften und die Nutzung von Ergebnisrückmeldungen aus Vergleichsarbeiten. (Empirische Erziehungswissenschaft.) Münster u.a.: Waxmann.

Kuper, H. (2002): Entscheidungsstrukturen in Schulen. Eine differenzielle Analyse der Schulorganisation. In: Zeitschrift für Pädagogik 48, H. 6, S. 856-878.

Lange, H. (2003): Schulaufsicht zwischen normativen Anforderungen und faktischen Wirkungsmöglichkeiten. In: Zeitschrift für Pädagogik, Beiheft 47, S. 137-155.

Laukart, J. (2009): Funktionen staatlicher Schulaufsicht im Kontext neuer Steuerungstheorien. Am Beispiel der unteren Schulaufsicht in Hessen. Frankfurt a.M.: Goethe Universität Frankfurt am Main.

Leithwood, K./Louis, K.S./Anderson, S./Wahlstrom, K. (2004): How Leadership Influences Student Learning. URL: http://www.wallacefoundation.org/knowledge-center/school-leadership/key-research/Documents/How-Leadership-Influences-Student-Learning.pdf; Zugriffsdatum: 29.12.2012.

Maier, U. (2010): Effekte testbasierter Rechenschaftslegung auf Schule und Unterricht: Ist die internationale Befundlage auf Vergleichsarbeiten im deutschsprachigen Raum übertragbar? In: Zeitschrift für Pädagogik 56, H. 1, S. 112-128.

Maier, U./Kuper, H. (2012): Vergleichsarbeiten als Instrument der Qualitätsentwicklung an Schulen. Überblick zum Forschungsstand. In: Die Deutsche Schule 104, H. 1, S. 88-99.

Mayring, P. (2010): Qualitative Inhaltsanalyse. In: Mey, G./Mruck, K. (Hrsg.): Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS, S. 601-613.

Mortimore, P./Sammons, P./Stoll, L./Lewis, D./Ecob, R. (1988): School Matters. Wells, ME: Open Books.

Rosenbusch, H.S. (2005): Organisationspädagogik der Schule. Grundlagen pädagogischen Führungshandelns. Neuwied: Luchterhand.

Sammons, P./Hillman, J./Mortimore, P. (1995): Key Characteristics of Effective Schools. A Review of School Effectiveness Research. London: OFSTED.

Scheerens, J. (2012): School Leadership Effects Revisited. Review and Meta-Analysis of Empirical Studies. Heidelberg/London/New York, NY: Springer.

Schleicher, A. (2009): Moderne Schulleitung im Wandel. Schlussfolgerungen aus OECD-Analysen. In: Die Deutsche Schule 101, H. 4, S. 311-322.

Schnell, H. (2006): Entwicklung und Perspektiven der Schulaufsicht als Steuerungsinstrument für Schulentwicklung in Hessen seit 1945. Dissertation. Kassel.

Spillane, J.P./Halverson R./Diamond J.B. (2001): Investigating School Leadership Practice: A Distributed Perspective. In: Educational Researcher 30, H. 23, S. 23-28.

Supovitz, J./Sirinides, P./May, H. (2010): How Principles and Peers Influence Teaching and Learning. In: Educational Administration Quarterly 46, H. 1, S. 31-56.

Terhart, E. (1997): Schulleitungshandeln zwischen Organisation und Erziehung. In: Wissinger, J. (Hrsg.): Schulleitung als pädagogisches Handeln. München: Oldenbourg, S. 7-20.

Thiel, F. (2008): Organisationssoziologische Vorarbeiten zu einer Theorie der Schulentwicklung. In: Journal für Schulentwicklung 12, H. 2, S. 31-39.

Tillmann, K.-J./Dedering, K./Kneuper, D./Kuhlmann, C./Nessel, I. (2008): PISA als bildungspolitisches Ereignis. Oder: Wie weit trägt das Konzept der „evaluationsbasierten Steuerung"? In: Brüsemeister, T./Euble, K.-D. (Hrsg.): Evaluation, Wissen, Nichtwissen. Wiesbaden: VS, S. 117-140.

Visscher, A.J./Coe, R. (2003): School Performance Feedback Systems. Conceptualisation, Analysis, and Reflection. In: School Effectiveness and School Improvement 14, H. 3, S. 321-349.

Warwas, J. (2009a): Schulleitung im Zeitalter der Beschleunigung. Reichen die zeitlichen Ressourcen für die Aufgabenerfüllung aus? In: SchulVerwaltung. Hessen/Rheinland-Pfalz 14, H. 3, S. 87-89.

Warwas, J. (2009b): Berufliches Selbstverständnis und Beanspruchung in der Schulleitung. Paralleltitel: Professional Self-Concept and the Demands on School Heads. In: Zeitschrift für Erziehungswissenschaft 12, H. 3, S. 476-498.

Weitzel, U. (2002): Berufsbild und Rollenverständnis der Schulaufsicht. In: Rolff, H.-G./Schmidt, H.-J. (Hrsg.): Brennpunkt Schulleitung und Schulaufsicht. Konzepte und Anregungen für die Praxis. Neuwied: Luchterhand, S. 323-335.

Wirtz, M./Caspar, F. (2002): Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Göttingen: Hogrefe.

Wissinger, J. (2011): Schulleitung und Schulleitungshandeln. In: Terhart, E./Bennewitz, H./Rothland, M. (Hrsg.): Handbuch der Forschung zum Lehrerberuf. Münster u.a.: Waxmann, S. 98-115.

Yin, R.K. (^{4}2009): Case Study Research. Design and Methods. Los Angeles, CA: Sage.

Harm Kuper, Prof. Dr., geb. 1966, Lehrstuhlinhaber im Arbeitsbereich Weiterbildung und Bildungsmanagement der Freien Universität Berlin.
E-Mail: harm.kuper@fu-berlin.de

Barbara Muslic, Dipl.-Päd., M.A., geb. 1982, wissenschaftliche Mitarbeiterin im Arbeitsbereich Weiterbildung und Bildungsmanagement der Freien Universität Berlin.
E-Mail: barbara.muslic@fu-berlin.de

Anschrift: Freie Universität Berlin, Arbeitsbereich Weiterbildung und Bildungsmanagement, Arnimallee 12, 14195 Berlin

Carolin Ramsteck, M.A., geb. 1982, wissenschaftliche Mitarbeiterin am Institut für Erziehungswissenschaft, Abteilung Schulpädagogik/Empirische Unterrichtsforschung der Pädagogischen Hochschule Schwäbisch Gmünd.
E-Mail: carolin.ramsteck@ph-gmuend.de

Anschrift: Pädagogische Hochschule Schwäbisch Gmünd | University of Education, Institut für Erziehungswissenschaft, Oberbettringer Str. 200, 73525 Schwäbisch Gmünd

SCHULINSPEKTIONSVERFAHREN

DDS – Die Deutsche Schule
Beiheft 12, S. 121-153
© 2013 Waxmann

Oliver Böhm-Kasper/Odette Selders

„Schulinspektionen sollten regelmäßig durchgeführt werden"[1]?

Ländervergleichende Analyse der Wahrnehmung und Akzeptanz von Schulinspektionsverfahren

Zusammenfassung
Im Fokus des Beitrages steht die Frage, ob länderspezifische Variationen der Schulinspektion mit entsprechend unterschiedlicher Akzeptanz und divergierenden Verarbeitungsstrategien von Schulinspektionsimpulsen auf Ebene der Einzelschule einhergehen. Im Rahmen einer standardisierten Befragung von rund 1.470 Schulleitungen in den vier Bundesländern Baden-Württemberg, Mecklenburg-Vorpommern, Niedersachsen und Thüringen zu Beginn des Jahres 2012 wurden Merkmale der Schulinspektion sowie die damit verbundene Wahrnehmung und Akzeptanz der Schulleitungen erfasst. Die Operationalisierung dieses Zusammenhangs orientiert sich an dem von Ehren und Visscher (2006) vorgelegten Rahmenmodell zur Wirkung von Schulinspektionen.
Schlüsselwörter: Schulinspektion, Schulleiterbefragung, Rahmenmodell zur Wirkung von Schulinspektionen

"*School Inspections Should Be Conducted Regularly*"?

Comparing the Perception and Acceptance of School Inspection Procedures across German Federal States

Summary
This article focuses on the question if variations in the inspection procedure come along with differing acceptance and processing strategies for school inspection impulses on the individual school level. By the use of a standardized survey from about 1,470 headmas-

1 Aussage aus einem Schulleiterinterview.

ters in the four federal states Baden-Württemberg, Mecklenburg-West Pomerania, Lower Saxony and Thuringia at the beginning of 2012, characteristics of the school inspection as well as the corresponding perception and acceptance by headmasters were measured. The operationalization of this correlation is based upon the framework of school inspection effects by Ehren and Visscher (2006).
Keywords: school inspection, headmaster survey, framework of school inspection effects

1. Einleitung

Die im Jahr 2004 in Deutschland eingeführte externe Evaluation von Schulen in Form der Schulinspektion hat eine übereinstimmende Grundstruktur, aber auch länderspezifische Ausprägungen. Diese zeigen sich bereits in einer unterschiedlichen Bezeichnung des Steuerungsinstrumentes, das als „Schulinspektion", „externe Evaluation", „Fremdevaluation", „Qualitätsanalyse" oder „Schulvisitation" bezeichnet wird. Die Gemeinsamkeiten der Schulinspektionskonzepte sind nach Kotthoff und Böttcher (2010) in den folgenden Merkmalen verankert:

Schulinspektion findet als Evaluation von Einzelschulen statt. Die daraus resultierenden Ergebnisse werden zunächst an die Einzelschulen mit der Intention zurückgemeldet, dass die Evaluationsergebnisse einen Impuls für Schulentwicklungsprozesse beinhalten. Zum Teil werden auch weitere Akteure des schulischen Steuerungssystems über die Ergebnisse unterrichtet und somit ergänzende Informationen für ein umfassendes, evidenzbasiertes schulisches Qualitätssicherungssystem bereitgestellt.

Weiterhin zeichnen sich die Schulinspektionsverfahren durch Datenbasierung und Standardisierung (wenn auch in unterschiedlicher Ausprägung) aus. Damit soll ein Vergleich der Evaluationsergebnisse zwischen Einzelschulen bzw. ein Abgleich mit vorab formulierten Normen ermöglicht werden.

Schließlich basieren die Bezugsnormen in der Regel auf einem öffentlich vorliegenden Orientierungsrahmen für Schulqualität, der normative Erwartungen an die Qualität von Schule und Unterricht formuliert und damit Schulen einer Rechenschaftslegung unterwirft (vgl. ebd., S. 295f.).

Trotz dieser Gemeinsamkeiten der Schulinspektionsverfahren lassen sich durchaus länderspezifische Variationen erkennen, die in der vorliegenden Literatur unterschiedlich systematisiert werden. Lassen sich bei Rürup (2008) mit einem Fokus auf die Wissensproduktion durch Schulinspektion wahrheits- und angemessenheitsorientierte Ansätze unterscheiden, so werden bei Kotthoff und Böttcher (2010) idealtypische Wirkungsmechanismen anhand der Formeln Qualitäts-/Schulentwicklung über Wettbewerb, über Konsequenzen oder über Einsicht unterschieden. Diese Systematisierungsversuche zeigen, dass trotz der dargestellten Gemeinsamkeiten bun-

desländerspezifische Ausprägungen von Schulinspektionsverfahren bestehen, deren Auswirkungen auf die Wahrnehmung und Akzeptanz bei Schulleitungen im Fokus des vorliegenden Beitrages stehen.

2. Gemeinsamkeiten und Unterschiede der Schulinspektionsverfahren in Baden-Württemberg, Mecklenburg-Vorpommern, Niedersachsen und Thüringen

Die nachfolgenden Ausführungen orientieren sich an den entsprechenden Verlautbarungen der jeweiligen Ministerien, Behörden und Institute der Bundesländer, wie sie auf ihren Webseiten zur Verfügung gestellt werden (Landesbildungsserver Baden-Württemberg, URL: http://www.schule-bw.de; Bildungsserver Mecklenburg-Vorpommern, URL: http://www.bildung-mv.de; Niedersächsischer Bildungsserver, URL: http://www.nibis.de; Thüringer Schulportal, URL: http://www.schulportal-thueringen.de).

Baden-Württemberg

Das in Baden-Württemberg *Fremdevaluation* genannte Inspektionsverfahren hat zum Ziel, der Schule eine externe, kriteriengeleitete und von der Schulaufsicht unabhängige Rückmeldung zum System Schule zu geben. Grundlage dazu bildet der Qualitätsrahmen zur Fremdevaluation in Baden-Württemberg, der für alle allgemein bildenden Schulen gültig ist. Dabei finden aber auch schulspezifische Ausprägungen in den gemessenen Indikatoren Berücksichtigung. Mit dem Verfahren sollen Impulse für die Qualitätsentwicklung gesetzt werden, indem sowohl Stärken als auch Verbesserungsbereiche aufgedeckt werden.

Den Schulen werden etwa zwei Jahre im Voraus das Schuljahr und mehrere Monate im Voraus der exakte Zeitpunkt der Fremdevaluation mitgeteilt. Ein Evaluationsteam besteht aus zwei bis drei eigens dafür qualifizierten Lehrkräften, die zeitlich befristet an das Landesinstitut für Schulentwicklung abgeordnet sind. Außerdem haben Schulen die Möglichkeit, eine assoziierte Person nach Wunsch zu benennen, die das Evaluationsteam ergänzt.

Das Verfahren besteht aus einem Vorgespräch mit der Schulleitung, einer Onlinebefragung, einer Dokumentenanalyse, einer zwei bis drei Tage dauernden Erhebung vor Ort und einer Rückmeldung der Evaluationsergebnisse. Die Rückmeldung findet dabei gestuft statt: Es gibt zuerst einen vorläufigen Bericht an die Schulleitung, den diese prüft; darauf folgt eine mündliche Präsentation der Ergebnisse im Rahmen einer Gesamtlehrerkonferenz; zuletzt erhält die Schule den endgültigen Bericht.

Nach Abschluss des Fremdevaluationsverfahrens und Erhalt des endgültigen Berichts ist die Schule verpflichtet, den Bericht der Schulaufsicht und dem Schulträger zu übermitteln und innerhalb der Schulgremien bekannt zu machen. Des Weiteren leitet die Schule aus dem Bericht Zielvorstellungen und Maßnahmen zur Schulentwicklung ab, legt diese der Schulaufsicht vor und trifft mit ihr verbindliche Zielvereinbarungen.

Mecklenburg-Vorpommern

Die *externe Evaluation* in Mecklenburg-Vorpommern bewertet die Qualität von Schule, Unterricht und Erziehung anhand von definierten Qualitätsmerkmalen und Leistungsdaten, welche sich aus dem „Referenzrahmen Schulqualität" generieren. Zielsetzung ist die Unterstützung der schulischen Qualitätsentwicklung durch das Aufzeigen von Stärken und Schwächen.

Etwa acht Wochen vor dem Schulbesuch erhält die Schule eine Ankündigung und Informationen über die externe Evaluation. Die speziell fortgebildeten Evaluationsteams bestehen je aus drei Mitgliedern: einem Vertreter oder einer Vertreterin des Dezernats für Qualitätssicherung im Landesinstitut für Schule und Ausbildung, einem Schulrat oder einer Schulrätin, der bzw. die nicht für die zu evaluierende Schule zuständig ist, und einem Schulleiter bzw. einer Schulleiterin einer anderen Schule.

Nach einer Vorabinformation und Vorstellung des Evaluationsteams an der Schule führt dieses eine Analyse der durch die Schule bereitgestellten Dokumente durch. Daran anschließend findet der zwei bis drei Tage dauernde Schulbesuch statt. Ca. drei Wochen später erhält die Schule den schriftlichen Evaluationsbericht, zu dem sie eine Stellungnahme abgeben kann, die ggf. an den Bericht angehängt wird. Der Bericht geht auch an die Schulaufsicht und das Dezernat für Qualitätssicherung im Landesinstitut für Schule und Ausbildung.

Die Schule übermittelt die Ergebnisse der externen Evaluation an die Schulkonferenz, welche sich mit den gewonnenen Ergebnissen auseinandersetzt und Vorschläge für konkrete Entwicklungsmaßnahmen generiert. Gemeinsam mit der Schulaufsicht werden anschließend verbindliche Zielvereinbarungen getroffen. Schulen haben im Verlauf des Qualitätsentwicklungsprozesses auch die Möglichkeit, Unterstützungssysteme einzufordern.

Niedersachsen

Zielsetzung der *Schulinspektion* in Niedersachsen ist der Gewinn detaillierter Kenntnisse über die Qualität der einzelnen Schulen und des Schulsystems insgesamt. Die zugrunde liegenden standardisierten Bewertungskriterien sind aus

dem Orientierungsrahmen „Schulqualität in Niedersachsen" abgeleitet, wobei aber auch das soziale Umfeld und die besonderen Rahmenbedingungen der jeweiligen Einzelschule beachtet und erfasst werden. Mit Hilfe der Ergebnisse sollen systematisch Maßnahmen der Qualitätsverbesserung entwickelt und durchgeführt werden.

Die Schulen werden ca. sechs bis acht Wochen im Voraus über den Termin der Schulinspektion informiert. In der Regel besteht das Inspektionsteam aus zwei Inspektoren bzw. Inspektorinnen, die durch die Niedersächsische Schulinspektion kontrolliert werden. Mindestens ein Inspektor/eine Inspektorin muss eine Lehramtsbefähigung besitzen und über umfangreiche Unterrichtserfahrung an der Schulform der zu inspizierenden Schule verfügen.

Die Schulinspektion startet mit einer schulöffentlichen Informationsveranstaltung durch die Niedersächsische Schulinspektion. Danach füllt die Schulleitung einen umfangreichen Erhebungsbogen aus und die Schule reicht die angeforderten Dokumente für die anschließende Analyse ein. Es folgt der, abhängig von der Schulgröße, drei bis fünf Tage dauernde Schulbesuch. Auf Basis einer vorläufigen Auswertung der gesammelten Daten vor Ort erhält die Schulleitung eine erste ausführliche mündliche Rückmeldung und die Schulöffentlichkeit eine Zusammenfassung der vorläufigen Ergebnisse. Etwa drei Wochen nach dem Schulbesuch erhält die Schulleitung einen vorläufigen schriftlichen Inspektionsbericht, zu dem sie eine Stellungnahme abgeben kann, die ggf. dem endgültigen Inspektionsbericht beigefügt wird. Ist die Schulinspektion inklusive Rückmeldung beendet, erhalten auch der Schulträger und die Landesschulbehörde den Inspektionsbericht. Der Bericht stellt zwar Handlungsbedarfe fest, gibt aber keine Handlungsempfehlungen.

Werden bei einer Schule gravierende Mängel festgestellt, wird eine *Nachinspektion* festgelegt, bei der eine Konzentration auf die ermittelten Defizitbereiche erfolgt. Falls eine Nachinspektion durch die Niedersächsische Schulinspektion angesetzt wird, stimmt sich die Landesschulbehörde mit der Schule über Verbesserungs- und Unterstützungsmaßnahmen ab. Dazu wird eine Vereinbarung mit der Schule getroffen, in der Ziele, konkrete Maßnahmen, eine Zeitplanung sowie Unterstützungsleistungen (z.B. Beratungs- und Fortbildungsangebote, Bereitstellung von Ressourcen) geregelt werden.

Thüringen

Ziel der *externen Evaluation* in Thüringen ist die systematische Entwicklung und Sicherung der Qualität schulischer Arbeit. Anhand der Evaluationsergebnisse werden mit dem Staatlichen Schulamt Entwicklungsziele im Rahmen abrechenbarer Zielvereinbarungen determiniert. Durch die regelmäßige Rechenschaftslegung der Schulen soll eine transparente Gestaltung der Entwicklungsprozesse forciert werden.

In dem dem Schulbesuch vorausgehenden Schuljahr erhalten die Schulen vom Expertenteam erste Informationen und die Termine für die externe Evaluation werden festgelegt. Das von der Koordinierungsstelle „Eigenverantwortliche Schule" des Thüringer Ministeriums für Bildung, Wissenschaft und Kultur beauftragte Expertenteam besteht in der Regel aus drei besonders qualifizierten Mitgliedern. Die Experten bzw. Expertinnen arbeiten in der Schulleitung, einem staatlichen Schulamt oder sind als Lehrkräfte mit besonderen Aufgaben tätig.

Wenn die Schulen vom Expertenteam kontaktiert wurden und die geforderten Dokumente eingereicht worden sind, findet zwischen beiden ca. vier Wochen vor dem Schulbesuch ein Vorbereitungsgespräch statt. Der Schulbesuch selbst dauert, abhängig von der Größe der Schule, drei bis vier Tage. Spätestens vier Wochen danach gibt es eine Rückmeldeveranstaltung, in der das Expertenteam die Evaluationsergebnisse präsentiert. Die Schulleitung erhält dazu auch den schriftlichen Evaluationsbericht mit Anlagen.

Innerhalb der Schulgemeinschaft sollen die Evaluationsergebnisse reflektiert und eigenverantwortlich Schwerpunkte für die Unterrichts- und Schulentwicklung gesetzt werden. Gemeinsam mit dem Staatlichen Schulamt werden daran anschließend Zielvereinbarungen getroffen. Die Schule kann sich vom Staatlichen Schulamt und vom Thüringer Institut für Lehrerbildung, Lehrplanentwicklung und Medien bei Bedarf Unterstützung holen. Über die Umsetzung der Zielvereinbarungen muss die Schule regelmäßig vor der Schulkonferenz Rechenschaft ablegen.

In Tabelle 1 ist eine Synopse der wichtigsten Charakteristika der Schulinspektionsverfahren in den vier Bundesländern Baden-Württemberg, Niedersachsen, Mecklenburg-Vorpommern und Thüringen aufgeführt.

Tab. 1: Ausgewählte Charakteristika von Schulinspektionsverfahren im Ländervergleich

	Baden-Württemberg	Niedersachsen	Mecklenburg-Vorpommern	Thüringen
Zielsetzung	• der Schule eine externe, kriteriengeleitete und von der Schulaufsicht unabhängige Rückmeldung zum „System Schule“ geben • Impulse für die Qualitätsentwicklung, indem Stärken und Verbesserungsbereiche aufgedeckt werden	• Hinweise, Anregungen und Impulse für die Verbesserung der schulischen Arbeit • Qualitätskontrolle	• Bewertung der Qualität von Schule, Unterricht und Erziehung anhand von definierten Qualitätsmerkmalen und Leistungsdaten • Unterstützung der schulischen Qualitätsentwicklung durch Aufzeigen von Stärken und Schwächen	• Festlegung von Entwicklungszielen im Rahmen abrechenbarer Zielvereinbarungen mit dem Staatlichen Schulamt auf der Grundlage der Evaluationsergebnisse
Inspektionsteam	• hauptamtlich	• hauptamtlich	• haupt- und nebenamtlich	• haupt- und nebenamtlich
Rückmeldeformat	• gestufte Rückmeldung (vorläufiger und endgültiger Evaluationsbericht) • Stellungnahme der Schule zu vorläufigem Bericht	• gestufte Rückmeldung (vorläufiger und endgültiger Evaluationsbericht) • Stellungnahme der Schule zu vorläufigem Bericht	• einfache Rückmeldung • Stellungnahme der Schule möglich	• einfache Rückmeldung
Konsequenzen	• Die Schule ist verpflichtet, aus dem Fremdevaluationsbericht Zielvorstellungen und Maßnahmen zur Schulentwicklung abzuleiten. Diese legt sie der Schulaufsicht vor und trifft mit ihr eine Zielvereinbarung.	• Bei gravierenden Mängeln wird eine Nachinspektion festgelegt (Konzentration auf ermittelte Defizitbereiche).	• In verbindlichen Zielvereinbarungen zwischen Schule und Schulaufsicht werden Maßnahmen beschlossen, die der konkreten Entwicklungsarbeit dienen.	• Schule setzt Schwerpunkte für die Unterrichts- und Schulentwicklung und bereitet zusammen mit dem Staatlichen Schulamt die Zielvereinbarung vor.

Quelle: Landesbildungsserver Baden-Württemberg (URL: http://www.schule-bw.de); Bildungsserver Mecklenburg-Vorpommern (URL: http://www.bildung-mv.de); Niedersächsischer Bildungsserver (URL: http://www.nibis.de); Thüringer Schulportal (URL: http://www.schulportal-thueringen.de)

Vergleicht man die Konzeptionen der Inspektionsverfahren der vier Bundesländer, fällt auf, dass sie sich in ihren Zielsetzungen kaum unterscheiden, während die

Durchführung und die Zusammensetzung der Inspektionsteams deutliche länderspezifische Ausprägungen aufweisen. Ebenso werden Unterschiede in den Rückmeldeverfahren und den Konsequenzen sichtbar.

Die Auswahl der vier Bundesländer erfolgte, um eine möglichst große konzeptuelle und verfahrensmäßige Differenz der Schulinspektionsmethoden abzudecken. Zudem liegt der Fokus auf Flächenländern sowie einer Repräsentation von ost- und westdeutschen Ländern. Wendet man zudem eine wissenschaftlich orientierte Systematisierung der Inspektionsmodelle an, dann können auf der Basis von aktuellen Erhebungsdaten zu den Schulinspektionskonzepten der deutschen Bundesländer (vgl. Döbert/Rürup/Dedering 2008; Rürup 2008) Niedersachsen und Mecklenburg-Vorpommern als Bundesländer identifiziert werden, in denen die Schulinspektion vordringlich einem wissenschaftsorientierten Erhebungsmodell folgt, während in Baden-Württemberg und Thüringen Elemente eines Peer-Reviews und einer einzelfallorientierten Evaluation eine stärkere Rolle spielen.

3. Forschungsstand

Laut Husfeldt (2011a) lassen sich Studien zu Wirkungen und Wirksamkeit externer Schulevaluation in drei Typen einteilen: Akzeptanzstudien, Studien zu Maßnahmen und Schulleistungsstudien. Akzeptanzstudien untersuchen die Einstellungen der Akteure und die damit verbundenen Wirksamkeitserwartungen oder -vermutungen. Studien zu Maßnahmen untersuchen die der externen Evaluation folgenden Reaktionen und anschließenden Entwicklungsmaßnahmen. Schulleistungsstudien analysieren direkt die Wirkung von externer Schulevaluation auf Leistungen bzw. Leistungssteigerungen (vgl. ebd., S. 14).

Auch wenn das unserer Untersuchung zugrunde liegende theoretische Rahmenmodell prinzipiell alle drei Typen abdeckt, beschränken wir uns in dieser Arbeit vorrangig auf den ersten Typ, wie Akteure Inspektion wahrnehmen und ob und wie sie sie akzeptieren. Dabei finden sich in der Literatur Studien zu verschiedenen schulischen Akteursgruppen. Der Fokus unserer Studie und damit auch dieses Beitrages liegt ausschließlich auf der Ebene der Schulleitungen. Im Folgenden werden daher empirische Studien berichtet, in denen Wahrnehmungen und Reaktionen von Schulleiterinnen und Schulleitern, aber teilweise auch von anderen Akteuren, auf Inspektionsverfahren thematisiert werden.

Die meisten Studien stammen aus dem angelsächsischen Raum, was wohl der Tatsache geschuldet ist, dass Schulinspektion dort bereits wesentlich länger implementiert ist als in Deutschland und sich dementsprechend auch die Forschungslandschaft länger entwickeln konnte.

Betrachtet man die Beziehung zwischen Inspektoren bzw. Inspektorinnen und Akteuren, lässt sich feststellen, dass sie meist als positiv und respektvoll wahrgenommen wird (vgl. Wilcox/Gray 1996). Trotzdem sind Lehrkräfte der Meinung, dass sie während der Inspektion einem hohen Druck ausgesetzt sind (vgl. Brimblecombe/Ormston/Shaw 1996). Es kann sogar dazu führen, dass Lehrkräfte durch Inspektionen so weit verängstigt werden, dass psychische und emotionale Probleme auftauchen können; teilweise ist dies dem Auftreten der Inspektoren bzw. Inspektorinnen geschuldet (vgl. Gray/Gardner 1999).

Leeuw (2002) betont, dass eine wechselseitige vertrauensvolle Beziehung zwischen Schulen und Inspektoren bzw. Inspektorinnen eine entscheidende Rolle bei der Schulinspektion spielt und nur Inspektionen, die durch eine reziproke Beziehung zu Schulen gekennzeichnet sind, effektiv sind.

Einen großen Einfluss auf die Schulreaktionen nehmen die Inspektionsergebnisse. Empfehlungen, die zur Schulpraxis, zur Schulkultur und zum Fokus passen, werden am erfolgreichsten akzeptiert und implementiert (vgl. Standaert 2000). Lehrkräfte halten sowohl das mündliche als auch das schriftliche Feedback nach einer Inspektion für einen wichtigen Impuls für Schulentwicklung (vgl. Fidler/Ouston/Earley 1994; Fidler/Earley/Ouston 1995). Das Feedback sollte als relevant, klar und brauchbar erlebt werden (vgl. Brimblecombe/Shaw/Ormston 1996).

In einer Evaluationsstudie der englischen Inspektion des „Office for Standards in Education, Children's Services and Skills (Ofsted)" konnte gezeigt werden, dass sich die Wirksamkeitsvorstellungen der unterschiedlichen Akteure unterscheiden (vgl. Kogan/Maden 1999). So vertrauen Eltern am stärksten auf die Bildungsqualität, die Standards und das Finanzmanagement im Zusammenhang mit der Inspektion, während das Vertrauen darauf bei den Schulleitungen am geringsten ausgeprägt ist. Auch politische Akteure haben eine positivere Einstellung zu Inspektion als Schulleitungen. Sie vermuten eine Wirkung auf die Schülerleistungen und eine Verbesserung der Rechenschaftslegung durch die Inspektion, was von kaum einer Schulleitung oder Lehrkraft geteilt wird.

Schulleitungen sehen vereinzelte positive Aspekte der Inspektion, mehrheitlich schätzen sie aber den Nutzen als dem Kostenaufwand nicht angemessen ein (vgl. Cuckle/Broadhead 1999). Werden die Inspektoren bzw. Inspektorinnen als kompetent und die Inspektion als qualitativ hochwertig wahrgenommen, sehen Schulleitungen eher einen Nutzen für ihre Schule.

O'Connor (2001) findet in seiner in Irland durchgeführten Studie, dass es einen Mangel an Konsens unter den Akteuren bezüglich der Ziele der Schulinspektion gibt. Lehrkräfte sehen Schulinspektion vorrangig als Kontrollinstrument, das Ängste generiert, während die Inspektoren und Inspektorinnen die Entwicklungsdimension beto-

nen. Trotzdem waren sich die Lehrkräfte in ihrer Wahrnehmung der Inspektoren und Inspektorinnen einig und sahen diese als sensibel, freundlich und glaubwürdig an.

Croxford, Gray und Ozga (2009) führten in England und Schottland eine Befragung von Lehrkräften zu ihrer Einstellung zu Verfahren der „Quality Assurance and Evaluation (QAE)" – darunter fällt auch Schulinspektion – durch. In beiden Ländern wurden die QAE-Prozesse positiver bewertet, wenn die Lehrkräfte selbst ein gewisses Maß an Kontrolle darüber innehatten, als wenn sie ihnen nach dem Top-down-Prinzip ausgesetzt waren. Sowohl Lehrkräfte aus England als auch aus Schottland sind gegen die Veröffentlichung von Ergebnissen und Rankings von Schulen, da sie glauben, dass dadurch die pädagogische Praxis verzerrt wird.

Eine Schweizer Studie, die zusätzlich noch die deutschen Bundesländer Hamburg und Bayern mit einbezog und die speziell die Unterrichtsevaluation als einen Teil der Schulinspektion in den Blick nahm, findet, dass eine hohe Gewichtung der Unterrichtsevaluation und des Feedbackgesprächs die Akzeptanz der Evaluation erhöht (vgl. Landwehr/Huber 2008). Die Unterrichtsevaluation und das nachfolgende Feedbackgespräch führten dazu, dass sich Lehrkräfte ernst genommen fühlen und Entwicklungsimpulse erhalten.

Die von Gärtner, Hüsemann und Pant (2009) in Deutschland durchgeführte Studie, bei der Schulleitungen in Brandenburg befragt wurden, kommt zu einem Ergebnis, das Untersuchungen aus England widerspricht: Schulleitungen berichten nur selten von Ängsten in Bezug auf die Inspektion. Eine mögliche Erklärung für die Diskrepanz der Ergebnisse in England und Deutschland könnte in den in England deutlich negativeren Konsequenzen bei schlechtem Inspektionsergebnis liegen.

Gärtner, Hüsemann und Pant (ebd.) fanden außerdem, dass die Schulleitungen überwiegend von einer hohen wahrgenommenen Qualität der Schulinspektion berichten – ein Effekt, der bei Schulen mit gutem Inspektionsergebnis noch deutlicher ausfiel.

Eine weitere deutsche Studie, die in Nordrhein-Westfalen durchgeführt wurde, kommt zu dem Ergebnis, dass Lehrer wahrscheinlicher die Inspektion positiv wahrnehmen, wenn sie das Feedback der Inspektoren und Inspektorinnen als fair, informativ und relevant wahrnehmen (vgl. Dedering/Müller 2011).

Eine Befragung von Lehrkräften in Niedersachsen kommt zu dem Fazit, dass Lehrkräfte Schulinspektion zwar insgesamt positiv bewerten, im Vergleich zu Schulleitungen, Eltern, schulischen Mitarbeitern und Mitarbeiterinnen sowie Schülern und Schülerinnen allerdings kritischer sind (vgl. Schwank/Sommer 2012). Die Lehrkräfte sind außerdem häufig der Meinung, dass die Schulinspektion keine neuen Erkenntnisse gebracht hat. Auch sind fast alle Lehrkräfte der Meinung, dass durch Schulinspektion zusätzliche Arbeit anfällt; die Hälfte der Befragten ist der

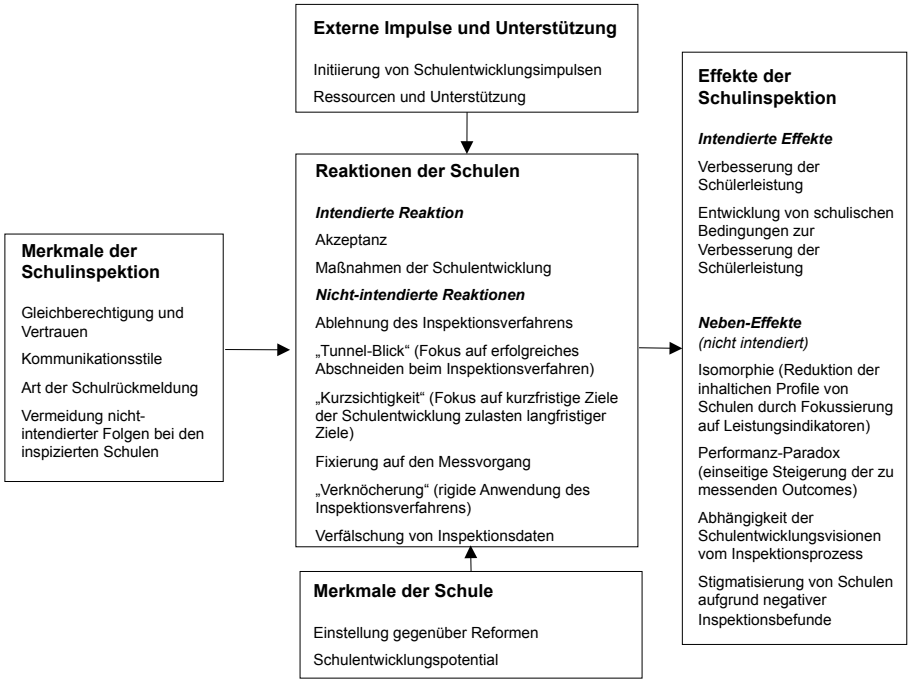
Ansicht, der Mehraufwand bewege sich in einem nicht mehr vertretbaren Rahmen. Trotzdem gibt es auch große Zustimmung bei positiven Empfindungen gegenüber der Schulinspektion. So werden die Inspektionsteams als schulfachlich kompetent wahrgenommen, Interviews verlaufen in einer angenehmen Atmosphäre, in der Lehrkräfte ihre Kenntnisse einbringen können, und mündliche Rückmeldungen nach der Schulinspektion werden als informativ bewertet (überall über 80 Prozent Zustimmung). 72 Prozent der befragten Lehrkräfte sind überwiegend mit der Inspektion zufrieden.

4. Wirkungsmodell von Ehren und Visscher

Die im Fokus dieses Beitrages stehende Frage, ob die aufgezeigten länderspezifischen Variationen der Schulinspektionsverfahren mit entsprechend unterschiedlicher Akzeptanz und divergierenden Verarbeitungsstrategien bei Schulleiterinnen und Schulleitern einhergehen, soll vor dem Hintergrund eines angemessenen Wirkungsmodells der Schulinspektion beantwortet werden. In der vorliegenden Studie wurde das Rahmenmodell der Wirkung von Schulinspektionen nach Ehren und Visscher (2006, vgl. Abb. 1) als Grundlage für die Operationalisierung entsprechender Fragebogen-Items und für die Formulierung darauf aufbauender Zusammenhangshypothesen verwendet. Für die hier zu berichtenden querschnittlichen Daten einer ersten Erhebungswelle muss jedoch die Frage nach der expliziten Wirkung von Schulinspektionen zunächst in den Hintergrund treten. Der Schwerpunkt dieser Studie liegt in der Identifizierung von potentiell variierenden Akzeptanz- und Wahrnehmungsmustern zwischen Schulleitungen, die unterschiedliche Verfahrensformen der Schulinspektion erlebt haben.

Im Gegensatz zu anderen Wirkungsmodellen (vgl. z.B. Reezigt/Creemers 2005) zielt das Modell von Ehren und Visscher (2006) weniger auf die Identifizierung der Gelingensbedingungen effektiver Schulentwicklung als vielmehr auf eine angemessene Beschreibung von sowohl erwünschten als auch unerwünschten Auswirkungen externer Schulevaluation (vgl. Husfeldt 2011b, S. 268). Die von den Autoren selbst gewählte Bezeichnung „Rahmenmodell" macht deutlich, dass mit dem formulierten Modell zunächst Zusammenhänge zwischen Charakteristika der Inspektionsverfahren und intendierten sowie nicht-intendierten Folgen auf Seiten der Schulen in den Blick genommen werden können. Eine konkrete theoretische Ausformulierung der schulinternen Wirkmechanismen enthält das Modell von Ehren und Visscher (2006) jedoch nicht.

Abb. 1: Rahmenmodell zur Wirkung von Schulinspektionen



Quelle: nach Ehren und Visscher (2006); eigene Übersetzung

Es lassen sich im von Ehren und Visscher (ebd.) vorgelegten Rahmenmodell fünf Dimensionen unterscheiden: (1) Merkmale der Schulinspektion, (2) schulentwicklungsrelevante Merkmale der Schule, (3) externe Impulse und Unterstützung, (4) Reaktionen der Schule auf die Schulinspektion sowie (5) Effekte und Nebeneffekte der Schulinspektion. Dabei werden die Merkmale der Schulinspektion als unabhängige Variablen betrachtet, die unterschiedliche Reaktionen bei den Schulen hervorrufen können. Diese sind wiederum durch externe Impulse und Unterstützungen beeinflussbar und vom generellen Innovationswillen der schulischen Akteure abhängig. Die schulischen Reaktionen auf die Schulinspektion führen letztendlich zu erwünschten und unerwünschten Wirkungen auf unterschiedlichen Handlungsebenen.

(1) Merkmale der Schulinspektion

Hierunter fassen Ehren und Visscher (ebd.) vier verschiedene Aspekte: eine Beziehung der Inspektoren und Inspektorinnen zu den Schulen, die auf Gleichberechtigung und Vertrauen basiert; die in den Inspektionsverfahren genutzten Kommunikationsstile; die Art und Weise der Ergebnisrückmeldung sowie die in einer Schulinspektion erkennbaren Maßnahmen zur Vermeidung nicht-intendierter

Folgen. Für die Vermeidung unerwünschter Nebenwirkungen schlagen Ehren und Visscher (ebd.) ein Set von Maßnahmen vor, das letztendlich auf ein umfangreiches Monitoring-System der Schule und ihrer Umwelt zielt. Nicht zuletzt kann auch die Androhung ernsthafter Konsequenzen unerwünschte Nebenwirkungen oder sogar Datenmanipulation vermeiden helfen.

(2) Merkmale der Schule

Die intermediären Dimensionen zwischen den Merkmalen der Schulinspektion und deren Auswirkungen beinhalten nach Ehren und Visscher (ebd.) zum einen die Merkmale der Schule selbst, ihre Reaktionen auf die Inspektion und die von außen gegebenen Impulse und Unterstützungen (vgl. Abb. 1). Zu den schulentwicklungsrelevanten Merkmalen einer Schule zählt vor allem, dass eine Schule dazu in der Lage ist, sowohl interne als auch externe Evaluationsergebnisse für ihre Entwicklung zu nutzen. Die Rolle der Schulleitung ist in diesem Zusammenhang verantwortungsvoll und komplex. Im *„transformative leadership concept"* nach Geijsel (2001) sind entsprechende Aspekte beschrieben: Dazu gehört u.a. die aktive Unterstützung der Lernprozesse des Lehrerkollegiums und der gesamten Schulorganisation. Eine effektive Schulleitung formuliert und unterstützt Visionen künftiger Entwicklungen und befähigt die schulischen Akteure, diese auch umzusetzen.

(3) Externe Impulse und Unterstützung

Nur ein kleiner Teil der Schulen ist jedoch in der Lage, aus sich heraus und ohne externe Impulse und Unterstützung, Innovationen zu initiieren. Gern werden Schulentwicklungsberater und -beraterinnen nachgefragt, und es wird auf entsprechende Materialien und Kurse zurückgegriffen. Die lokale Gemeinschaft, externe Schulentwicklungsberater und -beraterinnen oder auch Wissenschaftler und Wissenschaftlerinnen spielen eine wichtige Rolle beim Anstoß von Veränderungen. Sie können Schulen dazu anregen, über Schulentwicklung nachzudenken, oder bereits eingeschlagene Entwicklungsprozesse durch ihre Expertise unterstützen. Aber auch Entwicklungen in der schulischen Umwelt selbst, wie veränderte Schülerzahlen oder das Gebot der Inklusion von Schülern und Schülerinnen mit Förderbedarf, können schulische Entwicklungen in positiver, aber auch in negativer Weise forcieren.

(4) Reaktionen der Schule

Die häufige Inanspruchnahme des Ehren-Visscher-Modells (2006) in aktuellen empirischen Untersuchungen liegt sicherlich nicht zuletzt darin begründet, dass dieses Modell nicht nur nach den gelingenden Bedingungen entwicklungsfördern-

der Schulinspektion fragt, sondern auch deren nicht-intendierte Auswirkungen in den Blick nimmt. Reaktionen von Schulen auf die Schulinspektion können auch in einer Art und Weise erfolgen, die als diametral entgegengesetzt zu den von Inspektionsverfahren intendierten Reaktionen zu bezeichnen ist und aus der mangelnden Übereinstimmung zwischen den Zielen des Inspektorats und den je eigenen Zielen der Schule resultiert.

(5) Effekte und Nebeneffekte der Schulinspektion

Die Betrachtung der Effekte und Nebeneffekte der Schulinspektion verweist auf die abhängige Dimension im Rahmenmodell von Ehren und Visscher (ebd.). Ein hauptsächlich intendiertes Ziel der Inspektionsverfahren stellt im Modell die Verbesserung der Schülerleistung dar. Die erwähnten nicht-intendierten Reaktionen der Schulen führen jedoch auch zu einer Reihe von negativen Nebenwirkungen, die nicht im Interesse der Inspektion sind und denen durch Standardisierung und Optimierung von Inspektionsverfahren vorgebeugt werden soll.

Welche konkreten Mechanismen in den Schulen selbst notwendig sind, um von einer Akzeptanz der Inspektionsergebnisse zu einer gelingenden Schulentwicklung zu gelangen, wird im Ehren-Visscher-Modell selbst nicht explizit sichtbar. In anderen Inspektionswirkungsmodellen wird dagegen gerade diese innerschulische Verarbeitung der Inspektionsergebnisse thematisiert (vgl. z.B. Koch u.a. 2006; Sommer 2011). Ergänzend weist Sommer (ebd., S. 155f.) darauf hin, dass die verschiedenen schulischen Gruppen (Schulleitung, Kollegium, Schülerschaft, Eltern, Schulaufsicht) mit der Durchführung einer Inspektion durchaus divergierende Interessen verbinden. Ein Wirkungsmodell von Schulinspektion sollte daher die unterschiedlichen Perspektiven der schulischen Akteure aufnehmen und für die Phasen der Rezeption, Reflexion und Aktion im Anschluss an eine Inspektion zunächst getrennt betrachten und dann einer triangulativen Zusammenführung unterziehen.

Im Hinblick auf das Ehren-Visscher-Modell muss weiterhin einschränkend festgestellt werden, dass dieses Rahmenmodell vor allem auf die Wirksamkeit der Inspektion in Bezug auf die Schulentwicklung mit dem Ziel der Steigerung der Schülerleistung abzielt (vgl. Husfeldt 2011b). Alternative Schulentwicklungsziele werden im Modell nicht benannt. Dennoch stellt es mit seiner gleichzeitigen Berücksichtigung von intendierten und nicht-intendierten Reaktionen auf und Wirkungen von Schulinspektionen ein elaboriertes Modell zur Beschreibung der in Schulen beobachtbaren Beeinflussungen durch Inspektionsverfahren dar und bietet sich daher für eine kritische Überprüfung im Rahmen empirischer Untersuchungen an.

5. Methode

Standardisierte Befragung von Schulleitungen

Im Rahmen des BMBF-Verbundprojektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene" (vgl. Brüsemeister u.a. 2009) wurde das Ehren-Visscher-Modell (2006) als Grundlage für die Operationalisierung einer standardisierten Befragung von Schulleiterinnen und Schulleitern an allgemein bildenden Schulen in vier deutschen Bundesländern verwendet. Es sei bereits an dieser Stelle betont, dass die hier vorgestellten quantitativen Befunde nur einen Teil der Forschungsleistungen im Verbundprojekt ausmachen. Die innerschulischen Verarbeitungsprozesse und die subjektiven Deutungen der schulischen Akteure im Hinblick auf Schulinspektionen werden durch weitere qualitative Studien im Projektverbund erfasst. Die Zusammenführung der quantitativen und qualitativen Befunde im Rahmen einer Triangulation kann die Beschränkungen der jeweiligen methodischen Zugänge überwinden und eine umfassende Beschreibung schulischer Reaktionen auf externe Steuerungsimpulse liefern.

Die erste quantitative Befragung der Schulleitungen erfolgte im Winter 2011/2012 in Baden-Württemberg, Niedersachsen, Mecklenburg-Vorpommern und Thüringen. Eine zweite Befragung der Schulleitungen fand zur Jahreswende 2012/2013 statt. In diesem Beitrag werden die Befunde der ersten Erhebungswelle berichtet. Die standardisierte Befragung wurde in allen vier Bundesländern als Online-Erhebung durchgeführt.

Die bewusste Eingrenzung der Erhebung auf Schulleitungen bzw. auf in das Inspektionsverfahren involvierte Personen (z.B. Konrektoren/Konrektorinnen oder Steuergruppenleiter und -leiterinnen) hat mehrere Gründe:

- Schulleitungen (bzw. deren Vertreter/Vertreterinnen) sind auf jeden Fall Ansprechpartner bei Inspektionsverfahren. Inwieweit andere Mitglieder des Lehrerkollegiums involviert sind, ist vorab nicht bekannt und wäre nur mit unverhältnismäßigem Aufwand flächendeckend zu recherchieren.
- Eine Befragung von Lehrerkollegien ist logistisch aufwändiger und sollte eher postalisch durchgeführt werden, um einen akzeptablen Rücklauf zu erhalten. Diese Vorgehensweise ist für einen Ländervergleich unverhältnismäßig aufwändig und mit zu hohen Kosten verbunden, da bei einem solchen Vorgehen die Mehrebenenstruktur der zu erhebenden Daten durch entsprechend hohe Stichprobenumfänge berücksichtigt werden müsste.
- Als weiteres inhaltliches Argument ist anzuführen, dass ein wichtiger Einflussfaktor auf die Rezeption einer anstehenden Schulinspektion die emotionale Verarbeitung und professionelle Bearbeitung durch die Schulleitung ist (vgl. Brimblecombe/Ormston/Shaw 1995).

Die Methode der Vollerhebung wurde gewählt, um eine ausreichende Anzahl sowohl inspizierter als auch nicht inspizierter Schulen in die Analyse aufzunehmen. Durch eine vorab festgelegte Stichprobenziehung wäre es bei den gleichzeitig weiter voranschreitenden Schulinspektionsverfahren in den vier Bundesländern schwierig gewesen, diese beiden Gruppen über einen längeren Zeitraum für eine längsschnittliche Betrachtung konstant zu halten. Der nicht durch eine Zufallsauswahl hervorgerufene Response auf eine Online-Erhebung kann jedoch zu Stichproben führen, die mit Verzerrungen in bestimmten Merkmalen (z.B. der Motivation, sich mit Schulinspektion in reflexiver Weise auseinanderzusetzen) behaftet sein können. Da das vordergründige Interesse unserer Studien im Verbundprojekt jedoch darin besteht, die im Ehren-Visscher-Modell (2006) postulierten Zusammenhänge zu überprüfen und zu spezifizieren, erscheint uns die gewählte Vorgehensweise zumindest akzeptabel. Eine weitere Verzerrung der Stichprobe liegt sicher in den nicht einheitlich zu realisierenden Einladungsmodalitäten für die Erhebung. Die Länder handhaben die Adressweitergabe ihrer Schulen unterschiedlich, so dass wir nur in Baden-Württemberg, Mecklenburg-Vorpommern und Thüringen Adresslisten erhalten haben, mittels derer wir die Einladungen zur Online-Umfrage selbst direkt verschicken konnten. Niedersachsen stellte keine Adresslisten zur Verfügung, schickte aber unsere Einladung, ergänzt durch ein Anschreiben des Ministeriums, über einen landesinternen Verteiler an alle Schulen. Dieser deutlich höhere Aufforderungscharakter hat sicherlich zu dem im Vergleich zu den anderen Ländern überdurchschnittlichen Rücklauf geführt. Die Beschränkungen der vorliegenden Studie werden im Diskussionsteil noch einmal aufgegriffen und mögliche Alternativen aufgezeigt.

Stichprobe

An der ersten Erhebung nahmen insgesamt 1.470 Schulleiterinnen und Schulleiter (52 Prozent weiblich, 48 Prozent männlich) teil, die den Online-Fragebogen komplett ausgefüllt haben. Für die nachfolgenden Analysen greifen wir auf Schulleiterinnen und Schulleiter aus *Grundschulen und Gymnasien* zurück, da nur diese beiden Schulformen in allen vier untersuchten Bundesländern gleichermaßen vorfindbar sind. Im Bereich der Sekundarstufe I ist eine zu große organisatorische Differenzierung der Schulformen zu konstatieren, die für einen überschaubaren Ländervergleich eine zu hohe Komplexität aufweist. Weiterhin beschränken wir uns im Rahmen dieser Ergebnisdarstellung auf *bereits inspizierte Schulen*, um Konfundierungen mit dem Erleben bzw. Nichterleben einer Schulinspektion auszuschließen. In Tabelle 2 ist die resultierende Stichprobe (N=632) nach Bundesland und Schulform dargestellt.

Tab. 2: Stichprobe nach Bundesland und Schulform (erste Zeile: alle Schulen/inspizierte Schulen in der Stichprobe; zweite Zeile: Zahlen der Grundgesamtheit in Klammern)

Bundesland	Grundschule	Gymnasium	Gesamt
Baden-Württemberg	195 / *67* (2.433)	43 / *11* (378)	238 / *78* (2.811)
Niedersachsen	484 / *441* (1.765)	72 / *71* (221)	556 / *512* (1.986)
Mecklenburg-Vorpommern	11 / *9* (215)	7 / *7* (41)	18 / *16* (256)
Thüringen	33 / *24* (431)	3 / *2* (85)	36 / *26* (516)
Gesamt	723 / 541	125 / 91	848 / *632*

Anmerkung. Die *kursiv* hervorgehobenen Zahlen stellen die Analysestichprobe der nachfolgend berichteten Befunde dar.

Quelle: Quantitative Befragung von Schulleitungen in vier Bundesländern (2011/2012) im Rahmen des BMBF-Projektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene" und Statistische Landesämter; eigene Berechnungen

Für die beiden ostdeutschen Bundesländer Mecklenburg-Vorpommern und Thüringen sind nur sehr geringe Stichprobenumfänge zu verzeichnen. Trotz mehrfacher Erinnerungen und der Auslobung von Incentives konnte die Teilnahmebereitschaft in diesen beiden Ländern nicht weiter gesteigert werden. Ein Grund für die verhaltene Teilnahmebereitschaft ist sicherlich auch, dass mehrere empirische Untersuchungen zeitgleich in diesen Ländern an den Schulen stattfanden und daher die Motivation der Schulleiterinnen und Schulleiter, einen weiteren umfangreichen Fragebogen auszufüllen, entsprechend gering war.

Der Vergleich der Anteile von Grundschulen und Gymnasien in der jeweiligen Länderstichprobe mit den entsprechenden Anteilen in der Schulstatistik zeigt zudem, dass die länderspezifischen Verhältnisse zwischen Grundschul- und Gymnasialanteilen in der Stichprobe kaum abbildbar sind (vgl. Tab. 3). An dieser Stelle ist kritisch zu konstatieren, dass der Rücklauf aus unserer breit angelegten Online-Erhebung nicht den gewünschten Quoten entspricht und eine Zufallsauswahl oder ein theoretisches Sampling möglicherweise bessere (und repräsentative) Rücklaufe erbracht hätte. Die Gründe, warum eine Zufallsauswahl nicht vorgenommen werden konnte, wurden bereits erläutert. Gegen ein theoretisches Sample sprach zum Zeitpunkt der Projektplanung die Stellung der quantitativen Studie im Projektverbund. Durch die quantitativen Befunde sollten für alle vier Bundesländer möglichst breit angelegte und maximale Varianz aufweisende Vergleichszahlen generiert werden, die anschließend mit den qualitativen Befunden der anderen Verbundprojekte einer Triangulation unterzogen werden. Die gemachten Erfahrungen zeigen jedoch, dass vor dem

Hintergrund der derzeitigen starken „Beforschung“ von Schulen die Motivation, an einer Online-Befragung freiwillig teilzunehmen, nur noch sehr gering ausgeprägt ist.

Tab. 3: Vergleich der Anteile an Grundschulen und Gymnasien* in der Rücklauf-Stichprobe (inspizierte und nicht-inspizierte Schulen) mit den Anteilen der Grundgesamtheit im Schuljahr 2011/12 (in Prozent; nach Bundesland)

Bundesland	Anteil Grundschulen lt. Schulstatistik	*Anteil Grundschulen Stichprobe*	Anteil Gymnasien lt. Schulstatistik	*Anteil Gymnasien Stichprobe*
Baden-Württemberg	65,1	*45,9*	10,2	*10,1*
Niedersachsen	43,2	*52,3*	7,7	*7,8*
Mecklenburg-Vorpommern	43,4	*24,4*	8,3	*15,5*
Thüringen	52,8	*43,4*	10,4	*3,9*

Anmerkung. * Anteile der Grundschulen und Gymnasien an allen öffentlichen allgemein bildenden Schulen.
Quelle: Statistische Landesämter; eigene Berechnungen

6. Ergebnisse

Für die Betrachtung der potentiell länderspezifisch variierenden Einschätzungen von Schulinspektionsverfahren aus Sicht der Schulleitungen werden im Folgenden auf einer deskriptiven Ebene die Mittelwerte der entsprechenden Skalen berichtet. Die abhängigen Variablen orientieren sich an der Dimensionierung des Rahmenmodells von Ehren und Visscher (2006). Als Gruppierungsvariable wird die Länderzugehörigkeit der befragten Schulleiterinnen und Schulleiter betrachtet, um potentielle Unterschiede in den Wahrnehmungen zwischen den untersuchten Bundesländern aufdecken zu können. Darüber hinaus werden Differenzen dieser Mittelwerte zwischen den Bundesländern Niedersachsen und Baden-Württemberg auf ihre statistische Signifikanz mittels t-Tests überprüft. Die Beschränkung des Mittelwertvergleiches auf diese beiden Länder erfolgt aufgrund methodischer und inhaltlicher Überlegungen: Zum einen ist die Stichprobengröße der Bundesländer Thüringen und Mecklenburg-Vorpommern zu gering, um belastbare Signifikanztestungen und Berechnungen von Effektstärken (vgl. Bortz/Döring 1995, S. 574f.) vornehmen zu können. Zum anderen stellen die Verfahren in Niedersachsen und Baden-Württemberg einen maximalen Kontrast zwischen den in unserer Ländergruppe vorfindbaren Konzeptionen der Schulinspektion dar (siehe Abschnitt 2). Das wissenschaftsorientierte und mit Sanktionen (Nachinspektion) verbundene Verfahren ist in Niedersachsen etabliert, während das auf einem Peer-Review basierende Verfahren in Baden-Württemberg beheimatet ist.

Tab. 4: Merkmale der Schulinspektionsverfahren aus Schulleitungssicht im Ländervergleich

	BW (N= 73-78)	*MV (N= 14-15)*	*NI (N= 477-495)*	*TH (N= 23-26)*	*Ges. (N= 587-615)*	*t-Wert***	*Signifikanz*	*d*
Skala (Beispielitem)	*Mittelwert* Standardabweichung*							
Ziele der SI 1: Schul- und Unterrichtsentwicklung *(Schulinspektion dient der Verbesserung des Unterrichts.)*	2.69 .61	2.68 .68	2.56 .60	3.01 .55	2.60 .61	-1.83	n.s.	-
Ziele der SI 2: Wettbewerb zwischen Schulen *(Schulinspektion dient der Förderung des Wettbewerbs zwischen Schulen.)*	2.34 .84	2.23 .70	2.57 .79	2.06 .60	2.51 .80	2.36	.019	.29
Positive Wahrnehmung des Inspektionsteams *(Den Inspektoren war es wichtig, uns zu vermitteln, wie ihre Bewertung zustande kommt.)*	3.28 .68	3.23 .62	3.11 .68	3.56 .44	3.15 .67	-2.02	.043	.25
Qualität der SI *(Das Inspektionsteam hat die Kernpunkte für eine Verbesserung unserer Schule aufgezeigt.)*	3.01 .57	2.97 .55	2.88 .59	3.21 .49	2.91 .58	-1.85	n.s.	-
Charakteristik der SI *(Die Schulinspektion offenbart unsere blinden Flecken.)*	2.73 .56	2.65 .41	2.64 .50	2.86 .52	2.66 .51	-1.40	n.s.	-
SI-Bericht *(Der Bericht ist plausibel.)*	3.46 .57	3.54 .55	3.41 .50	3.61 .44	3.43 .51	-.729	n.s.	-

Anmerkungen. BW = Baden-Württemberg, MV = Mecklenburg-Vorpommern, NI = Niedersachsen, TH = Thüringen.

* Antwortskala: 1 „trifft überhaupt nicht zu" – 4 „trifft voll und ganz zu".

** Die t-Tests wurden zwischen den Ländern Niedersachsen und Baden-Württemberg berechnet.

Quelle: Quantitative Befragung von Schulleitungen in vier Bundesländern (2011/2012) im Rahmen des BMBF-Projektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene"; eigene Berechnungen

In Tabelle 4 sind die *Merkmale der Schulinspektion* aus Sicht der Schulleitungen im Ländervergleich aufgeführt. An den Mittelwerten der Gesamtstichprobe ist zu erkennen, dass alle aufgeführten Merkmale der Schulinspektion von den Schulleitungen eher positiv bewertet werden (im Vergleich zum theoretischen Skalenmittelwert von 2.5 für alle vierstufigen Antwortskalen). Besonders trifft dies für die Wahrnehmung des Inspektorenteams und die Einschätzung des Inspektionsberichtes zu. Auch die Qualität der Schulinspektion erscheint in allen vier untersuchten Bundesländern aus Sicht der Schulleitungen in einem eher positiven Licht. Im Vergleich dazu wird von den befragten Schulleitungen einem potentiellen Wettbewerbsgedanken der Schulinspektion weniger Zustimmung entgegengebracht. Nur in zwei der sechs dargestellten Merkmale sind statistisch signifikante Länderdifferenzen zwischen Niedersachsen und Baden-Württemberg erkennbar. Wie die aufgeführte Effektgröße d (.20 = kleiner Effekt; .50 = mittlerer Effekt; .80 = großer Effekt) zeigt, sind diese signifikanten Differenzen jedoch nur mit geringer praktischer Bedeutsamkeit verbunden. Als übergreifendes Muster ist festzustellen, dass die Mittelwerte der niedersächsischen Schulleiterinnen und Schulleiter in fünf von sechs erhobenen Merkmalen die niedrigsten Ausprägungen im Vergleich zu den Mittelwerten der Kolleginnen und Kollegen aus Baden-Württemberg, Mecklenburg-Vorpommern und Thüringen aufweisen. Ausnahme bildet das eher negativ konnotierte Merkmal des Wettbewerbsgedankens von Schulinspektion. Der Aussage, dass der Wettbewerb zwischen den Schulen durch Inspektionsverfahren forciert wird, stimmen die niedersächsischen Schulleitungen häufiger zu als ihre Kolleginnen und Kollegen der anderen drei untersuchten Bundesländer.

Tab. 5: Externe Impulse und Unterstützung aus Schulleitungssicht im Ländervergleich

	BW (N= 73-77)	*MV (N=14)*	*NI (N= 489-495)*	*TH (N= 23-24)*	*Ges. (N= 600-609)*	*t-Wert***	*Signifikanz*	*d*
Skala (Beispielitem)	*Mittelwert* Standardabweichung*							
Unterstützung 1								
(Die regionale Schulaufsicht/das Schulamt arbeitet mit unserer Schule zusammen, um die Unterrichtsqualität zu verbessern.)	2.20 .72	2.35 .86	2.13 .61	2.58 .68	2.16 .64	-.869	n.s.	-
Unterstützung 2								
(Die regionale Schulaufsicht/das Schulamt berät unsere Schule aufgrund der in Leistungsuntersuchungen festgestellten Ergebnisse.)	1.89 .73	1.95 .92	1.77 .64	2.15 .72	1.80 .67	-1.51	n.s.	-

Anmerkungen. BW = Baden-Württemberg, MV = Mecklenburg-Vorpommern, NI = Niedersachsen, TH = Thüringen.

* Antwortskala: 1 „trifft überhaupt nicht zu" – 4 „trifft voll und ganz zu".

** Die t-Tests wurden zwischen den Ländern Niedersachsen und Baden-Württemberg berechnet.

Quelle: Quantitative Befragung von Schulleitungen in vier Bundesländern (2011/2012) im Rahmen des BMBF-Projektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene"; eigene Berechnungen

In Tabelle 5 werden die erhobenen *Externen Impulse und Unterstützungsleistungen* aus Schulleitungssicht im Ländervergleich wiedergegeben. Im Vergleich zu den oben dargestellten Merkmalen der Schulinspektion zeigt sich bei dieser Modelldimension eine deutlich weniger positiv ausgeprägte Zustimmung der befragten Schulleitungen. Besonders wenig Zustimmung erhält die Frage nach der beratenden Tätigkeit der regionalen Schulaufsicht bzw. des Schulamtes nach Bekanntgabe der Ergebnisse aus zentralen Leistungsuntersuchungen. Auch in dieser Tabelle wiederholt sich das Muster der vorangegangenen Darstellung: Die niedersächsischen Schulleitungen schätzen die Unterstützungsleistungen von regionaler Schulaufsicht bzw. vom Schulamt am geringsten ein. Die Differenzen zum Bundesland Baden-Württemberg sind jedoch nicht signifikant.

Tab. 6: Merkmale der Schule aus Schulleitungssicht im Ländervergleich

	BW (N= 75-78)	*MV (N= 14-15)*	*NI (N= 487-500)*	*TH (N= 24-25)*	*Ges. (N= 600-616)*	*t-Wert***	*Signifikanz*	*d*
Skala (Beispielitem)	*Mittelwert* Standardabweichung*							
Kollegialer Zusammenhalt								
(In unserem Kollegium herrscht Konsens über das pädagogische Selbstverständnis der Schule.)	3.36 .47	2.95 .40	3.23 .50	3.28 .45	3.24 .50	-2.12	.034	.26
Innovationsbereitschaft								
(An unserer Schule ist das Kollegium stets bemüht, die Arbeit am schuleigenen pädagogischen Konzept voranzutreiben.)	2.85 .54	2.34 .57	2.83 .51	2.95 .52	2.83 .52	-.203	n.s	-
Stellenwert Fortbildungen								
(Regelmäßig an Fort- und Weiterbildungen teilzunehmen ist in unserem Kollegium selbstverständlich.)	3.22 .55	3.00 .55	3.12 .54	3.33 .55	3.14 .54	-1.58	n.s.	-
Maßnahmen Qualitätsentwicklung/-steigerung								
(Festlegung von Qualitätsindikatoren, die Auskunft über die Erreichung der Ziele der Schule geben.)	2.68 .52	2.98 .62	2.64 .48	2.91 .45	2.66 .49	-.792	n.s.	-

Anmerkungen. BW = Baden-Württemberg, MV = Mecklenburg-Vorpommern, NI = Niedersachsen, TH = Thüringen.

* Antwortskala: 1 „trifft überhaupt nicht zu" – 4 „trifft voll und ganz zu".

** Die t-Tests wurden zwischen den Ländern Niedersachsen und Baden-Württemberg berechnet.

Quelle: Quantitative Befragung von Schulleitungen in vier Bundesländern (2011/2012) im Rahmen des BMBF-Projektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene"; eigene Berechnungen

Eine weitere Dimension des Ehren-Visscher-Modells (2006) stellen entwicklungsrelevante Merkmale der Schulen selbst dar. Diese sind in Tabelle 6 im Ländervergleich aufgeführt. Bis auf den von den Schulen wahrgenommenen Zusammenhalt im Kollegium variieren die erhobenen Merkmale nicht signifikant im Ländervergleich. Dieser Befund zeigt, dass die für Schulentwicklung förderlichen Einstellungen des Kollegiums gegenüber Innovationen und Fortbildungsaktivitäten zwar zwischen einzelnen Schulen durchaus variieren können, aber im Ländervergleich kei-

ne bedeutsamen systematischen Unterschiede auftreten. Auch für die weiteren statistischen Analysen ist dieser Befund erfreulich, da die Reaktionen der Schulen auf Schulinspektionsverfahren nicht bereits durch länderspezifisch variierende Entwicklungsmerkmale der Schulen konfundiert werden.

Tab. 7: Reaktionen der Schule aus Schulleitungssicht im Ländervergleich

	BW (N= 76-77)	*MV (N=15)*	*NI (N= 486-495)*	*TH (N= 24-26)*	*Ges. (N= 603-611)*	*t-Wert***	*Signifikanz*	*d*
Skala (Beispielitem)	*Mittelwert* Standardabweichung*							
Allgemeine Akzeptanz der SI *(Schulinspektionen sind für die Arbeit der Schulen sehr wichtig.)*	2.65 .65	2.82 .58	2.50 .68	2.99 .64	2.55 .68	-1.80	n.s.	-
Negative Folgen der SI *(Die Schulinspektion hatte zusätzliche Belastungen zur Folge.)*	2.40 .69	2.49 .67	2.66 .60	2.10 .59	2.60 .62	3.47	.001	.43
Ablehnende Sichtweise auf SI *(Ich beobachte in meinem lokalen Umfeld, dass die Schulinspektion zu einer Stigmatisierung von Schulen führt.)*	1.79 .58	1.77 .47	2.14 .53	1.90 .37	2.08 .54	5.45	.000	.66

Anmerkungen. BW = Baden-Württemberg, MV = Mecklenburg-Vorpommern, NI = Niedersachsen, TH = Thüringen.

* Antwortskala: 1 „trifft überhaupt nicht zu" – 4 „trifft voll und ganz zu".

** Die t-Tests wurden zwischen den Ländern Niedersachsen und Baden-Württemberg berechnet.

Quelle: Quantitative Befragung von Schulleitungen in vier Bundesländern (2011/2012) im Rahmen des BMBF-Projektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene"; eigene Berechnungen

In Tabelle 7 werden *Reaktionen der Schule* auf Schulinspektionsverfahren aus Schulleitungssicht berichtet. Die im Fragebogen mittels Skalen operationalisierten Reaktionen der Schule beziehen sich auf die allgemeine Akzeptanz von Schulinspektionsverfahren, deren wahrgenommene negative Folgen und eine eher ablehnende Haltung gegenüber Schulinspektionen von Seiten der Schulleitungen.

Für die in Tabelle 7 vorgestellten Skalen ist zunächst bemerkenswert, dass – über alle Schulleitungen betrachtet – eine durchschnittlich bis leicht überdurchschnittlich ausgeprägte allgemeine Akzeptanz von Schulinspektionsverfahren zu konstatieren ist. Dieser Befund wird dadurch untermauert, dass eine generell ablehnende Haltung gegenüber der Schulinspektion in den vier untersuchten Bundesländern

nicht beobachtet werden kann. Gleichwohl ist die Schulinspektion nicht frei von zusätzlichen Belastungen in den Schulen; diese werden vor allem von den niedersächsischen Schulleitungen thematisiert. Auch sind es wieder die niedersächsischen Schulleiterinnen und Schulleiter, die insgesamt eine neutrale Haltung gegenüber der Schulinspektion aufweisen, aber sich jeweils signifikant in der Ablehnung und dem Belastungsempfinden von ihren Kolleginnen und Kollegen in Baden-Württemberg unterscheiden. Für die Wahrnehmung negativer Folgen und eine eher ablehnende Sichtweise auf Schulinspektion sind die signifikanten Länderdifferenzen mit mittleren Effektstärken verbunden, so dass eine Diskussion um mögliche länderspezifische Ursachen dieser Unterschiede unumgänglich erscheint (siehe Diskussionsteil dieses Beitrages).

Tab. 8: Nicht-intendierte Reaktionen der Schule aus Schulleitungssicht im Ländervergleich

	BW (N=81-85)	*MV (N=34)*	*NI (N=499-500)*	*TH (N=57)*
Item	*Mittelwert* Standardabweichung*			
Schulinspektionen führen zu einer bloßen Fixierung auf ein gutes Abschneiden der Schule.	2.42	2.29	2.82	2.12
	.93	.80	.88	.78
Ich beobachte an meiner Schule, dass die Schulinspektion die pädagogische Kreativität von Lehrkräften beschneidet.	1.66	1.38	1.94	1.54
	.73	.49	.77	.60
Ich beobachte an meiner Schule, dass wir unsere Planungen zur Schulentwicklung am Zyklus der Schulinspektion orientieren.	1.88	1.82	1.85	2.04
	.96	.83	.79	.84
Die Schulinspektion bestätigt unsere bisherigen eigenen Einschätzungen.	3.17	3.15	2.97	3.11
	.77	.56	.71	.59

Anmerkungen. BW = Baden-Württemberg, MV = Mecklenburg-Vorpommern, NI = Niedersachsen, TH = Thüringen.
* Antwortskala: 1 „trifft überhaupt nicht zu" – 4 „trifft voll und ganz zu".

Quelle: Quantitative Befragung von Schulleitungen in vier Bundesländern (2011/2012) im Rahmen des BMBF-Projektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene"; eigene Berechnungen

In der Online-Erhebung wurde auch der Versuch unternommen, die im Ehren-Visscher-Modell (2006) darüber hinausgehenden nicht-intendierten Reaktionen von Schulen zu operationalisieren. Da zum Zeitpunkt der Fragebogenerstellung noch keine entsprechenden Skalen in der Forschung berichtet wurden, haben wir den Versuch unternommen, die nicht-intendierten Reaktionen zumindest auf der Ebene einzelner Items abzubilden. In Tabelle 8 sind die damit verbundenen deskriptiven Kennwerte dargestellt.

Tab. 9: Effekte der Schulinspektion aus Schulleitungssicht im Ländervergleich

	BW (N= 70-72)	*MV (N= 13-14)*	*NI (N= 478-482)*	*TH (N= 25-26)*	*Ges. (N= 586-594)*	*t-Wert***	*Signifikanz*	*d*
Skala (Beispielitem)	*Mittelwert* Standardabweichung*							
Effekte auf Schul- und Unterrichtsentwicklung I *(Effekte der Schulinspektion: Umsetzung systematischer Schulentwicklung)*	1.94 .68	1.89 .47	1.99 .59	2.35 .58	2.00 .60	.568	n.s.	-
Effekte auf Schul- und Unterrichtsentwicklung II *(Effekte der Schulinspektion: Verbesserung der Unterrichtskompetenz der Lehrkräfte)*	2.18 .65	2.15 .66	2.23 .60	2.45 .55	2.23 .60	.679	n.s.	-

Anmerkungen. BW = Baden-Württemberg, MV = Mecklenburg-Vorpommern, NI = Niedersachsen, TH = Thüringen.

* Antwortskala: 1 „trifft überhaupt nicht zu" – 4 „trifft voll und ganz zu".

** Die t-Tests wurden zwischen den Ländern Niedersachsen und Baden-Württemberg berechnet.

Quelle: Quantitative Befragung von Schulleitungen in vier Bundesländern (2011/2012) im Rahmen des BMBF-Projektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene"; eigene Berechnungen

Die letzte Dimension des Rahmenmodells von Ehren und Visscher (2006) enthält die *intendierten und nicht-intendierten Effekte* der Schulinspektion. Im Rahmen der Online-Befragung des Verbundprojektes haben wir vor allem die intendierten Effekte im Bereich der Schul- und Unterrichtsentwicklung in den Blick genommen. In Tabelle 9 sind die entsprechenden Mittelwerte der eingesetzten Skalen im Ländervergleich aufgeführt. Die Befunde zeigen ein ernüchterndes Bild der mit Schulinspektion verbundenen Effekte. Die Auswirkungen der Inspektion auf Schul- und Unterrichtsentwicklung sind aus Sicht der Schulleitungen als marginal zu beschreiben. Signifikante Länderdifferenzen sind bei der Umsetzung systematischer Schulentwicklung nach Inspektionen nicht beobachtbar.

Wie bereits bei der Darstellung des Ehren-Visscher-Modells (ebd.) referiert (s.o.), ist von einem gewissen zeitlichen Abstand zwischen Schulinspektion und der Institutionalisierung einer Innovation auszugehen. Aus diesem Grund haben wir die dargestellten Wirkungen nicht nur nach Länderdifferenzen untersucht, sondern auch überprüft, ob es Unterschiede in den Wirkungen hinsichtlich des Zeitpunktes der Schulinspektion gibt (vgl. Tab. 10).

Tab. 10: Effekte der Schulinspektion aus Schulleitungssicht im Zeitvergleich

	Zeitpunkt (Jahr) der letzten Schulinspektion			*Ges. (N= 610-611)*	*F-Wert*	*Signifikanz*	η^2
	2003 bis 2007 (N=127)	*2008 bis 2009 (N= 204-206)*	*2010 bis 2011 (N= 278-279)*				
Skala *(Beispielitem)*	*Mittelwert* Standardabweichung*						
Effekte auf Schul- und Unterrichtsentwicklung I	2.58	2.50	2.68	2.60	5.10	.006	.02
(Effekte der Schulinspektion: Umsetzung systematischer Schulentwicklung)	.63	.61	.59	.61			
Effekte auf Schul- und Unterrichtsentwicklung II	2.69	2.51	2.42	2.51	5.09	.006	.02
(Effekte der Schulinspektion: Wettbewerb zwischen den Schulen)	.76	.84	.77	.80			

Anmerkung. * Antwortskala: 1 „trifft überhaupt nicht zu" – 4 „trifft voll und ganz zu".

Quelle: Quantitative Befragung von Schulleitungen in vier Bundesländern (2011/2012) im Rahmen des BMBF-Projektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene"; eigene Berechnungen

Auch hier ist der Befund überraschend eindeutig: Es zeigen sich zwar signifikante, aber letztendlich kaum bedeutsame Differenzen in den Wirkungen der Schulinspektion bei Betrachtung des seit der Inspektion verstrichenen Zeitraumes. Die nach einer Schulinspektion wahrgenommenen Auswirkungen sind ähnlich randständig, wie zuvor im Ländervergleich dargestellt.

Tab. 11: Mehrebenen-Regressionen* auf Reaktionen der Schule und Effekte der Schulinspektion (Grundschulen und Gymnasien)

	Allgemeine Bewertung der SI		*Negative Folgen der SI*		*Effekte der SI auf Kooperation & SE*	
Prädiktor	b	p	b	p	b	p
Merkmale der SI						
Ziele 1: Schul- und Unterrichtsentwicklung	*.34*	.000	*-.23*	.000	*.20*	.000
Ziele 2: Wettbewerb zwischen Schulen	-.04	n.s.	*.09*	.002	.02	n.s.
Positive Wahrnehmung der Inspektoren	-.07	n.s.	-.07	n.s.	.03	n.s.
Qualität der SI	*.16*	.000	*-.12*	.040	.09	n.s.
Charakteristik der SI	*.42*	.000	*-.31*	.000	*.24*	.002
SI-Bericht	-.04	n.s.	.02	n.s.	-.06	n.s.
Externe Impulse und Unterstützung						
Unterstützung 1	-.05	n.s.	.01	n.s.	*.12*	.007
Unterstützung 2	*-.08*	.034	-.01	n.s.	.02	n.s.
Merkmale der Schule						
Kollegialer Zusammenhalt	-.05	n.s.	.04	n.s.	.-02	n.s.
Innovationsbereitschaft	.02	n.s.	*-.19*	.000	-.05	n.s.
Stellenwert Fortbildungen	.02	n.s.	-.01	n.s.	.04	n.s.
Maßnahmen QE	.02	n.s.	-.05	n.s.	.02	n.s.
ICC zwischen den Ländern	.08		.12		>.01	
R^2 (innerhalb der Länder)	.54		.49		.23	
R^2 (zwischen den Ländern)	.73		.81		.99	

Anmerkung. * Da keine signifikanten Kontexteffekte vorliegen, werden nur die Prädiktoren auf Schulleitungsebene dargestellt.

Quelle: Quantitative Befragung von Schulleitungen in vier Bundesländern (2011/2012) im Rahmen des BMBF-Projektes „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene"; eigene Berechnungen

In Tabelle 11 sind Befunde aus Mehrebenenanalysen (Random Coefficient Models) aufgeführt, da die vorliegenden Daten eine hierarchische Struktur aufweisen: Die befragten Schulleiterinnen und Schulleiter bilden dabei die Individualebene, während die Bundeslandzugehörigkeit (und damit die bundeslandspezifische Umsetzung der Schulinspektion) die Kontextebene bildet. Die abhängigen Variablen sind Teile der Modelldimensionen „Reaktionen der Schule" sowie „Intendierte Effekte der Inspektion" und wurden daher modellkonform mit den „Merkmalen der Schulinspektion", „Externen Unterstützungsmaßnahmen" und „Merkmalen der Schule" in Zusammenhang gebracht. Alle kursiv hervorgehobenen Koeffizienten verweisen auf

statistisch signifikante Beziehungen ($\alpha<0.05$) zwischen den abhängigen und unabhängigen Variablen.

Das Merkmal *allgemeine Akzeptanz der Schulinspektion* variiert nur geringfügig zwischen den untersuchten Bundesländern. Die Intraklassenkorrelation liegt bei rund 8 Prozent. Für die Erklärung der Variation innerhalb der Bundesländer (*within-group*) spielen vor allem die Merkmale der Inspektion selbst eine bedeutende Rolle. Wird die Schulinspektion als Instrument wahrgenommen, das auf Schul- und Unterrichtsentwicklung zielt sowie eine angemessene Beschreibung der Schulwirklichkeit vornimmt, so ist dies mit einer positiven Bewertung des Verfahrens verknüpft. Mit dem Einbezug der dargestellten Prädiktoren ist auf der Individualebene eine Varianzaufklärung von 54 Prozent verbunden. Die länderspezifisch variierenden Ausprägungen der Prädiktoren erklären auf der Ebene der Bundesländer bereits 73 Prozent der *Between-Group*-Varianz. Es lassen sich aus der Reihe der Prädiktoren keine auf Länderebene aggregierten Merkmale identifizieren, die im Sinne eines Kontexteffektes signifikant über der Erklärungskraft der Individualprädiktoren liegen. Dieser Befund ist dahingehend zu interpretieren, dass die Wahrnehmung der aufgeführten Merkmale von Schulinspektion zwar innerhalb der Länder bedeutsam variiert, aber keine systematischen Unterschiede in den aufgeführten Prädiktoren zwischen den Ländern erkennbar sind.

Die Analyse *negativer Folgen* von Schulinspektionen als abhängiger Variable zeigt zunächst eine höhere Variation zwischen den untersuchten Bundesländern. Der Intraklassenkoeffizient beträgt rund 12 Prozent. Betrachtet man die signifikanten Regressionskoeffizienten auf der Individualebene, so liefern die vorliegenden Ergebnisse folgendes Bild: Sobald Schulleiterinnen und Schulleiter das Gefühl haben, dass die Inspektionsverfahren auf eine Verbesserung von Schule und Unterricht zielen und die Stärken ihrer Schule angemessen abbilden, werden deutlich seltener negative Effekte der Inspektion berichtet. Aber auch die Innovationsbereitschaft des Kollegiums selbst trägt dazu bei, ob Schulleitungen das Inspektionsverfahren als zusätzliche Belastung erleben oder nicht. Sind Schulinspektionen mit einem Wettbewerbskalkül verbunden, so führt dies zu einer signifikant höheren Artikulation negativer Folgen der Schulinspektion bei den befragten Schulleitungen. Die Varianzaufklärung dieses Modells liegt bei rund 50 Prozent auf der Individualebene. Die zwischen den Schulleitungen variierenden Einstellungsmuster tragen zudem zu einer Varianzaufklärung auf der Länderebene bei (81 Prozent der *Between-Group*-Varianz), ohne dass bereits Kontexteffekte spezifiziert worden sind. Aber auch bei einer entsprechenden Modellierung von potentiellen Kontexteffekten lassen sich keine systematischen länderspezifischen Einflüsse erkennen, die sich signifikant auf die Wahrnehmung negativer Folgen der Schulinspektion auswirken.

Als letzte abhängige Variable sollen abschließend die *Effekte von Schulinspektion* betrachtet werden. Für diese Variable zeigt sich keinerlei signifikante Variation

zwischen den Bundesländern. Der entsprechende Intraklassenkoeffizient ist kleiner als 1 Prozent. Kontexteffekte sind für diese Dimension daher nicht zu erwarten. Auf der Individualebene ist jedoch ein interessanter Befund beobachtbar: Lediglich für diese abhängige Variable ist ein signifikanter Einfluss von externen Impulsen und Unterstützungsmaßnahmen erkennbar. Mit steigender wahrgenommener Unterstützung durch Schulaufsicht bzw. Schulamt steigt offensichtlich die Wahrscheinlichkeit der Initiierung von Schulentwicklungsmaßnahmen. Wie bereits bei der Darstellung der Akzeptanz von Schulinspektion beobachtbar, ist auch die Initiierung von Schulentwicklungsmaßnahmen positiv mit einer auf Schul- und Unterrichtsentwicklung abzielenden und die jeweiligen Gegebenheiten der Schule angemessen abbildenden Schulinspektion assoziiert. Insgesamt liefern die aus dem Ehren-Visscher-Modell abgeleiteten Prädiktoren für dieses Modell die schlechteste Varianzaufklärung. Diese liegt bei rund 23 Prozent und fällt damit um die Hälfte geringer aus als im ersten Modell.

7. Diskussion

Die diesem Beitrag zugrunde liegende Fragestellung, ob die länderspezifischen Ausprägungen von Schulinspektionsverfahren eine messbare Auswirkung auf die Wahrnehmung und Akzeptanz dieses Steuerungsimpulses aus der Sicht von Schulleitungen haben, kann insgesamt eher verneint werden. Es zeigen sich zwar in einigen der dargestellten Skalen signifikante Differenzen zwischen den Bundesländern Baden-Württemberg und Niedersachsen; diese Unterschiede sind aber mehrheitlich mit geringen praktischen Bedeutsamkeiten verbunden. Lediglich ein Muster zieht sich durch alle verwendeten Indikatoren: Die Schulleitungen in Niedersachsen äußern eine niedrigere Akzeptanz der Schulinspektion insgesamt, thematisieren häufiger die mit einer Inspektion verbundenen Belastungen und nehmen weniger externe Unterstützungen als ihre Kollegen und Kolleginnen in Baden-Württemberg, Mecklenburg-Vorpommern und Thüringen wahr. In der Darstellung der Spezifika der einzelnen Inspektionsverfahren in den vier untersuchten Bundesländern wird deutlich, dass Niedersachsen das einzige der einbezogenen Bundesländer ist, in dem das Instrument der Nachinspektion Verwendung findet. Im Zeitraum von der Mitte des Jahres 2005 bis Ende des Schuljahres 2007/2008 wurde in Niedersachsen für 67 Schulen eine Nachinspektion angesetzt (dies entspricht 6,2 Prozent aller in diesem Zeitraum inspizierten Schulen; vgl. Niedersächsische Schulinspektion 2008). Die mit den niedersächsischen Schulinspektionsverfahren für die Schulen einhergehende Option, als *„failing school"* identifiziert zu werden und daraufhin eine Nachinspektion zu erfahren, wirkt sich offenbar auf die gesamte Akzeptanz des Verfahrens und damit einhergehende Bewertungen von Seiten der Schulleitungen aus. Andererseits zeigt sich aber auch, dass die Androhung ernsthafter Konsequenzen (Nachinspektion bei Nichterreichen bestimmter Qualitätskriterien) nicht zu messbar höheren Effekten

im Nachgang einer Inspektion führt. In allen vier untersuchten Bundesländern sehen die Schulleitungen kaum Auswirkungen der Schulinspektion auf Schul- und Unterrichtsentwicklungsmaßnahmen. Die Ausprägungen der entsprechenden Skalen zeigen nicht nur ein unterdurchschnittliches Niveau, sondern auch keinerlei signifikante Varianz zwischen den vier untersuchten Bundesländern. Auf die Ergebnisse einer Schulinspektion bezogene Schul- und Unterrichtsentwicklung findet offensichtlich in der Praxis nicht statt. Die niedersächsischen Schulleitungen zeigen in dieser Dimension keine auffällig abweichenden Werte vom Gesamtmittelwert. Etwas differenzierter erscheint das Bild, wenn nur Schulen betrachtet werden, die einer Nachinspektion unterzogen wurden. Im Report der Niedersächsischen Schulinspektion (ebd.) wird von 14 vollzogenen Nachinspektionen im Berichtsjahr 2008 berichtet. Im Rahmen der Nachinspektion konnte festgestellt werden, dass sich 12 der 14 nachinspizierten Schulen soweit verbessert hatten, dass sie zumindest die Norm für eine ausreichende Schulqualität überschreiten konnten. Inwieweit hier gesonderte Unterstützungsmaßnahmen oder sensibilisierte Bewertungen der Inspektoren und Inspektorinnen wirksam waren, kann dem Bericht der Niedersächsischen Schulinspektion nicht entnommen werden.

Insgesamt erstaunt es in den deutschsprachigen Untersuchungen (vgl. u.a. Landwehr/Huber 2008; Gärtner/Hüsemann/Pant 2009; Sommer 2011; Schwank/Sommer 2012), dass die Akzeptanz der Schulinspektionsverfahren insgesamt positiv ausfällt, während in angelsächsischen Studien eher die Belastungen und der Kontrollaspekt von Schulinspektionen thematisiert werden (vgl. z.B. Brimblecombe/Ormston/Shaw 1996; Cuckle/Broadhead 1999; O'Connor 2001). Selbst im Bundesland Niedersachsen – mit der dort praktizierten Option der Nachinspektion – überwiegen bei den Schulleitungen die positiven Einschätzungen des Inspektionsverfahrens und der darin involvierten Inspektoren und Inspektorinnen. Die auch in anderen Studien bereits sichtbaren Zusammenhänge zwischen der Passung von Inspektionsergebnissen und dem Selbstbild der Schule einerseits und einer damit einhergehenden Akzeptanz und Wertschätzung von Schulinspektionen andererseits lassen sich auch anhand der hier dargestellten Befunde replizieren. Weiterhin tragen eine positive Wahrnehmung des Inspektorenteams und die von den Schulleitungen empfundene Angemessenheit des Inspektionsberichtes zur Anerkennung von Schulinspektionsverfahren bei. Als problematisch ist die kaum wahrgenommene externe Unterstützung durch die Schulbehörden im Nachgang von Schulinspektionen oder externen Leistungsvergleichen zu bewerten. Alle untersuchten Schulinspektionsverfahren zielen in ihrem Selbstverständnis letztendlich auf eine Anregung der Verbesserung schulischer Arbeit. Dieser intendierte Effekt ist anhand der hier berichteten Daten nicht nachweisbar. Inwieweit auch die in den Augen der Schulleitungen defizitären Unterstützungsleitungen durch die Schulbehörden in diesem Zusammenhang eine Rolle spielen, wird Gegenstand weiterer Forschungsbemühungen unseres Verbundprojektes sein.

Die Grenzen der vorliegenden Untersuchung sind in zwei Bereichen zu verorten: Zum einen kann eine methodisch angemessene Überprüfung der Kausalwirkungen im Ehren-Visscher-Modell (2006) nur über experimentelle oder zumindest quasi-experimentelle Designs gewährleistet werden. Dies ist aufgrund des Forschungsgegenstandes jedoch mittlerweile ausgeschlossen. Bereits 2011 konstatiert Sommer: „[…] der Nachweis kausaler Zusammenhänge ist schwierig, das Zeitfenster für einen quasi-experimentellen Vergleich der Entwicklungen in den Ländern mit und ohne Schulinspektion hat sich in Deutschland nahezu geschlossen" (Sommer 2011, S. 138). Im Rahmen dieser Studie ist es nur möglich, bereits inspizierte Schulen mit noch nicht inspizierten Schulen innerhalb eines Bundeslandes zu vergleichen. Dieser Vergleich kann vor allem im Bundesland Baden-Württemberg erfolgen, da hier zum Zeitpunkt der Befragung rund 74 Prozent der Schulen in der Stichprobe noch nicht inspiziert waren. Beim Vorliegen der Daten aus der zweiten Befragungsrunde können dann Schulen identifiziert werden, in denen die Schulinspektion zwischenzeitlich erfolgte.

Zum zweiten entspricht der im Rahmen der Online-Erhebung erfolgte Rücklauf nicht unseren Erwartungen. Es ist daher mit Selbstselektionseffekten zu rechnen, die die Generalisierbarkeit der gewonnenen Ergebnisse in Frage stellen.[2] Leider besteht noch zu wenig Wissen darüber, ob und welche systematischen Motivlagen bei Schulleitungen dazu führen, sich reflexiv mit Schulinspektionsverfahren auseinanderzusetzen oder nicht. Vor diesem Hintergrund erscheint es auch nicht weiterführend, durch Matching-Verfahren (z.B. *Propensity Score Matching*) statistisch äquivalente Vergleichsgruppen zu bilden, wenn theoretisch oder empirisch fundiertes Wissen über entsprechende Kovariaten fehlt.

Trotz dieser Beschränkungen der vorliegenden Studie halten wir die Ergebnisse für geeignet, einen Vergleichshorizont (vor allem für die qualitativen Studien unseres Projektverbundes) aufzuspannen, in dem sowohl beabsichtigte als auch nicht-intendierte Effekte bzw. Nebenwirkungen der Schulinspektionsverfahren sichtbar werden.

Literatur und Internetquellen

Bildungsserver Mecklenburg-Vorpommern. URL: http://www.bildung-mv.de; Zugriffsdatum: 04.01.2013.

Bortz, J./Döring, N. (1995): Forschungsmethoden und Evaluation für Sozialwissenschaftler. Berlin: Springer.

Brimblecombe, N./Ormston, M./Shaw, M. (1995): Teachers' Perceptions of School Inspection: A Stressful Experience. In: Cambridge Journal of Education 25, H. 1, S. 98-108.

2 Das Problem der Selbstselektion wäre auch bei der Ziehung von Zufallsstichproben aufgetreten, da sich gerade zu Beginn der Etablierung von Schulinspektionsverfahren die Schulen freiwillig für die Durchführung einer Inspektion zur Verfügung gestellt haben.

Brimblecombe, N./Ormston, M./Shaw, M. (1996): Teachers' Perceptions of Inspections. In: Ouston, J./Earley, P./Fidler B. (Hrsg.): Ofsted-Inspections: the Early Experience. London: Fulton, S. 126-134.

Brimblecombe, N./Shaw, M./Ormston, M. (1996): Teachers' Intention to Change Practice as a Result of OFSTED School Inspections. In: Educational Management & Administration 24, H. 4, S. 339-354.

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".

Croxford, L./Gray, J./Ozga, J. (2009): Teacher Attitudes to Quality Assurance and Evaluation (QAE) in Scotland and England. CES Briefing No. 51. Edinburgh: Centre for Educational Sociology, University of Edinburgh.

Cuckle, P./Broadhead, P. (1999): Effects of Ofsted Inspection on School Development and Staff Morale. In: Cullingford, C. (Hrsg.): An Inspector Calls: Ofsted and Its Effect on School Standards. London: Kogan, S. 176-187.

Dedering, K./Müller, S. (2011): School Improvement Through Inspections? First Empirical Insights from Germany. In: Journal of Educational Change 12, H. 3, S. 301-322.

Döbert, H./Rürup, M./Dedering, K. (2008): Externe Evaluation von Schulen in Deutschland. In: Döbert, H./Dedering, K. (Hrsg.): Externe Evaluation von Schulen. Münster u.a.: Waxmann, S. 63-151.

Ehren, M.C.M./Visscher, A.J. (2006): Towards a Theory on the Impact of School Inspections. In: British Journal of Educational Studies 54, H. 1, S. 51-72.

Fidler, B./Earley, P./Ouston, J. (1995): Ofsted School Improvement through Inspection. Vortrag: European Conference on Educational Research, Bath University, Sept.

Fidler, B./Ouston, J./Earley, P. (1994): Ofsted Inspections and Their Contribution to School Development. Vortrag: Annual Conference of the British Educational Research Association, Oxford University, Sept.

Gärtner, H./Hüsemann, D./Pant, H.A. (2009): Wirkungen von Schulinspektion aus Sicht betroffener Schulleitungen. Die Brandenburger Schulleiterbefragung. In: Empirische Pädagogik 23, H. 1, S. 1-18.

Geijsel, F. (2001): Schools and Innovations: Conditions Fostering the Implementation of Educational Innovations. Nijmegen: University Press.

Gray, C./Gardner, J. (1999): The Impact of School Inspections. In: Oxford Review of Education 25, H. 4, S. 455-468.

Husfeldt, V. (2011a): Externe Schulevaluation. Ein länderübergreifender Blick auf Forschungen und Modelle. In: Quesel, C./Husfeldt, V./Landwehr. N./Steiner P. (Hrsg.): Wirkungen und Wirksamkeit der externen Schulevaluation. Bern: hep, S. 13-33.

Husfeldt, V. (2011b): Wirkungen und Wirksamkeit der externen Schulevaluation. Überblick zum Stand der Forschung. In: Zeitschrift für Erziehungswissenschaft 14, H. 2, S. 259-282.

Koch, U./Groß Ophoff, J./Hosenfeld, I./Helmke, A. (2006): Von der Evaluation zur Schul- und Unterrichtsentwicklung – Ergebnisse der Lehrerbefragungen zur Auseinandersetzung mit den VERA-Rückmeldungen. In: Eder, F./Gastager, A./Hofmann, F. (Hrsg.): Qualität durch Standards? Beiträge zur 68. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF), Salzburg. Münster u.a.: Waxmann, S. 187-199.

Kogan, M./Maden, M. (1999): An Evaluation of Evaluators: The OFSTED System of School Inspection. In: Cullingford, C. (Hrsg.): An Inspector Calls: Ofsted and Its Effect on School Standards. London: Kogan, S. 9-32.

Kotthoff, H.-G./Böttcher, W. (2010): Neue Formen der „Schulinspektion": Wirkungshoffnungen und Wirksamkeit im Spiegel empirischer Bildungsforschung. In: Altrichter,

H./Maag Merki, K. (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS, S. 295-325.
Landesbildungsserver Baden-Württemberg. URL: http://www.schule-bw.de; Zugriffsdatum: 04.01.2013.
Landwehr, N./Huber, A. (2008): Die Unterrichtsevaluation als Teil der externen Schulevaluation. Eine Bestandsaufnahme. Aarau: Institut Forschung und Entwicklung der Pädagogischen Hochschule Nordwestschweiz, Zentrum Bildungsorganisation & Schulqualität.
Leeuw, F.L. (2002): Reciprocity and Educational Evaluations by European Inspectorates: Assumptions and Reality Checks. In: Quality in Higher Education 8, H. 2, S. 137-150.
Niedersächsische Schulinspektion (2008): Periodischer Bericht 2008. Bad Iburg.
Niedersächsischer Bildungsserver. URL: http://www.nibis.de; Zugriffsdatum: 04.01.2013.
O'Connor, P.P. (2001): The Policy, Process and Impact of Whole School Inspection at Primary Level in the Republic of Ireland from the Perspective of some Inspectors and Teachers. EdD Thesis, The Open University.
Reezigt, G.J./Creemers, B.P. (2005): A Comprehensive Framework for Effective School Improvement. In: School Effectiveness and School Improvement 16, H. 4, S. 407-424.
Rürup, M. (2008): Typen der Schulinspektion in den deutschen Bundesländern. In: Die Deutsche Schule 100, H. 4, S. 467-477.
Schwank, E./Sommer, N. (2012): Wirkung der Schulinspektion anhand der Wahrnehmung der Lehrkräfte. Ergebnisse einer Befragung im Rahmen der Inspektionsevaluation. In: Schulverwaltung Niedersachsen 23, H. 4, S. 106-110.
Sommer, N. (2011): Wie beurteilen schulische Gruppen die erlebte Schulinspektion? Ergebnisse einer Befragung. In: Müller, S./Pietsch, M./Bos, W. (Hrsg.): Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht. Münster u.a.: Waxmann, S. 137-164.
Standaert, R. (2000): Inspectorates of Education in Europe: a Critical Analysis. Flanders, Ministry of Education. Leuven: Acco.
Thüringer Schulportal. URL: http://www.schulportal-thueringen.de; Zugriffsdatum: 04.01. 2013.
Wilcox, B./Gray, J. (1996): Inspecting Schools: Holding Schools to Account and Helping Them to Improve. Buckingham: Open University Press.

Oliver Böhm-Kasper, Prof. Dr., geb. 1972, Universitätsprofessor in der AG 9: Medienpädagogik, Forschungsmethoden und Jugendforschung, Fakultät für Erziehungswissenschaft an der Universität Bielefeld.
E-Mail: oliver.boehm-kasper@uni-bielefeld.de

Odette Selders, Dipl.-Päd., geb. 1981, wissenschaftliche Mitarbeiterin in der AG 9: Medienpädagogik, Forschungsmethoden und Jugendforschung, Fakultät für Erziehungswissenschaft an der Universität Bielefeld.
E-Mail: odette.selders@uni-bielefeld.de

Anschrift: Universität Bielefeld, Fakultät für Erziehungswissenschaft, Universitätsstraße 25, 33615 Bielefeld

DDS – Die Deutsche Schule
Beiheft 12, S. 154-171
© 2013 Waxmann

Bianca Preuß

Akteurkonstellation zwischen Schulträger und Schule

Empirische Analysen zur Governance von Schulinspektion

Zusammenfassung
Der vorliegende Artikel widmet sich dem Schulträger als bildungspolitischem Akteur in seiner Governance-Beziehung zur Schule im Kontext der Einführung von Schulinspektionen in vier verschiedenen Bundesländern. Rekurriert wird dabei auf länderübergreifende Daten im Rahmen des BMBF-Projekts „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene" (2010-2013) für das Teilprojekt der Universität Gießen. Vor dem Theorie- und Methodenhintergrund der Educational-Governance-Forschung wird die Akteurkonstellation zwischen Schulträger und Schule beobachtet und damit einer der blinden Flecken der empirischen Bildungsforschung erhellt. Als wesentlicher Befund zeichnet sich ab, dass der Schulträger seine Einflussmöglichkeiten und -chancen als Steuerungsakteur im Zusammenhang von Schulinspektion reflektiert, hinterfragt sowie durch spezifische Formen der „Grenzüberschreitung" potenziell erhöht.
Schlüsselwörter: Akteurkonstellation, Schulträger, Schulinspektion, BMBF-Projekt, schulische Governance

The Constellation of Actors between Local Education Agencies and Schools

Empirical Analyses with Regard to the Governance of School Inspection

Abstract
This article explores the governance relationship between local education agencies as education political actors and schools with regard to the introduction of school inspections in four different German states. It is based on data from the across-federal-states BMBF-project "School-Inspections as Control Impulses for School Development and Its Implementation on the Individual School Level" (2010-2013), which was partly accomplished by the University of Gießen. The constellation of the actors from education agen-

cies and schools will be documented on the theoretical and methodical background of Educational Governance. So it will explore one of the "blind spots" of educational research. The essential result is that the local education agencies are reflecting their assertiveness and opportunities as an actor of governance in the context of school-inspection. In addition, they increase their impact and opportunities through specific forms of "transgression".
Keywords: constellation of actors, local education agencies, school-inspection, BMBF-project, educational governance

1. Governancetheoretische Eingangsbemerkungen

Mit der Einführung der Schulinspektion in den Ländern als Element schulischer Qualitätssicherung (vgl. u.a. Dedering 2012; Bos u.a. 2007) wurde die formale Steuerungs- und Legitimationsarchitektur im Mehrebenensystem Schule umgebaut. Zum einen wurde dem bestehenden Gefüge der Steuerung von Schulsystemen ein neuer Akteur aufgesetzt, mit dem die anderen Akteure umgehen und zu dem sie sich positionieren müssen. Zum anderen erhielten die Schulen veränderte bzw. neue Rechenschaftspflichten. Neben bzw. zusätzlich zu internen Selbstevaluationsprozessen und -erfahrungen sollen sie mit datenbasierten Berichten umgehen, die schulische Qualität von externer Seite spiegeln. Der Schulaufsicht wird nun die neue Aufgabe abverlangt, mit den Schulen auf Basis der Inspektionsberichte Zielvereinbarungen durchzuführen. Die traditionell intermediäre Rolle der Schulaufsicht als ein Akteur der Bildungsverwaltung, die Aufsicht und Beratung an die Schule vermittelt, differenziert sich zu einer triadischen Governance-Beziehung zwischen Schulinspektion, Aufsicht und Schule. Der Schulträger wiederum wird durch Schulinspektion zwar adressiert, weil er in der Regel von Seiten der Schulaufsicht oder der Schulen den Inspektionsbericht erhält, er in Form von Interviews befragt wird oder an schulischen Begehungen teilnimmt,[1] ihm wird jedoch darüber hinaus keine explizite Rolle bzw. Aufgabe innerhalb von Schulinspektion zugedacht. Er steht, wie die Eltern, mittelbar im Gefüge von Schulinspektion als neue institutionalisierte Form der Handlungskoordination. Sowohl Bildungswissenschaften als auch Politik und schulische Steuerungspraxis interessieren sich dafür, ob die mit dieser Transformationsstrategie und den evaluationsbasierten Steuerungsinstrumenten verbundenen Reformerwartungen und Optimierungsversprechen auf allen Makro- und Mikroebenen des schulischen Mehrebenensystems erfüllt werden (vgl. Wacker/Maier/Wissinger 2011, S. 9). Die Einführung von Schulinspektion steht deutlich im

1 Da die bildungspolitische Reform *Schulinspektion* in föderalistische Strukturen eingebettet ist, gilt auch für den Schulträger, dass seine Rolle innerhalb von Schulinspektion zwischen den Bundesländern divergieren kann und er unterschiedlich in den Evaluationsprozess involviert ist. Zum Beispiel wird die Praxis, den Inspektionsbericht von Seiten der Schulen an den Schulträger zu geben, in den einzelnen Ländern unterschiedlich gehandhabt: Zum Teil geschieht dies auf freiwilliger Basis, zum Teil aber aus Verpflichtung.

Zusammenhang der 2001 durch die Veröffentlichung der PISA-Studie ausgelösten Qualitätsdiskussionen im deutschen Schulwesen (vgl. Rürup/Lambrecht 2012).

Anzunehmen ist, dass sich im Zuge der Einführung neuer Steuerungsinstrumente auch die vielfältigen Interdependenzen der in dem durch Komplexität gekennzeichneten Mehrebenensystem ‚eigensinnig' handelnden Akteure verändern. Eine solche Forschungsperspektive u.a. zur Schulinspektion verfolgt die Educational-Governance-Forschung (vgl. Kussau/Brüsemeister 2007; Altrichter/Heinrich 2007). Die bisherige Schulinspektionsforschung hat sich überwiegend damit beschäftigt, ob die in den bildungspolitischen Programmatiken formulierten Zielsetzungen des Instruments in der schulischen Praxis erreicht werden (vgl. Rürup/Lambrecht 2012, S. 172). Sie kommt national und international aber weder zu eindeutigen noch kohärenten Ergebnissen (vgl. u.a. Bruggen 2006; Böttcher/Kotthoff 2007; Müller 2010; Müller/Pietsch/Bos 2011; Husfeldt 2011) und stellt eher Nebenwirkungen als Wirkungen fest (vgl. Kotthoff/Böttcher 2010). Es bedarf damit einer Vertiefung und Fundierung der Evaluations- und Wirkungsforschung, vor allem im Hinblick auf detaillierte Prozessanalysen, die untersuchen, wie Schulinspektionsimpulse in Form von Schulentwicklung verarbeitet werden und welche „Nebenwirkungen" entstehen zwischen dem externen Impuls von Schulinspektion und den internen Reaktionen und Verarbeitungen auf der Ebene der Einzelschule.

Mit ihrem analytischen Potenzial der Beobachtung, wie sich Reformkonzepte in institutionalisierten Akteurkonstellationen durchsetzen, bietet die Governance-Perspektive in diesem Kontext einen besonderen Forschungsansatz. Sie setzt paradigmatisch voraus, dass bildungspolitische Reformen eingebettet sind in wechselseitige Handlungskoordinationen im Mehrebenensystem (vgl. Benz u.a. 2007). Aus dieser Perspektive erhält jeder einzelne Akteur Subjektstatus mit eigenen Handlungskapazitäten und -rationalitäten, was bedeutet, dass Steuerungsprozesse auf allen Ebenen des Mehrebenensystems gebrochen werden können (vgl. Wacker/Maier/Wissinger 2011, S. 23). Ein so zu verstehender Steuerungsbegriff führt zur Unterscheidung von Gelingen/Scheitern, indem er nahelegt, „Abweichungen" aus Sicht des jeweiligen steuernden Akteurs zu beobachten (vgl. Kussau/Brüsemeister 2007). Probleme der Implementierung neuer Steuerungsinstrumente werden in der Governance-Perspektive über einzelne Akteure und ihre Koordinierungsprozesse beobachtet.

Auf den Schulträger und dessen institutionalisierte Interaktionsbeziehungen ist in der Governance-Forschung und auch in der empirischen Bildungsforschung bislang kaum geschaut worden. Darauf weist etwa Schnell (vgl. 2007, S. 262) für den Blickwinkel regionaler Governance hin. Er hebt hervor, dass in der Bildungsforschung kommunale Schulträger als Akteure zwar gesehen werden, explizite Forschungen aber fehlen.

2. BMBF-Forschungsprojekt und Methode

Im Rahmen des vorliegenden Beitrags soll ein Ausschnitt aus den vielfältigen Konstellationen zwischen verschiedenen Akteuren im schulischen Mehrebenensystem, nämlich die Akteurkonstellation zwischen Schulträger und Schule, länderübergreifend vorgestellt werden. Damit wird einer der ‚blinden Flecken' in der Schulinspektionsforschung (vgl. Rürup/Lambrecht 2012, S. 172) und auch in der Bildungsforschung beleuchtet. Dass dem Schulträger als bildungspolitischem Akteur generell nur wenig Aufmerksamkeit gewidmet wird, mag mit dem Tenor der gegenwärtigen Bildungsforschung zu tun haben, die traditionell auf die Analyse von einzelnen Akteuren abstellt und nicht auf Konstellationsgefüge, Mehrebenensysteme oder Governance-Regime. Zweifelsohne ist auch die Einbettung des Akteurs Schulträger in institutionalisierte Akteurkonstellationen nicht von der Hand zu weisen.

Rekurriert wird auf Daten, die im Rahmen des BMBF-Projekts „Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene" in insgesamt 52 qualitativen Experteninterviews erhoben wurden (vgl. Brüsemeister u.a. 2009)[2]. Am Verbund beteiligt sind die Universitäten Gießen, Bielefeld, Hannover und Hamburg. Übergreifend für die einzelnen Standorte wird die Frage nach dem Zusammenhang zwischen der Einführung von Schulinspektion und Schulentwicklung untersucht. Die hier vorgelegten Befunde basieren auf Zwischenergebnissen des Teilprojekts der Universität Gießen, die auf Interviews mit Schulleitungen, Lehrkräften, Eltern, Schulträgern, Schulaufsicht und Inspektionsteams in vier verschiedenen Bundesländern zurückgehen. Untersucht wurde jeweils eine Schule pro Bundesland. Nach der Erhebung und Transkription wurden die Interviews thematisch kodiert (vgl. Flick 2007), in Fallvignetten zusammengefasst sowie anhand institutionenanalytischer Aspekte u.a. nach Scott (2008) governancetheoretisch ausgewertet.

3. Der Schulträger in der Governance von Schulinspektion

Der Schulträger ist ein regionaler Akteur bzw. lokaler Intermediär. Im Kontext von Schulinspektion teilt er sich mit der Schulaufsicht, der Schule und dem Inspektionsteam eine Ebene mit regionalen Bezügen. Die Schule wiederum ist Teil der örtlichen Gemeinschaft, hat sowohl staatliche als auch kommunale Akteureigenschaften (letztere z.B. den Standortfaktor und Sozialraum betreffend). Sie ist eine nicht rechtsfähige Anstalt des Landes und ihres Schulträgers.

Die kommunalen Schulträger verwalten ihre Schulangelegenheiten als pflichtige Selbstverwaltungsaufgabe. Sie sind die Körperschaft, die das Schulgebäude errich-

2 Siehe URL: http://www.verbundprojekt-schulinspektion.de/; Zugriffsdatum: 22.07.2013.

tet und für die Ausstattung mit sächlichen Mitteln zuständig ist. In dieser Funktion leistet der Schulträger einen Beitrag zur Realisierung der in den Schulgesetzen der Länder verankerten Bildungs- und Erziehungsziele und hat im Rahmen der zugewiesenen Aufgaben in Abstimmung mit den Kultusministerien, den örtlichen Gremien und den Schulleitungen Entscheidungsbefugnisse.

> „Der Schulträger ist für die Errichtung, Unterhaltung und Verwaltung der Schule verantwortlich und trägt in der Regel die Sachkosten (während die Personalkosten für Lehrer an öffentlichen Schulen vom Land übernommen werden)" (Sekretariat der Kultusministerkonferenz 2012).

Die Schule dagegen hat Entscheidungshoheit über die inhaltliche Ausgestaltung und operative Umsetzung der pädagogischen Ziele. Beide Akteure, Schulträger und Schule, stehen unter dem Prinzipal Staat, wobei die Schule als ‚nicht rechtsfähige Anstalt' ein in gewissem Maße verselbständigter Teil der Organisation ihrer Träger ist und somit abhängig bleibt von den materiellen Ressourcen ihres kommunalen Schulträgers und den inhaltlichen und personellen Vorgaben des Staates (vgl. Bräth/Eickmann/Galas 2009, S. 54).

Konflikte mit Auswirkungen auf die pädagogischen Handlungsfähigkeiten der Schulen materialisieren sich insbesondere im kostenintensiven Personalbereich, in der Frage, welche Aufgaben das vom Land oder das vom Schulträger zu finanzierende Personal wahrnehmen darf/muss. Beispielhaft seien an dieser Stelle zwei Themen bzw. Fragestellungen genannt, die in Niedersachsen auch in der Öffentlichkeit und/oder im Landesparlament behandelt wurden. Die Bestimmungen des dortigen Schulgesetzes zur „Aufbringung der Kosten"[3] scheinen die Interpretationen zuzulassen, dass die Betreuung von Computern und Computernetzwerken in Schulen vom Schulträger zu gewährleisten ist (als nicht pädagogische Aufgabe) oder aber von den vom Land zu bezahlenden Lehrpersonen (da Schüler und Schülerinnen auch didaktisch an die Arbeit mit Computern herangeführt werden sollen).[4] Ähnliche Abgrenzungsprobleme lassen sich finden bezüglich der Verwaltungsaufgaben in den Schulen, die von Schulsekretärinnen oder Lehrkräften sowie Schulleiterinnen und Schulleitern zu erledigen sind.[5] Zur Optimierung der offensichtlich unbefriedigenden schulischen Arbeitssituation wird „nach tragfähigen Lösungen" gesucht (so die

3 Vgl. §§ 112 und 113 des Niedersächsischen Schulgesetzes (NSchG in der Fassung vom 3. März 1998, zuletzt geändert durch das Gesetz vom 12. Juli 2007).

4 Vgl. Hannoversche Allgemeine Zeitung (HAZ): „Kein Netz. Dass Schulen Computer brauchen, ist Konsens – nur wer sich um sie kümmern soll, weiß niemand. Ein Hilferuf" (Nr. 20 vom 24. Januar 2013): Niedersächsischer Landtag, Stenografischer Bericht der 127. Sitzung am 20.01.2012, Anlage 13, S. 16441f.

5 Siehe Landtagsdrucksache 16227 (Antwort der Landesregierung vom 29.05.2008 auf eine Kleine Anfrage): Niedersächsischer Landtag, Stenografischer Bericht der 127. Sitzung am 20.01.2012, Anlage 20, S. 16451f.

Antwort des Niedersächsischen Kultusministeriums am 20.01.2012 auf eine parlamentarische Anfrage)[6].

Durch die Einführung von Schulinspektion werden die exemplarisch dargestellten Unstimmigkeiten zwischen Schulträger und Land nicht unmittelbar berührt, da die Kostenlastverteilung des Schulgesetzes unangetastet bleibt. Allerdings könnten Konflikte zwischen Schulträger und Schule dann entstehen, wenn Schulen auf Basis des Inspektionsberichtes kostenträchtige Entwicklungen initiieren möchten.

Gegenwärtig wird insbesondere im Kontext neu ausgerufener regionaler Bildungslandschaften das strukturelle Faktum getrennter Verantwortlichkeiten zwischen Schulträger und Schule als unbefriedigend wahrgenommen. In jenen soll eine Gesamtqualität des Bildungssystems dann entstehen, wenn die Handlungskoordination zwischen Ländern, Schulträger und Schule über das bisherige Maß hinausgeht. Dies beinhaltet bildungsbezogene Handlungsabstimmungen und ein integriertes Bildungsmanagement zwischen Schulträger und Schule, in dem Bildung als (gemeinsamer) Standortfaktor gesehen wird. An der grundsätzlichen Problematik zwischen den „inneren und äußeren Schulangelegenheiten" hat sich dadurch jedoch nichts geändert (vgl. z.B. Bertelsmann Stiftung 2012).

Schulinspektionen zielen als vom Land extern gesetzter Steuerungsimpuls auf die einzelne Schule als Handlungseinheit. Konzipiert werden aber auch regionale Akteure wie Schulträger und Schulaufsicht als nutznießende und mitverantwortliche Akteure eines gelingenden Schulinspektionsprozesses. Sie sind stellvertretende *stakeholder* des Staates, der Gesellschaft, der Eltern und haben eigenständige Interessen: im Falle des Schulträgers etwa, Chancengleichheiten in der Region zu verbessern, schulische Bildungsabschlüsse zu erhöhen oder den Elternbedürfnissen mit ausreichenden Kindergarten- und Kindertagesstätten- bzw. Ganztagsschulangeboten gerecht zu werden.

Ob sich durch die Einführung von Schulinspektion „systematische Grenzüberschreitungen" (vgl. Kussau/Brüsemeister 2007) ergeben, sollen untenstehende empirische Analysen anhand der Projektdaten aus dem BMBF-Verbund klären. Den Fokus auf Grenzüberschreitungen zu setzen, ist charakteristisch für die Governance-Perspektive. Mit Rekurs auf Emmerich (vgl. 2010, S. 357f.) interessiert in diesem Zusammenhang die Frage nach der Handlungsfähigkeit von Akteuren im regionalen Kontext, nach Formen der Interdependenzbewältigung zwischen ihnen sowie Verkoppelungen zwischen intermediären und schulischen Steuerungsebenen durch die Einführung der Schulinspektion.

6 Vgl. Fußnote 5.

4. Empirische Befunde: Institutionenanalytische Aspekte

Die Analyse der Daten aus dem BMBF-Projekt basiert auf Scott (2008), da sein Ansatz sowohl institutionenbezogen als auch sozialwissenschaftlich angelegt ist. Eigenschaften von Institutionen entstehen nach Scott (vgl. ebd., Kapitel 3, S. 47-71) erst in der Relation zum sozialen Handeln. Institutionen konstituieren sich explizit und im Besonderen durch soziale Interaktionen, die in der ‚Governance-Sprache' Handlungskoordinationen genannt werden. Allerdings, so räumt Scott (vgl. ebd., S. 22) ein, folgen die sozialen Akteure in Institutionen meist bürokratischen Mustern und Routinen, d.h. soziales Handeln in Institutionen findet in erster Linie als Anweisung der Hierarchie statt: „[…] people do what they are supposed to do" (ebd.).

Untersucht wird im Folgenden, wie sich die Akteurkonstellation zwischen Schulträger und Schule institutionenanalytisch gestaltet. Dies erfolgt in Bezug auf drei Aspekte, die Scott (2008) für die Analyse von Organisationen differenziert: 1) Mit dem *„normative pillar"* meint Scott normative Regeln wie Routinen, Handlungsabläufe, Konventionen, Rollen, Strategien, organisationale Formen, Überzeugungen, Paradigmen, Kodes, Kulturen oder Wissen. 2) Der *„regulative pillar"* betrachtet Individuen und Organisationen, die Regelsysteme konstruieren, um ihre eigenen Interessen und Handlungsweisen zweckdienlich und instrumentell zu bestimmen. 3) Mit dem *„cultural-cognitive pillar"* einer Institution fokussiert Scott geteilte Bedeutungen und Bedeutungsgebungen (z.B. Leitideen), die soziale Realität konstruieren.

Methodisch begründet sich dieses spezifische analytische Vorgehen in der Educational-Governance-Perspektive, die Organisationen als Akteurkonstellationen betrachtet.

4.1 Der normative Aspekt

Mit dem *„normative pillar"* nach Scott (2008) wird das normative Rollenhandeln des Schulträgers im Kontext von Schulinspektion fokussiert. Analytisch interessiert dabei im Besonderen, wie er sein Rollenhandeln definiert und eingrenzt und an welchen Stellen er sich durch die Einführung der Schulinspektion konkret adressiert (bzw. nicht adressiert) sieht. Vor diesem Hintergrund zeigen die empirischen Daten, dass der Schulträger aus dem Anspruch des Schulgesetzes ein klares Rollenhandeln ableitet:

> *[…] das Schulgesetz definiert ja auch die Aufgaben des Trägers, das heißt eben technische, materiale Voraussetzungen zu stellen* (B_ST_1, S. 3, Z. 15-17).[7]

Entsprechend der präskriptiven Dimension des Normativen, die dem Schulträger keinen Einfluss auf die Bildungsinhalte von Schule zugesteht, folgert dieser, dass er im Kontext von Schulinspektion nur dann tangiert sei, wenn es um die Fragen geht, *„[…] wie ist das Schulgebäude, wie ist die Ausstattung, ist das entsprechend oder gibt es da vielleicht noch Handlungsbedarfe […]?“* (D_ST_1, S. 2, Z. 16-21) In einem Interview mit einer Schulleitung erwähnt diese, dass der Schulträger in einem Gespräch betonte, die Zuständigkeitsbereiche von Schule und Schulträger seien nach *„Software“* und *„Hardware“* getrennt (vgl. A_SL_1, S. 26, Z. 15-18). Wie sich im weiteren Verlauf des Interviews herausstellt, trennt damit der Schulträger auch im Zusammenhang von Schulinspektion den Aufgabenbereich zwischen Schule und Schulträger. Er klammert in puncto „Software“ alles aus, was mit Didaktik und Unterricht zu tun hat, und hält sich an das Gesetz, das für den Schulträger pädagogische Zuständigkeiten nicht vorsieht.

Im Zusammenhang von Schulinspektion aber wird der Schulträger mit pädagogischen Inhalten des Inspektionsberichts konfrontiert. Aus dem normativen Aspekt heraus interessiert sich der Schulträger nicht dafür. Er sieht sich angesprochen, wenn es um das Thema Ausstattung und Räume geht (vgl. A_ST_2, S. 11, Z. 27f.). Dem nach pädagogischen Kriterien abgefassten Inspektionsbericht schreibt er Informationen für Bereiche zu, für die er nicht zuständig ist (vgl. ebd., S. 6, Z. 150f.) und kommt zu der Aussage, *„[…] es bringt im Moment nix, weil wir sowieso nix machen können“* (D_ST_1, S. 17, Z. 8). So sieht sich der Schulträger nicht als Steuerungsakteur, sondern in einer *„untergeordneten Rolle“* (B_ST_1, S. 21, Z. 22f.). Schulinspektion ist folglich nicht sein *„Thema“*:

> *[…] die Wahrnehmung über [Schulinspektion] auf Seiten des Schulträgers, ich wills mal so sagen die geht doch eigentlich gegen Null […]. Wir nehmen das nicht wahr, das ist nicht unser Thema, das ist ein Thema was durch das Kultusministerium möglicherweise hier bestimmt wird, aber nicht zu unserm Thema gemacht worden ist* (D_ST_1, S. 19, Z. 12-14 u. 14-17).

Deutlich wird an dieser Stelle die affektive Komponente des normativen Aspekts nach Scott (2008): Der Schulträger reagiert beinahe gekränkt auf die Tatsache, dass ihn die Bildungsverwaltung nicht als Steuerungsakteur adressiert.

---

7 Sämtliche im vorliegenden Artikel angeführten Interviewzitate wurden geglättet und nach Personen, Geschlecht, Institution und Länderzugehörigkeit anonymisiert. Der Begriff „Schulinspektion“ wurde unabhängig von länderspezifischen Bezeichnungen von Schulinspektion einheitlich verwendet, um Rückschlüsse auf das jeweilige Bundesland zu vermeiden.

In den nachfolgenden Textabschnitten wird jedoch deutlich, dass diese normative Präskription nicht handlungsdeterminierend sein muss, sondern sich Veränderungen von Normen, Regeln, Konventionen, Rollen oder Überzeugungen durchaus ergeben können, wenn die Aspekte nach Scott (2008) zusammengedacht werden: Es ist möglich, dass sie interagieren und sich dadurch Akteurkonstellationen umformen.

4.2 Der regulative Aspekt

Dieser analytische Aspekt nach Scott fokussiert den Schulträger als Interpreten von formellen und informellen Regeln. Zum Beispiel gibt das Schulrecht vor, dass der Schulträger den Inspektionsbericht in seiner Rolle als Sachaufwandsträger liest und handhabt. Da der Bericht aus seiner Sicht „Regeln" für ein pädagogisches Handeln enthält, führt dies dazu, dass der Inspektionsbericht eher einen negativen Anreiz darstellt. Der Akteur Schulträger sieht sich nicht betroffen und kommt mittels Abgrenzung – wie er also seine Rolle nicht sieht – zu der Interpretation, dass er im Kontext von Schulinspektion lediglich die Position eines marginalen Akteurs einnimmt. Die im Folgenden dargestellten empirischen Beispiele machen dies deutlich.

Der Schulträger nimmt sich in seiner Rolle und Bedeutung stark zurück, indem er seine Profession und Rollenaufgabe von der der Schule abgrenzt. Er sieht sich nicht als pädagogischer Akteur. Für ihn steht im Rahmen der Schulinspektion die pädagogische Schularbeit im Vordergrund, welche ihn nur sekundär betrifft. Denn pädagogische Fragen liegen für ihn als „Sachträger" außerhalb seiner professionellen Bewertungsschemata; er hat auf sie nur eine *„begrenzte Sicht"*:

> *Ja wie gesagt, wir haben [als Schulträger] nur eine begrenzte Sicht, das heißt, also die Dinge, die pädagogisch gemacht werden, oder die Vorgaben, die da pädagogisch, umgesetzt werden müssen, entziehen sich, im Grunde genommen, meiner Beurteilung* (C_ST_1, S. 18, Z. 3-8).

Der Schulträger interpretiert die maßgeblichen Normen des Schulgesetzes auch im Rahmen von Schulinspektion so, dass er im Gegensatz zur Schule *„keinen direkten Steuerungseinfluss"* (ebd., S. 6, Z. 12f.) bis *„keinen Einfluss"* (A_ST_1, S. 13, Z. 26; C_ST_1, S. 6, Z. 17) hat. Für ihn liefert der Inspektionsbericht lediglich ‚pädagogisches' Wissen, das er nicht verwenden kann. Für den Möglichkeitsraum des Handelns im Kontext von Schulinspektion auf Seiten des Schulträgers hat dies einerseits Selbstbehauptung als Konsequenz – der Schulträger beansprucht für sich *„relevantere"* und *„drängendere"* (A_ST_1, S. 11, Z. 31) Themen – und andererseits Zurücknahme der eigenen Rolle; er ist *„nur der Schulträger"*:

> *So stellt sich das im Grunde dann in dem Inspektionsbericht dar. Und das ist dann natürlich das, was uns als Schulträger immer besonders interessiert. Wir bekommen den Entwurf dieses Inspektionsberichtes zur Kenntnis, wie gesagt das sind viele pädagogische Ansätze oder die Masse sind eigentlich pädagogische Ansätze. Die sind also für uns nicht so interessant, weil wir keinen direkten Einfluss natürlich auf bestimmte Dinge, wo vielleicht Mängel festgestellt werden, selber ja nicht haben. Also wir bedauern das und das entspricht unserem Eindruck vielleicht. Aber ändern können wir daran nichts. Wir sind nur der Schulträger. Wir haben darauf keinen Einfluss* (C_ST_1, S. 6, Z. 6-17).

Gerne wäre der Schulträger im Zusammenhang von Schulinspektion nicht bloß „Randbeteiligter" (vgl. D_ST_2). Dies ist ihm aber aufgrund normativer Begrenzungen (vgl. Kap. 4.1) per se nicht zugedacht. Das heißt: Die Schulinspektion hat den Akteur Schulträger nicht zu einem *‚composite actor'* gemacht, was bedeutet, dass der Schulträger nicht von vorneherein ein kollektiver Akteur ist bzw. eine Konstellation bildet, weil er nicht gleich konstitutiv mit anderen Akteuren zusammenarbeitet und nicht unbedingt Interdependenz erzeugt. Bezogen auf den empirischen Fall des regulativen Aspekts nach Scott bleibt der Schulträger in der Rolle des *„Außenbeobachters"* (D_ST_2, S. 9, Z. 22). Vorwiegend in der materiellen Verfügungsmacht, den Ressourcen, sieht er Steuerungsmöglichkeiten:

> *Wie wir steuern könnten zum Beispiel indem wir Schulen zusätzliche Ressourcen zur Verfügung stellen könnten* (A_ST_1, S. 21, Z. 20-22).

Empirisch kommt an dieser Stelle zum Ausdruck, dass der Schulträger gerne über seine Einflussmittel, die Ressourcen, steuern würde, dieses aber nur als „goldenen Zügel" andeutet.

In der Sichtweise des Schulträgers tun sich aber auch „Grenzüberschreitungen" auf, indem bzw. wenn er die Chance sieht, die normative – im Sinne der gesetzlich bestimmten – ‚Trennung der Ebenen' in Aufgaben und Funktionen aufzuweichen. Modi hierfür nennt der Schulträger im Folgenden:

- Generell eine gute Beziehung zur Schule herstellen, unabhängig von Hierarchie- und Ebenenzugehörigkeiten:

  > *[…] wichtig ist, dass auch die Zusammenarbeit stimmt zwischen Schulträger und Schule, auch wenn es nicht unsere Bediensteten sind, sondern Landesbedienstete, aber die Zusammenarbeit muss auch stimmen […]* (B_ST_1, S. 23, Z. 20-23).

- Über die gesetzlich definierte Rolle hinaus handeln:

  […] wir wollen ja zum guten Gelingen von Schule auch beitragen, und dazu gehört es manchmal, dass man über das hinaus geht, was gesetzlich normiert ist (C_ST_1, S. 21, Z. 23-25).

- Auf der Grundlage der Inspektionsergebnisse nach Unterstützungsbedarf fragen:

  Gibt es da Möglichkeiten wo wir als Schulträger vielleicht die Situation so verbessern können, dass die pädagogische Arbeit, die vielleicht schlecht beurteilt worden ist, etwas bessere Startchancen hat (ebd., S. 11, Z. 22-25).

Durch die Koppelung des regulativen Aspekts mit dem normativen kommt es zu einer Grenzüberschreitung. Normativ betrachtet ist der Schulträger zwar nicht für die (pädagogischen) Inhalte von Schule zuständig, sondern für deren Sachausstattung. Im regulativen Handlungsrahmen kann sich der Schulträger aber als Akteur sehen, der durch seine normative Funktion, nämlich durch eine optimale materielle Ausstattung von Schule, Beitragender für ein gutes Gelingen von Schule ist. Dadurch entsteht ein neuer ‚*frame*', ein weiterer Interpretationsrahmen für sein Handeln innerhalb des Inspektionsverfahrens. Der Schulträger leistet einen Beitrag, die pädagogische Arbeit von Schule zu verbessern, wenn er auf der Grundlage des Inspektionsberichts seine Handlungsmöglichkeiten auf qualitative Verbesserungen der Schule ausrichtet.

So positiv dieses Exempel für Grenzüberschreitung in der regionalen Governance auch klingt, es darf nicht übersehen werden, dass sich der Schulträger gezwungen sieht, in Begrenzungen zu handeln. Aus seiner Sicht bleibt die Schule auch durch Schulinspektion eine *„black box"*, die seiner Sorge um eine kontinuierliche Unterstützung der Bildungsbiografie von Kindern und Jugendlichen entzogen ist:

Die Schulen sind ja nicht in die Hierarchie hier eingebunden, sondern sind ja im Prinzip so eine Art black box. Die Schüler gehen vom Kindergarten rein und kommen dann raus und davor und danach ist der Schulträger oder die Kommune verantwortlich und während der Schulzeit, was dann drin passiert, weniger (A_ST_2, S. 6f., Z. 167-174).

Empirisch zeigt sich an dieser Stelle, dass der Schulträger eine Theorie der kommunalen Einbettung von Schule entwirft, d.h. vor und nach der Schulzeit ‚gehören' die Schüler und Schülerinnen in die Obhut der Kommune: Von dort kommen sie her und dorthin gehen sie nach der Schulzeit wieder zurück. Nur die Schulzeit erlaubt eine Auszeit von der kommunalen Adressierung, während andere Bildungsprozesse sehr wohl von der Kommune mitgestaltet werden.

Die Herausforderung für eine Überschreitung der Begrenzungen, die der Schulträger als ungewollt verpasste Chance von Steuerung wahrnimmt, bestünde nun darin, das klassische Denken in Zuständig- und Verantwortlichkeiten aufzugeben und Schule in übergreifende Kommunikationsstrukturen einzubinden.[8] Nach Scott wären dann normative und regulative Elemente verkoppelt.

4.3 Der kulturell-kognitive Aspekt

Mit diesem Aspekt nach Scott (2008) spielen, wie beim regulativen, Interpretationen eine Rolle. Hervorzuheben ist in diesem Zusammenhang – und dies zeigen die Daten –, dass das eigene Handeln im Rahmen von Schulinspektion in Prozesse der Bedeutungsgebung (z.B. Leitideen, Benefits) eingebunden wird. Hierdurch wird eine (neue) soziale Realität konstruiert, die sowohl Handlungen als auch Akteure potenziell ermächtigt. Mittels kulturell-kognitiver Deutung wird dem Instrument Schulinspektion ein hoher Stellenwert beigemessen. Denn:

> *Bildung hat eben ein hohen Stellenwert, das muss man eben sehen und deswegen haben wir also doch einiges an Investitionen* (C_ST_1, S. 2, Z. 27-29).

Damit drückt der Schulträger aus, dass es für ihn Sinn macht, in Bildung zu investieren. Auch in diesem Fall interpretiert der Schulträger seine normativ-regulative Rolle mit einer weiteren, der kulturellen.

Dadurch gewinnt der Schulträger einen Benefit. Er verlässt seine bislang *„untergeordnete"* Rolle (B_ST_1, S. 21, Z. 25) und wertet seinen Steuerungseinfluss auf. Er rückt vom Rand des Spielfeldes mehr zur Mitte des Spielfeldes vor. Auf der Ebene kulturell-kognitiver Deutungen wird gleichsam ein ideelles Bündnis mit den Schulen eingegangen, obwohl diese Anbindung in regulativer Hinsicht so nicht vorgesehen ist. Auf diese Weise eröffnen sich Chancen, zu einem Mitspieler zu werden und eigene Interessen in der Region voranzutreiben.

Ein weiterer Benefit ist, dass das Evaluationsverfahren hilft, durch mehr Wissen über die Schule seinen Einfluss zu erhöhen:

> *Einfluss oder wenigstens die Möglichkeit, davon ein Stück besser Kenntnis zu erlangen, bedeutet ja an der Stelle für uns ein besseres Verständnis für die Schulen zu entwickeln, für ihre Arbeit und natürlich auch Erkenntnisse in die Notwendigkeit bestimmter Voraussetzungen zu erlangen* (D_ST_1, S. 9, Z. 5-9).

8 So zum Beispiel in Bildungsregionen.

Schulinspektion ist geeignet, eine *„[…] langfristige Planung“* der Investitionen durchzuführen (vgl. D_ST_2, S. 2, Z. 12-16), und ist als Hilfsmittel zu sehen, *„[…] um uns nach vorne zu bringen. So sehe ich das“* (B_ST_1, S. 2, Z. 16-19). Der Schulträger deutet an dieser Stelle sein originäres Interesse an, seine Einflussmöglichkeiten und seine Position im kommunalen Standortwettbewerb zu erhöhen. Entscheidend ist, dass er grundsätzlich eine aufgewertete Steuerungsposition anstrebt und diese nicht nur im Zusammenhang von Schulinspektion. Schulinspektion kann ihm dafür ein wichtiges Instrumentarium sein.

Die mehrdimensionale Betrachtung nach Scott (2008) zeigt, dass Schulinspektion aus der Sicht des Schulträgers im Allgemeinen kein Steuerungsinstrument ist, das ihm eine bedeutende Rolle im Agieren und Entscheiden zuschreibt, dass diese aber doch Informationen produziert, die er sich im Rahmen seiner Interessen und zur Unterfütterung seiner Strategien aneignen kann. Schulinspektion provoziert zudem Konstellationen, in denen Schulen größere Bereitschaft zur Zusammenarbeit mit Schulträgern zeigen können. Schulinspektion ist *„doch ein sinnvolles Instrument“* (B_ST_1, S. 20, Z. 1f.), betont ein Schulträger, um beste Bedingungen für die pädagogische Arbeit zu erbringen. Er leistet dazu einen Beitrag neben den anderen *‚composite actors‘* wie der Schule, dem Inspektionsteam und der Schulaufsicht. Hierbei sieht sich der Schulträger vornehmlich als solitärer Akteur, der im Zusammenhang von Schulinspektion programmatisch zwar eine Rolle spielt, weil er überhaupt adressiert wurde und insofern Mitglied der *‚composite actors‘* ist, sich schlussendlich aber nur in einer Randposition verortet. Er besinnt sich auf sein professionelles Selbstverständnis, zwar kein Pädagoge zu sein (vgl. Kap. 4.1 und 4.2), seiner nur mittelbaren Rolle aber gleichwohl einen Gewinn zuzuschreiben:

> *Ich bin kein Pädagoge* (B_ST_2, S. 10). Aber: *Ist die pädagogische Arbeit gut, ist es ja auch für uns gut* (ebd., S. 6, Z. 41).

Der Schulträger formuliert sogar Wünsche und Bedürfnisse als eine Form neuen Wissens, denn die Einführung der Schulinspektion macht ihm deutlich, wie wichtig es wäre, eine *„stärkere Bildungsverantwortung der Kommune zu übernehmen“* (D_ST_1, S. 10, Z. 4) oder *„die Vernetzung der Schule mit ihrem Sozialraum zu intensivieren“* (ebd., S. 10f., Z. 34/1). Schulinspektion als flächendeckend bildungspolitisches Reformprogramm für das ganze Bundesland macht dem Schulträger an dieser Stelle eklatant bewusst, dass er nur in Experimentier- und Ausnahmefällen eine solche gewünschte Position inmitten gemeinsamer Bildungsverantwortung von Schule und Kommune einnimmt, in der die Kommune als ‚Networker‘ eine bedeutendere Steuerungsposition einnimmt als bislang.

Ein Schulträger pointiert die vorstehend beschriebenen Zusammenhänge von Vorteilen durch veränderte Perspektiven auf den Tatbestand:

> *Wenn man natürlich mit so einem Instrument Schnittstellen definieren könnte zwischen Kommune und Schule, wenn man Synergien herausfinden könnte wie einzelne Bereiche miteinander besser zusammenarbeiten, im Sinne der Verbesserung von Bildungsqualität für die Kommune, wäre das natürlich eine gute Sache* (D_ST_1, S. 11, Z. 6-10).

Hiermit bestätigt sich noch einmal, dass es dem Schulträger hinsichtlich der Schnittstelle und Zusammenarbeit mit Schule auch um kommunale Interessen geht. Er möchte die Bildungsqualität insgesamt aufgewertet wissen, und in dem Punkt sieht er Schulinspektion als chancenreiches und gewinnbringendes Instrumentarium.

5. Fazit

Resümierend kann erstens konstatiert werden, dass sich die drei Institutionenaspekte nach Scott (2008) in methodischer Hinsicht als hilfreich für die Governance-Forschung erweisen. Sie schlüsseln die thematischen Kodierungen (nach Flick 2007) feindifferenziert auf und verleihen der Analyse vor allem im Hinblick auf die Perzeptions- und Rollenebene des Akteurs Schulträger sowohl Deutungs- als auch Tiefendimensionen. Interessant wäre es, diese Analysemittel auch für andere Akteurkonstellationen, wie zum Beispiel zwischen Schulleitung und Schule oder zwischen Schulaufsicht und Schulleitung, zu verwenden. Damit ließen sich tiefendifferenzierte Aufklärungen in Evaluationsdaten bringen. Der Bedarf hierfür ist groß.

Zweitens zeigen die Aspekte nach Scott in analytischer Hinsicht, dass sich durch Schulinspektion die Akteurkonstellation zwischen Schulträger und Schule grundsätzlich nicht verändert. Von Bedeutung sind jedoch Grenzüberschreitungen im Rollenhandeln des Schulträgers, wenn dieser sich jenseits von klassischen Zuständigkeits- und Ressortfragen Handlungsmöglichkeiten erschließt. Als wesentlicher Befund zeichnet sich dabei ab, dass der Schulträger seine Einflussmöglichkeiten und -chancen als Steuerungsakteur im Zusammenhang von Schulinspektion gewahrt sehen will, obwohl oder gerade weil er eigentlich im ‚Spiel' Schulqualität ‚draußen' steht.

Damit ist gemeint, dass der Schulträger im Kontext von Schulinspektion gerne ein echter ‚Mitspieler' wäre und für ihn die Frage nach seinem Steuerungseinfluss eine Grundsatzfrage ist (vgl. A_ST_1, S. 6, Z. 157f.). Faktisch sieht er sich jedoch nicht angeschlossen und in einer Rolle als Außenbeobachter am Rande des Geschehens. Die Analyse nach Scott (2008) schlüsselt dabei differenziert auf, dass es empirische Fälle gibt, bei denen die Schulträger es dennoch schaffen, ihren Steuerungseinfluss – zumindest und zunächst – subjektiv zu erhöhen: Verkoppeln sich etwa normative und kognitive Dimensionen des Handelns, wird es möglich, dass der regionale Akteur kognitiv ‚über seine Rolle hinaus' agiert. In seiner Interpretation und folglich

im Handeln tun sich veränderte Deutungs-, Legitimations- und Herrschaftsmuster[9] auf, wenn sich der Schulträger zum Beispiel nicht mehr allein auf seine gesetzlich determinierte Rolle beruft. Unter diesem veränderten bzw. erweiterten ‚*frame*' sieht sich der Schulträger nicht länger nur als ‚materiellen Belieferer' von Schule, sondern in einem ideellen Bündnis mit Schule, in dem sich beide um Optimierung der Bedingungen für gute Pädagogik bemühen. In diesem Fall legt der Schulträger die normative Begrenzung weit aus, ohne aber an deren Grenzen zu rütteln. Bildung hat aus dieser Perspektive einen so hohen Stellenwert, dass es sich für den Schulträger lohnt, zu investieren und Fragen jenseits von Zuständigkeiten zu beantworten. Schulinspektion bedeutet dann für den Schulträger ein nützliches Instrument, das auch hilft, eigene Positionen und Interessen in der Region zu stärken. Dies ist prinzipiell gesetzeskonform, da Gesetze sowohl auslegungsbedürftig als auch -fähig sein sollen. Auf diese Weise ergeben sich nicht nur neue Schnittstellen zwischen Schule und Region, sondern auch ein Wandel der bislang „*unbefriedigenden Rolle*" (A-ST, S. 18, Z. 542) des Schulträgers.

Allerdings wird von Seiten der Schule eine solche Umformung der Akteurkonstellation nicht unbedingt gesehen, was sich in der Aussage einer Schulleitung widerspiegelt:

> *Das heißt, der Schulträger hat natürlich den [Inspektionsbericht] bekommen, aber den Schulträger betrifft er eigentlich nicht. Der wurde auch nirgends genannt, dass der Schulträger, äh, mehr für das oder jenes sorgen müsste. Von daher hat der Schulträger jetzt im Ergebnis bei uns, das muss nicht überall so sein, aber bei unserer Schule hat jetzt der Schulträger eigentlich keine Rolle gespielt* (A_SL_3, S. 19, Z. 22).

In jedem Fall müsste ein Bündnis zwischen Schulträger und Schule auf einem gemeinsamen Aushandlungsprozess basieren, der Fragen jenseits von üblichen Zuständigkeitsfragen thematisiert. Denkbar ist, dass dieser Impuls auch von staatlichen Bestimmungen ausgeht, was einer heterarchischen Selbstregulation gleichkäme, die temporäre Hierarchien ermöglicht, wodurch sich Systeme selbst regeln und sich dabei an die gemäß ihrer eigenen Differenzierung erzeugten Wirklichkeitskonstrukte anpassen (vgl. Hejl 1992). Bislang zeigen die Daten des BMBF-Projekts, dass in allen empirischen Fällen die beteiligten Akteure mittels Schulinspektion ein Interdependenzbewusstsein entwickeln (vgl. Preuß/Wissinger/Brüsemeister, im Erscheinen), sie sich also ihre Rollen und Funktionen im Zusammenhandeln mit den anderen vergegenwärtigen und zu klären versuchen.

9 Vgl. zu veränderten Deutungs-, Legitimations- und Herrschaftsmustern am Beispiel von „kommunalen Bildungsprogrammen und die Bearbeitung von Akteurkonstellationen" den Vortrag von Sebastian Niedlich, Julia Klausing, Felix Brümmer und Thomas Brüsemeister auf der KBBB-Herbsttagung, 01.10.2012, Justus-Liebig-Universität Gießen.

Zur Vervollständigung der empirischen Perspektive müsste an diesem Punkt ferner das Verhältnis des Schulträgers im Governance-Dreieck zwischen Schule und Schulaufsicht geprüft werden, das sich in einem Interview mit einer Vertreterin/einem Vertreter der Schulaufsicht wie folgt andeutet:

> *[…] der Schulträger sagt „Wir hätten gerne an der Stelle eine Ganztagesschule, weil wir die Infrastruktur schon haben". Jetzt sagt die Schule „Das möchten wir aber nicht". Jetzt stehen wir dazwischen. Da sagt der Schulträger „Haben wir das Sagen oder haben Sie das Sagen? Wie läuft das?"* (A_SAF_1, S. 17, Z. 30 - S. 18, Z. 4).

Mitentscheidend dafür, den Schulträger im Kontext von Schulinspektion in seinem „Steuerungseinfluss" aufzuwerten und die Umgewichtungen in der Akteurkonstellation nicht überwiegend ungesteuert zu lassen, wird zudem sein, dass ihm seitens der Bildungspolitik erweiterte Handlungsspielräume zugestanden werden. Dies wäre nicht nur für die Effektivität von evaluationsbasierter Schulentwicklung gut, sondern auch für andere Reformmaßnahmen, wie z.B. die der Eigenverantwortlichen Schule, für die Mitverantwortung und Partnerschaft wesentliche Parameter gelungenen Interdependenzmanagements sind (vgl. Winter 2007). Vielleicht funktioniert dies gar nicht so kompliziert, wie gewöhnlich Bildungspolitik mit komplexen Reformbemühungen ‚denkt', sondern schon durch die Bereitschaft der Länder loszulassen, vorausgesetzt die Kommunen sind bereit „anzupacken" (vgl. Hennecke 2011, S. 226).

Literatur

Altrichter, H./Brüsemeister, T./Wissinger, J. (Hrsg.) (2007): Educational Governance. Handlungskoordination und Steuerung im Bildungssystem. Wiesbaden: VS.

Altrichter, H./Heinrich, M. (2007): Kategorien der Governance-Analyse und Transformation der Systemsteuerung in Österreich. In: Altrichter, H./Brüsemeister, T./Wissinger, J. (Hrsg.): Educational Governance. Handlungskoordination und Steuerung im Bildungssystem. Wiesbaden: VS, S. 55-103.

Altrichter H./Maag Merki, K. (Hrsg.) (2010): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS.

Benz, A./Lütz, S./Schimank, U./Simonis, G. (2007): Handbuch Governance. Wiesbaden: VS.

Bertelsmann Stiftung (2012): Bildungsregionen. URL: http://www.bertelsmann-stiftung.de/cps/rde/xchg/bst/hs.xsl/10117.htm; Zugriffsdatum: 12.12.2012.

Böttcher, W./Kotthoff, H.-G. (2007): Schulinspektion zwischen Rechenschaftslegung und schulischer Qualitätsentwicklung: Internationale Erfahrungen. In: Böttcher, W./Kotthoff, H.-G. (Hrsg.): Schulinspektion: Evaluation, Rechenschaftslegung und Qualitätsentwicklung. Münster u.a.: Waxmann, S. 9-20.

Bos, W./Dedering, K./Holtappels, H.G./Müller, S./Rösner, E. (2007): Schulinspektion in Deutschland – Eine kritische Bestandsaufnahme. In: Van Buer, J./Wagner, C. (Hrsg.): Qualität von Schule – Entwicklungen zwischen erweiterter Selbstständigkeit, defi-

nierten Bildungsstandards und strikter Ergebniskontrolle. Ein kritisches Handbuch. Frankfurt a.M.: Lang, S. 241-257.
Bräth, P./Eickmann, M./Galas, D. (2009): Niedersächsisches Schulgesetz. Kommentar. Überarb. Aufl. Köln: LinkLuchterhand.
Brüsemeister, T./Preuß, B./Wissinger, J. (2012): Schulinspektion im regionalen und institutionellen Kontext – Qualitative Befunde zu schulinspektionsbezogenen Akteurkonstellationen. Vortrag zum Workshop „Schulinspektion im Mehrebenensystem: Länderübergreifende Projektergebnisse" am 07.12.2012, Leibniz Universität Hannover.
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".
Bruggen, J.C. v. (2006): Schulinspektion in den Niederlanden. Metaevaluation und punktuelle Tiefe auf Wunsch der Schule. In: Buchen, H./Horster, L./Rolff, H.-G. (Hrsg.): Schulinspektion und Schulleitung. Berlin: Raabe, S. 107-124.
Bußhoff, H. (Hrsg.) (1992): Politische Steuerung: Steuerbarkeit und Steuerungsfähigkeit. Beiträge zur Grundlagendiskussion. Baden-Baden: Nomos.
Dedering, K. (2012): Schulinspektion als wirksamer Weg der Systemsteuerung? In: Zeitschrift für Pädagogik 58, H. 1, S. 70-88.
Emmerich, M. (2010): Regionalisierung und Schulentwicklung: Bildungsregionen als Modernisierungsansätze im Bildungssektor. In: Altrichter H./Maag Merki, K. (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS, S. 355-375.
Flick, U. (2007): Qualitative Sozialforschung. Eine Einführung. Reinbek: Rowohlt.
Hejl, P. (1992): Politik, Pluralismus und gesellschaftliche Selbstregelung. In: Bußhoff, H. (Hrsg.): Politische Steuerung: Steuerbarkeit und Steuerungsfähigkeit. Beiträge zur Grundlagendiskussion. Baden-Baden: Nomos, S. 107-142.
Hennecke, H.-G. (2011): Kommunale Bildungskompetenzen. Bericht über das DLT-Professorengespräch 2011 im Rhein-Sieg-Kreis. In: Der Landkreis, H. 6, S. 223-232.
Husfeldt, V. (2011): Externe Schulevaluation. Ein länderübergreifender Blick auf Forschungen und Modelle. In: Quesel, C./Husfeldt, V./Landwehr, N./Steiner, P. (Hrsg.): Wirkungen und Wirksamkeit der externen Schulevaluation. Bern: hep, S. 13-33.
Kotthoff, H.-G./Böttcher, W. (2010): Neue Formen der „Schulinspektion": Wirkungshoffnungen und Wirksamkeit im Spiegel empirischer Bildungsforschung. In: Altrichter, H./Maag Merki, K. (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS, S. 295-325.
Kussau, J./Brüsemeister, T. (2007): Educational Governance: Zur Analyse der Handlungskoordination im Mehrebenensystem der Schule. In: Altrichter, H./Brüsemeister, T./Wissinger, J. (Hrsg.): Educational Governance. Handlungskoordination und Steuerung im Bildungssystem. Wiesbaden: VS, S. 15-54.
Müller, S. (2010): Erste Effekte von Schulinspektion – Eine Zwischenbilanz. In: Berkemeyer, N./Bos, W./Holtappels, H.G./McElvany, N./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung, Bd. 16. Daten, Beispiele, Perspektiven. Weinheim/München: Juventa, S. 289-308.
Müller, S./Pietsch, M./Bos, W. (Hrsg.) (2011): Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht. Münster u.a.: Waxmann.
Preuß, B./Wissinger, J./Brüsemeister, T. (im Erscheinen): Einführung der Schulinspektion: Struktur und Wandel regionaler Governance im Schulsystem. In: Abs, H.J./Brüsemeister, T./Schemmann, M./Wissinger, J. (Hrsg.): Mixed Modes of Governance in Education. Band zur Herbsttagung der KBBB 2012. Wiesbaden: Springer VS.

Rürup, M./Lambrecht, M. (2012): Deregulierung durch Schulinspektion? Zur Berechtigung einer Fragestellung. In: Hornberg, S./Parreira do Amaral, M. (Hrsg.): Deregulierung im Bildungswesen. Münster u.a.: Waxmann, S. 165-186.

Schnell, H. (2007): Perspektiven der Bildungsverwaltung in Deutschland. Verfassungsauftrag und Aufsicht über Schulen. In: SchulVerwaltung, S. 261-262.

Scott, W.R. (^{3}2008): Institutions and Organizations. Ideas and Interests. Los Angeles, CA, u.a.: Sage Publications.

Sekretariat der Kultusministerkonferenz (2012). URL: http://www.bildungsserver.de/glossarset.html?sp=0&Id=144; Zugriffsdatum: 12.12.2012.

Van Buer, J./Wagner, C. (Hrsg.) (2007): Qualität von Schule – Entwicklungen zwischen erweiterter Selbstständigkeit, definierten Bildungsstandards und strikter Ergebniskontrolle. Ein kritisches Handbuch. Frankfurt a.M.: Lang.

Wacker, A./Maier, U./Wissinger, J. (2011): Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung. Empirische Befunde und forschungsmethodische Implikationen. Wiesbaden: Springer VS.

Winter, R. (2007): Einbinden und Rat holen: Schulträger. In: Busemann, B./Oelkers, J./Rosenbusch, H.S. (Hrsg.): Eigenverantwortliche Schule – ein Leitfaden. Konzepte, Wege, Akteure. Köln: LinkLuchterhand.

Bianca Preuß, Dr., geb. 1965, wissenschaftliche Mitarbeiterin an der Justus-Liebig-Universität Gießen.
E-Mail: Bianca.E.Preuss@sowi.uni-giessen.de

Anschrift: Karl-Glöckner-Straße 21 E, 35394 Gießen

ERWEITERTE SCHULISCHE VERANTWORTUNG UND INTERNE EVALUATION

DDS – Die Deutsche Schule
Beiheft 12, S. 172-196
© 2013 Waxmann

Stefan Brauckmann/Christoph Herrmann

Belastungserleben von Schulleiterinnen und Schulleitern im Rahmen erweiterter schulischer Eigenständigkeit

Erste empirische Befunde aus der SHaRP-Studie

Zusammenfassung
Mit der Veränderung schulrechtlicher Zuständigkeiten und Pflichten sollen neue intra- wie interinstitutionelle Verantwortungsgefüge (shift of powers) *entstehen, die in letzter Konsequenz auch eine Reorganisation des bisherigen Schulleitungshandelns bedeuten* (shift of tasks). *Es ist jedoch derzeit noch offen, ob, wie und in welchem Umfang der* shift of powers *und der* shift of tasks*, also die (aktive) Wahrnehmung von mehr Verantwortung in zunehmend komplexeren Tätigkeitsfeldern, in der sozialen Realität der einzelnen Schule gelingen können. Vor diesem Hintergrund sollen im vorliegenden Beitrag Hürden schulleiterischen Handelns im Rahmen erweiterter schulischer Eigenständigkeit anhand tätigkeitsbezogener zeitlicher Belastung und dem daraus resultierenden Belastungsempfinden von Schulleiterinnen und Schulleitern untersucht werden. Hierzu erfolgt zunächst eine forschungsorientierte wie praxisgeleitete Bestimmung der im Rahmen des „Neuen" Steuerungskontextes relevanten Tätigkeitsbereiche von Schulleitern und Schulleiterinnen, um sie anschließend als Gegenstand der Datenerhebung empirisch zu konkretisieren. Aufgrund des explorativen Charakters der Studie wird mit Hilfe von Korrelationen auf Basis von theoriegeleiteten Annahmen der Zusammenhang zwischen ausgeübten Tätigkeiten und Belastungsfaktoren untersucht.*
Schlüsselwörter: Autonomie, Schulleitungshandeln, Belastung, Kontext

Workload of School Principals in the Context of Greater School Autonomy

First Empirical Findings from the SHaRP-Study

Summary

Changes of school legislation responsibilities and obligations are aimed at the development of new intra- and inter-institutional structures of responsibility (shift of powers), which ultimately mean a re-organization of previously existing school leadership actions (shift of tasks). So far it is yet unknown if, how and to what degree the shift of powers and shift of tasks, that means actively taking on more responsibility in increasingly complex fields of activity, can succeed in the social reality of individual schools. Against this background, the contribution aims to investigate obstacles to school leadership actions in the framework of enhanced school autonomy, concerning activity-related demands on time and the resulting experience of workload by school leaders. To this end, a research-oriented and practically driven determination of school leadership activities is presented that is relevant in the context of "new governance". Subsequently these fields are empirically defined as objects of data assessment. Owing to the explorative character of the study, correlations based on theory-directed assumptions serve to investigate the relationship between the tasks that are carried out and experience of workload factors.
Keywords: autonomy, school leadership action, workload, context

1. Ausgangspunkt der SHaRP-Studie

Vor dem Hintergrund der Ergebnisse der internationalen Leistungsvergleichsstudien, die sich methodisch dem Schuleffektivitäts- und Schulentwicklungsparadigma verschrieben haben, haben die deutschen Bundesländer mit der Umstellung der Steuerung des Schulwesens von der Input- auf die evaluationsbasierte Outputsteuerung den jüngsten Modernisierungsschub im Schulsystem in Gang gesetzt. In den vergangenen 20 Jahren haben die deutschen Bundesländer in unterschiedlichem Maße die Eigenverantwortung der Schule erweitert und in diesem Zusammenhang auch die formale Stellung der Schulleitung gestärkt. Aus Sicht der Schuleffektivitäts- und Schulentwicklungsforschung hat der Gesetzgeber schulrechtlich Rechnung getragen, dass der Schulleitung bei der Qualitätssicherung und -entwicklung eine Schlüsselrolle zukommt (vgl. Wissinger 2011). So attestiert eine zunehmende Zahl empirischer Studien dem Schulleitungshandeln einen messbaren, meist direkten Einfluss auf die Lehr- und Lernqualität (vgl. Louis u.a. 2010; Day u.a. 2009; Robinson/Lloyd/Rowe 2008; Marzano/Waters/McNulty 2005; Witziers/Bosker/Kruger 2003; Bonsen u.a. 2002).

Die Umsetzung der auf Qualitätssicherung gerichteten neuen Entscheidungsmöglichkeiten und Gestaltungsoptionen schließt die zentrale Rolle von Schulleitung ein,

die sich mit erweiterten rechtlichen Kompetenzen, Verantwortlichkeiten und neuen Aufgaben konfrontiert sieht. Schulische Qualitätsrahmen und vermehrte Rechenschaftspflicht führen zu normativen Handlungsvorgaben für Schulleitungsmitglieder. Angesichts der Schulaufsichtspflicht des Staates erzeugt die damit vorgenommene Erweiterung der schulischen Eigenverantwortung somit eine Spannung im Verhältnis zur einzelnen Schule, die insbesondere die Schulleitung betrifft. Die Schulleiter und Schulleiterinnen sind an der Schnittstelle zwischen außerschulischer und schulischer Umwelt positioniert und tragen die Gesamtverantwortung für die Schule. Die sog. „Neue Steuerung", die dem Prinzip „Eigenverantwortung und Rechenschaftslegung" folgt, verändert die Rahmenbedingungen schulischen Handelns, insbesondere von Schulleitungen, in einer Weise, dass sich grundsätzlicher als bislang Fragen nach Voraussetzungen, Bedingungen und Wirksamkeit ihres Handelns stellen. Verantwortung wird auf die einzelschulische Ebene verlagert, Aufgaben und neue Ansprüche werden dorthin übergeben, Lösungen von dort erwartet.

Die neuen Steuerungsansätze, die mehr Autonomie und Eigenverantwortung fordern, bleiben nicht ohne Folgen für die Managementaufgaben der Schulleitungsmitglieder. An Schulen mit mehr Eigenverantwortung stehen Schulleitungen vor neuen Aufgaben, wie etwa strategischen Führungsaufgaben, da ihnen nicht nur das operative Management im Sinne des administrativen und pädagogischen Alltagsgeschäfts zufällt, sondern sie auch die Ziele der jeweiligen Schulentwicklung festzulegen und deren konzeptionelle Umsetzung mit Blick auf die extern herangetragenen Standards zu sichern haben. Somit steht die Einzelschule vor der Aufgabe, ihre Entwicklung eigenständig, zielgerichtet und systematisch zu betreiben. Hierfür benötigt sie eine entsprechende Selbststeuerungsfähigkeit. Den Schulleitungen obliegt es, diese aufzubauen und zu fördern.

Dass Schulleiter und Schulleiterinnen damit überfordert sind, sich allein gelassen und ohne Unterstützung empfinden, wird dabei oft übersehen. Unterstrichen wird dies durch Befunde internationaler Studien, die feststellen, dass die Arbeitsbelastung im subjektiven Empfinden der Schulleitungen aufgrund ihres erweiterten Tätigkeitsspektrums zugenommen hat (vgl. Leithwood u.a. 1999; Dworking 2009). Vor diesem Hintergrund können eine empfundene zeitliche Mehrbelastung und das erweiterte Aufgabenspektrum zu Widerständen beim zentralen *Change Agent* Schulleitung führen, da sie als Fehlen individueller Handlungsressourcen wahrgenommen werden und somit die Erprobung und Erlernung neuer Rollen und Handlungen nachhaltig behindern beziehungsweise negativ beeinflussen (vgl. Winkler 2009) und darüber hinaus die Rekrutierung neuer Schulleitungsmitglieder und die Besetzung vakanter Schulleiterpositionen nachhaltig erschweren (vgl. OECD 2008). Was die Verteilung der Arbeitszeit auf zentrale Schulleitungstätigkeitsbereiche angeht, so betonen Autoren und Autorinnen vielfach vor allem die Belastung durch Verwaltungs- und organisatorische Aktivitäten zu Lasten konzeptioneller Tätigkeiten (vgl. Warwas

2013; Landert/Partner 2009; Rosenbusch/Braun-Bau/Warwas 2007; Hasenbank 2000; Bonsen/Iglhaut/Pfeiffer 1999; Rosenbusch/Wissinger 1989).

Deutsche Schulleiterverbände betonen in diesem Zusammenhang, dass es unter den gegenwärtigen Rahmenbedingungen zusehends schwieriger wird, eine ausreichende Zahl von bereitem Personal für die vakanten Schulleiterpositionen zu gewinnen. Erschwerend kommt ihrer Meinung nach hinzu, dass unter den keinesfalls zahlreichen Bewerbern und Bewerberinnen nur wenige sind, die über die notwendigen Einstellungsvoraussetzungen verfügen. Mit beiden Seiten – den gestiegenen Erwartungen und Ansprüchen an Bildungseinrichtungen einerseits und den Anforderungen an das dafür benötigte Personal sowie dem Belastungserleben der derzeit tätigen Professionellen und deren Ansprüchen an ihre Tätigkeiten andererseits – wird auch gegenwärtig der Professionalisierungsdiskurs im Bereich des Schulleitungshandelns in Deutschland bestimmt und legitimiert (vgl. Füssel 2009).

Von wissenschaftlicher Seite kommt erschwerend hinzu, dass nur höchst vereinzelt Forschungsergebnisse zum Schulleitungshandeln in erweiterter Selbständigkeit vorliegen, die zudem keine eindeutige Sprache aufweisen. Dies liegt daran, dass der Reformprozess der eigenverantwortlichen Schule in allen befragten Bundesländern noch in Bewegung ist. Zu beachten ist zudem, dass sich die Transferstrategien und die gesetzliche Verankerung der schulischen Eigenverantwortung in den Ländern nicht entsprechen, auch wenn sich die für Schule zuständigen Ministerinnen und Minister in der Bundesrepublik Deutschland auf eine Gesamtstrategie verständigt haben und dies in gemeinsamen Beschlüssen ausformulieren (vgl. dazu Altrichter/Maag Merki 2010; Rürup 2007).

Nur in einigen Studien werden die Funktionsausübungen von Schulleitungen und die Rahmenbedingungen, denen sie unterliegen, untersucht (vgl. Horng/Klasik/Loeb 2009; May/Goldring/Hiff 2008; Rosenbusch/Braun-Bau/Warwas 2007; Goldring u.a. 2006). Nur wenige Befunde liegen zur Ausschöpfung dieser neuen Entscheidungsmöglichkeiten und Gestaltungsoptionen vor (vgl. Senkbeil u.a. 2004). Ferner blieben die Zusammenhänge zwischen kontext-/situationsbezogenen Bedingungen und Formen des Schulleitungshandelns bislang weitgehend unbestimmt (vgl. OECD 2009).

Varianzen in der Ein- bzw. Verteilung der Arbeit auf strategische und operative Tätigkeitsfelder werden teilweise auf systemische (z.B. Autonomiegrad), auf institutionelle (z.B. Schulform, Schulgröße) (vgl. Goldring u.a. 2008) oder aber auf situative Einflussfaktoren zurückgeführt. Somit ist auf der bisherigen schmalen Befundlage nicht erkennbar, ob und inwieweit der Kontext auf System-, einzelschulischer und personenbezogener Ebene das Schulleitungshandeln normiert oder diversifiziert. Angesichts der in der Regel geringen Stichprobengröße bisheriger Untersuchungen ist deren explorativer Charakter zu betonen. Gerade vor diesem Hintergrund empfehlen

sich für den deutschsprachigen Forschungsraum weitergehende empirische Ansätze, die einerseits repräsentative Aussagen in einem länderübergreifenden Vergleich gestatten und andererseits auch der bisherigen Befundlage zu institutions- und situationsspezifischen/individuellen Akzentsetzungen in der Aufgaben- und Zeitstruktur von Schulleitungen Rechnung tragen (vgl. Brauckmann/Herrmann 2012).

Die diesbezüglichen Implikationen können – gerade hinsichtlich innerschulischer und externer Unterstützungssysteme – neue oder modifizierte Schwerpunktsetzungen für verschiedene Akteure im bildungspolitischen, verwaltungsbezogenen, (schul-) aufsichtlichen und -praktischen Arbeitskontext nahelegen. Benötigt werden daher Studiendesgins, die empirisch fundierte Erkenntnisse für die gegenwärtig stattfindende Diskussion über die Gestaltungsspielräume und Belastungsfaktoren von Schulleitungen im Rahmen erweiterter Eigenverantwortung liefern. Solche Designs können ferner Hinweise auf die mit der schulischen Eigenverantwortung verbundenen Herausforderungen und Problemlagen von Schulleitern und Schulleiterinnen geben.

2. Zielsetzung und konzeptioneller Bezugspunkt der SHaRP-Studie

Das vom BMBF geförderte Forschungsprojekt „Schulleitungshandeln zwischen erweiterten Rechten und Pflichten“ (SHaRP) will zu einer weiteren empirisch fundierten Klärung des Belastungserlebens von Schulleitern und Schulleiterinnen im Kontext neuer Steuerungsansätze beitragen.[1] Zu diesem Zweck wurden in Grundschulen und Gymnasien von sechs Bundesländern, die laut eines Schulautonomieindizes (vgl. Rürup 2007) einen unterschiedlichen Grad an Eigenverantwortlichkeit der Einzelschule ausweisen, quantitative und qualitative Daten (a) zum Schulleitungshandeln selbst sowie (b) zum Belastungsempfinden von Schulleitern und Schulleiterinnen erhoben.

Die Befunde sollen es erlauben, Belastungsfaktoren und Tätigkeiten von Schulleitungen sowie regionale, institutionelle und auch persönlichkeitsbezogene Unterschiede zu identifizieren. Weiterhin soll der Zusammenhang von Belastungsfaktoren und Schulleitungshandeln aufgezeigt werden.

Das allgemein gehaltene und für die Akteursgruppen Lehrkräfte und Schüler und Schülerinnen validierte theoretische Modell von Böhm-Kasper (2004) bietet einen Ausgangspunkt, um die komplexen Belastungs- und Beanspruchungs-Zusammenhänge bei Schulleitungen modellhaft darzustellen und zu untersuchen. Für

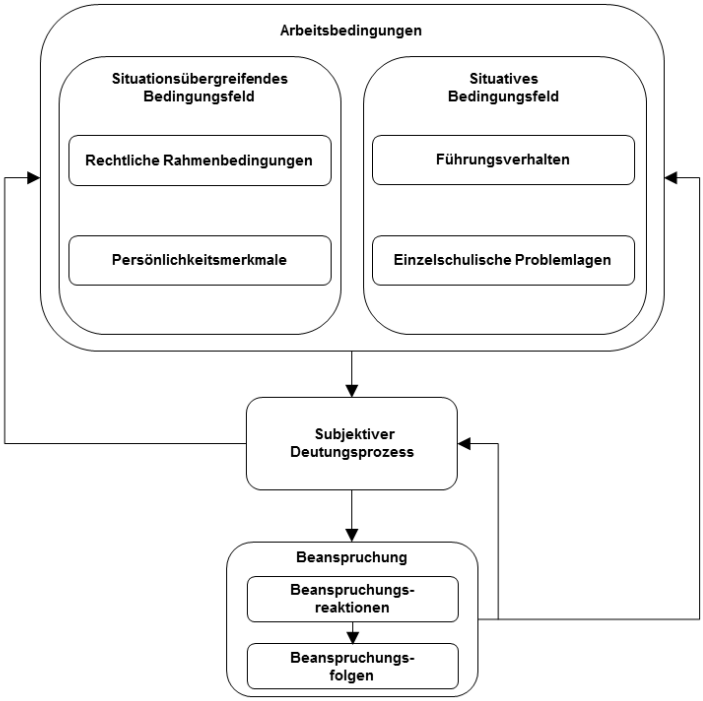
1 Unser Dank gilt dem Bundesministerium für Bildung und Forschung für die Förderung des Einzelvorhabens „Schulleitungshandeln zwischen erweiterten Rechten und Pflichten“ (SHaRP) im Rahmen der Förderung der empirischen Bildungsforschung im Forschungsschwerpunkt „Steuerung im Bildungssystem“ (SteBis).

den Bereich der Belastungsfaktoren betont er die Notwendigkeit einer Differenzierung zwischen objektiven Tätigkeitsanforderungen, individuellen Handlungsvoraussetzungen und subjektiven und sachlichen Handlungsbedingungen. Zusammen bilden diese Kategorien die quantifizierbare Belastung ab, die vermittelt über subjektive Deutungsprozesse zu kurzzeitig (und zeitnah) auftretenden Beanspruchungsreaktionen und langfristig (überdauernd) auftretenden Beanspruchungsfolgen führt (vgl. ebd.).

Im Sinne handlungstheoretischer Rahmenmodelle zur Belastung und Beanspruchung (vgl. ebd.; Rudow 1994) entsteht subjektive Belastung bzw. Beanspruchung aus der Bewertung und Interpretation von objektiven Anforderungen durch das Individuum mit all seinen Handlungsvoraussetzungen. „*Objektive Belastungen*" sind einer transparenten Messung bzw. Erfassung nur schwer zugänglich. Näherungsweise sind sie durch die Normierung „neuer" rechtlicher Anforderungsprofile bestimmbar, die auf die erhöhte Eigenverantwortlichkeit und Rechenschaftspflicht der Schulleitungen abzielen. Der der subjektiven Belastung vorausgehende interpretative Bewertungsprozess kann nicht direkt erfasst werden, sondern nur das Ergebnis desselben.

Für die Entstehung von subjektiv messbarer Belastung kann daher die Bewertung von Eigenschaften oder Merkmalen aus dem beruflichen Kontext herangezogen werden, um so auf latente Belastungsfaktoren zu schließen. Belastungsfaktoren sind danach zu differenzieren, ob sie situationsübergreifend oder situativ sind, d.h. ob es sich um stabile Merkmale bzw. *Kontextfaktoren* (wie z.B. die Schulform) oder um situative Faktoren (z.B. den konkreten Ausfall von Personal an einem Schultag) handelt, die für die Interpretation von entstandener Beanspruchung herangezogen werden können. Während der Umgang mit situativen Belastungsfaktoren einstellungsbezogen ist, da u.a. die Einschätzung von Sachaspekten (wie z.B. das Vorhandensein von materieller Ausstattung) oder Beziehungsaspekten (z.B. die Einschätzung der Kooperation im Kollegium) vor dem Hintergrund der eigenen Erfahrung reflektiert wird, sind situationsübergreifende Belastungsfaktoren, wie z.B. die Schulform, als bestehende Größen zu betrachten (vgl. Abb. 1).

Abb. 1: Modell der Belastung und Beanspruchung



Quelle: nach Böhm-Kasper (2004), modifiziert für den Schulleitungsberuf

Vor dem Hintergrund des verwendeten Belastungs- und Beanspruchungsmodells stellt in SHaRP die aufgewendete Zeit in einer Tätigkeit bzw. in einem Tätigkeitsfeld die „objektive Belastung" dar, während der mit der Ausübung der Tätigkeit einhergehende Beanspruchungsgrad das Ergebnis des subjektiven Deutungsprozesseses, mit anderen Worten: die „subjektive Belastung", ist. Letztere stellt in der Analyse des vorliegenden Aufsatzes die abhängige Variable dar. Das „subjektive Belastungserleben" wird sowohl tätigkeitsbezogen als auch tätigkeitsübergreifend im Sinne beruflicher Belastung nach Enzmann und Kleiber (1989) erfasst.

3. Anlage und Durchführung der SHaRP-Studie

Die SHaRP-Studie gliedert sich insgesamt in vier Forschungsphasen:

1) eine dokumentengestützte Ausgangsanalyse: Beschreibung der Tätigkeitsfelder von Schulleitung aus rechtlicher, theoretischer und professionstheoretischer Perspektive (Ermittlung des Soll-Zustandes);

2) eine quantitative Feldphase: Erfassung von Arbeitsstrukturen und -beanspruchung von Schulleitungen in den zuvor bestimmten Tätigkeitsbereichen mittels einer elektronischen standardisierten Befragung und Selbsteinschätzung der Aufgabenstrukturen (Ermittlung des Ist-Zustandes);
3) eine qualitative Feldphase: vertiefende, weitergehende Interviews mit ausgewählten Schulleitern und Schulleiterinnen, deren Tätigkeits- und Belastungsbefunde entweder im Widerspruch zu theoretischen Annahmen und Studien stehen oder aber über die bisher kein ausreichendes theoretisches oder aber empirisches Wissen vorliegt;
4) eine Verschränkung bzw. Kontrastierung der wesentlichen Befunde des qualitativen und quantitativen Forschungszugangs.

3.1 Dokumentengestützte Ausgangsanalyse

Um die Rahmenbedingungen, denen Schulleitungshandeln unterliegt, wissenschaftlich seriös erfassen zu können, sollten die Anforderungsprofile an Schulleitungshandeln über die theorie- und professionsbasierten Perspektiven hinaus vor allem auch aus rechtlicher Perspektive beschrieben werden. Hierzu wurde ein schulrechtliches Gutachten in Auftrag gegeben (vgl. Hanßen 2012). Die in Auftrag gegebene schulrechtliche Expertise, die in der Zwischenzeit auf sechs Bundesländer ausgeweitet wurde, liegt inzwischen für alle sechs Bundesländer vor. Diese Analyse beinhaltet nicht nur die rechtswissenschaftliche Darstellung von Kompetenzen und Weisungsbefugnissen von Schulleitung, sondern darauf aufbauend auch die Beschreibung unterschiedlicher Beteiligungs-, Mitwirkungs- sowie Mitbestimmungsformen von Führungskräften in schulinternen wie schulexternen Prozessen der Entscheidungsfindung. Diese Perspektive erbrachte somit neue Informationen über inhaltliche und insbesondere bisher vernachlässigte prozessuale Zusammenhänge eines auf qualitätsorientierte Selbststeuerung bezogenen Schulleitungshandelns. In diesem Zusammenhang wurden auf der Grundlage einer qualitativen Inhaltsanalyse Tätigkeitskategorien erstellt.

3.2 Quantitative Feldphase

Im Rahmen der Forschungsphase II mussten die in Phase I schulrechtlich hergeleiteten Tätigkeitskategorien operationalisiert werden, d.h. sie mussten mit einzelnen Tätigkeiten unterlegt werden, um die Tätigkeitskategorien empirisch überprüfbar zu machen. Gleichzeitig musste die inhaltliche Trennschärfe der einzelnen Tätigkeitskategorien untereinander gewährleistet bleiben. Hierzu wurden Sekundäranalysen mit Daten[2] aus dem Gutachten zur Orientierungsstufe in

2 Für die Bereitstellung dieser Daten möchten wir Herrn Prof. Dr. Weishaupt besonderen Dank aussprechen.

Niedersachsen (vgl. Avenarius u.a. 2001) durchgeführt und ein international wie national ausgerichteter forschungsbasierter Literaturbericht erstellt (siehe Brauckmann 2012). Ferner wurden teilstandardisierte Experteninterviews mit solchen Schulleitern und Schulleiterinnen durchgeführt, die nach Möglichkeit über Vorerfahrungen im Umgang mit der eigenverantwortlichen Schule verfügen. Daher wurden Schulleitungen kontaktiert, die sich mit ihren Schulen an den Modellvorhaben zur erweiterten schulischen Eigenverantwortung in den Ländern Berlin und Brandenburg im Zeitraum von 2003-2006 beteiligt hatten (vgl. Sroka u.a. 2007; Avenarius u.a. 2006).

Zudem sollten die Schulleitungen entweder Gymnasien oder aber Grundschulen vorstehen, die sich in relevanten einzelschulspezifischen Kontextbedingungen (Schulgröße, Ballungsraum, ländlicher Raum) unterscheiden sollten. Zielsetzung der Interviews war es, mehr über das Tätigkeitsspektrum und die Tätigkeitsdichte von Schulleitern und Schulleiterinnen sowie über ihre Arbeitsabläufe zu erfahren. Ferner sollten sie ihre Wahrnehmung von Belastungs- und Beanspruchungsmomenten schildern.

Mit jeweils einer Grundschul- und einer Gymnasialleitung in Berlin und Brandenburg wurden Gespräche geführt. Um die inhaltliche Trennschärfe der einzelnen Tätigkeitskategorien untereinander zu gewährleisten, wurden Sekundäranalysen durchgeführt. Die in der rechtlichen Analyse ermittelten und aus dem wissenschaftlichen Literaturbericht ergänzten Tätigkeitsfelder von Schulleitungen, die sich in den Sekundäranalysen bestätigt haben, konnten auch im Pretest mit sowohl veränderten als auch eigens formulierten und definierten Tätigkeiten identifiziert werden. Das anschließend entwickelte Instrument zur empirischen Tätigkeitsanalyse bildete den Kern des SHaRP-Schulleiterfragebogens, mit dem die Arbeitsbedingungen und die Tätigkeiten innerhalb ihres Aufgabenbereiches so genau wie möglich erfasst werden sollten. Dies gilt sowohl für den zeitlichen Aufwand von Schulleitungstätigkeiten als auch für das damit einhergehende Beanspruchungsempfinden.

Der Fragenkatalog an Schulleiter und Schulleiterinnen umfasst darüber hinaus die folgenden Bereiche:

I. Struktur und Organisation der Schule,
II. Schulleitung als Profession,
III. Arbeitssituation der Schulleiterin und des Schulleiters,
IV. Eigenschaften der Schulleiterin und des Schulleiters.

3.3 SHaRP-Haupterhebung

Nach dem in Niedersachsen erfolgreich durchgeführten Pretest zwecks Überprüfung der methodischen Güte des für die quantitative Forschungsphase entwickelten

Erhebungsinstrumentariums lief Anfang August die Haupterhebung in den Ländern Berlin, Brandenburg, Nordrhein-Westfalen, Niedersachsen, Hessen und Hamburg an.

3.3.1 Stichprobenbeschreibung

Die Stichprobenziehung erfolgte in zwei Erhebungswellen in den Bundesländern Nordrhein-Westfalen, Hamburg, Niedersachsen, Hessen, Brandenburg und Berlin. Alle Schulleiter und Schulleiterinnen sind aufgrund einer Zufallsauswahl aus den Gesamtlisten der Schulverzeichnisse ausgewählt worden. Die Teilnahme an der Erhebung erfolgte freiwillig. Von insgesamt 1.454 kontaktierten Grundschul- und Gymnasialleitungen waren 227 bereit, an der Befragung teilzunehmen. Dies entspricht einer länderübergreifenden Rücklaufquote von ca. 16 Prozent. Tabelle 1 zeigt die Verteilung der Personen auf die Schulformen nach Geschlecht in Prozent (11 Personen haben keine Angabe gemacht).

Tab. 1: Stichprobenverteilung (Anzahl, in %)

			Schulform				
			Grundschule	Gymnasium ohne Oberstufe	Gymnasium mit Oberstufe	Andere	Gesamt
Geschlecht	weiblich	Anzahl	87	1	29	7	124
		% i. v. Geschlecht	70.2%	.8%	23.4%	5.6%	100.0%
		% i. v. Schulform	79.1%	100.0%	30.9%	63.6%	57.4%
		% der Gesamtzahl	40.3%	.5%	13.4%	3.2%	57.4%
	männlich	Anzahl	23	0	65	4	92
		% i. v. Geschlecht	25.0%	.0%	70.7%	4.3%	100.0%
		% i. v. Schulform	20.9%	.0%	69.1%	36.4%	42.6%
		% der Gesamtzahl	10.6%	.0%	30.1%	1.9%	42.6%
Gesamt		Anzahl	110	1	94	11	216
		% i. v. Geschlecht	50.9%	.5%	43.5%	5.1%	100.0%
		% i. v. Schulform	100.0%	100.0%	100.0%	100.0%	100.0%
		% der Gesamtzahl	50.9%	.5%	43.5%	5.1%	100.0%

Quelle: Schulleiterdatensatz SHaRP, eigene Berechnungen

Die Mehrheit der befragten Schulleiter und Schulleiterinnen steht Grundschulen vor (ca. 51 Prozent). Diese werden vorwiegend von weiblichen Personen (70,2 Prozent; schulformunabhängig ist dies mit ca. 57 Prozent die größte Gruppe der Befragten) geleitet. Bei Gymnasien ist es umgekehrt; hier ist die Mehrheit der Befragten männlich (71 Prozent; diese stellen mit ca. 43 Prozent die zweitgrößte Gruppe der Befragten dar). Insgesamt sind ca. 44 Prozent der Befragten Schulleiter und Schulleiterinnen an Gymnasien tätig. Das Verhältnis von männlichen zu weiblichen Befragten beträgt ca. 44 Prozent zu 51 Prozent.

Es ist darauf hinzuweisen, dass die Schüleranzahl stark an die Schulform gebunden ist. Grundschulen weisen in der Regel weniger Schüler und Schülerinnen auf als Gymnasien.

Aufgrund der positiv selektierten Stichprobe (freiwillige Teilnahme) und der geringen Fallzahlen der befragten Schulleitungen pro Bundesland können die berechneten Koeffizienten und Mittelwertdifferenzen, trotz der Anwendung weitgehend standardisierter Instrumente und Verfahren, nur bedingt generalisiert werden.

So ist es z.B. möglich, dass die anhand der vorliegenden Werte durchgeführten Berechnungen aufgrund der nicht kontrollierbaren Durchführungsbedingungen und der Selektivität der Stichprobe (freiwillige Teilnahme der Schulleitungen) zu signifikanten Ergebnissen führen, obwohl in der Population kein Zusammenhang bzw. Unterschied besteht. Auch sind situative Einflüsse, die zum Zeitpunkt der Befragung mit in das Antwortverhalten der Befragten eingehen, nicht auszuschließen. Insofern können die Ergebnisse nur für einen bestimmten Zeitpunkt bzw. Zeitraum unter spezifischen Umständen gelten, die nur eingeschränkt Rückschlüsse auf wechselseitige Bedingungen im Antwortverhalten der Schulleiter und Schulleiterinnen zulassen.

Anderseits gelten diese Voraussetzungen jedoch für die meisten Fragebogenstudien (vgl. Lienert/Raatz 1998; zit. nach: Katzenburg 2008), so dass es fraglich ist, inwieweit die Aussagekraft der nachfolgend berichteten Befunde beeinträchtig ist.

3.3.2 Güte der selbst entwickelten Instrumente

Instrument zur Bestimmung von schulleitungsbezogenen Tätigkeitsfeldern

Mit den in der Haupterhebung bislang erhaltenen Daten (n=227; Dezember 2012) wurden ebenfalls Faktoren- und Reliabilitätsanalysen mit allen Skalen vorgenommen. Es zeigte sich im Vergleich zum Pretest, dass sich die Skalen u.a. aufgrund der größeren Stichprobe besser reproduzieren ließen, gerade auch im Hinblick auf die Faktorenstruktur und die Reliabilität.

Insgesamt ist für alle Instrumente von einer hohen Verlässlichkeit und Reproduzierbarkeit auszugehen, dies gilt insbesondere für jene Instrumente, die für die Erfassung der Tätigkeiten und der Führungsstile entwickelt wurden.

Tabelle 2 zeigt die entwickelten Tätigkeitsdimensionen und die Reliabilität der Skalen zur Erfassung der aufgewendeten Zeit. Zur Messung der Tätigkeiten im Sinne eines Belastungs- und Beanspruchungsmodells wurden aufgewendete Zeit je Tätigkeit und die damit einhergehende Beanspruchung differenziert. Ähnlich hohe Reliabilitäten ließen sich auch für die Beanspruchung feststellen. Das Tätigkeitsfeld „Eigener Unterricht" zeigt dort eine deutlich höhere Reliabilität (Cronbachs Alpha = 0.730), ebenso das Tätigkeitsfeld „Verwaltungs- und Organisationaufgaben" (Cronbachs Alpha = 0.864). Alle andere Tätigkeitsbereiche zur Erfassung der tätigkeitsspezifischen Beanspruchung zeigen eine ähnlich hohe interne Konsistenz im Vergleich zur Erfassung tätigkeitspezifisch aufgewendeter Zeit (Cronbachs Alpha > 0.7).

Tab. 2: Tätigkeitsfelder (aufgewendete Zeit pro Woche)

Skala	Beispiel-Item	Reliablität (Chronbachs Alpha)	Anzahl Items
Eigener Unterricht (EU)	Korrekturarbeiten	0.547	6
Unterrichtsbezogene Führungsarbeit (UFA)	Konfliktschlichtung zwischen Lehrer/inne/n und Schüler/inne/n sowie deren Eltern	0.812	7
Schüler- und elternbezogene Arbeit (SEA)	Eltern individuell beraten (Schullaufbahnberatung)	0.728	5
Personalführung und -entwicklung (PF/PE)	Beratung von Lehrkräften im Hinblick auf ihre berufliche Entwicklung	0.791	5
Organisationsführung und -entwicklung (OF/OE)	Mit dem Kollegium auf die Lösung gemeinsamer Probleme hinarbeiten	0.788	4
Verwaltungs- und Organisationsaufgaben (VOA)	Aufträge und Anfragen der Schulaufsichtsbehörden erledigen	0.649	9
Vertretung der Schule nach außen (VS)	Die Schule in der Öffentlichkeit vertreten (z.B. Schulwettbewerbe, Sportturniere)	0.837	5

Quelle: Schulleiterdatensatz SHaRP, eigene Berechnungen

Weitere Erhebungsinstrumente

Die auf der Grundlage verschiedener Studien und Instrumente (vgl. hierzu u.a. Sroka u.a. 2007; van Dick 2006; Häbler/Kunz 1985; Schubert 1983) entwickelte Itembatterie zur Erfassung schulischer Problemlagen befasst sich mit den von Schul-

leitern und Schulleiterinnen wahrgenommenen Problemen an ihrer Schule. In diesem Zusammenhang konnte überdies eine Skala gebildet werden, welche die derzeitige Belastung durch unterschiedliche Lernausgangslagen der Schülerschaft misst (Cronbachs Alpha = .873).

Auch bewährte Skalen, wie z.B. die der beruflichen Belastung nach Enzmann und Kleiber (1989) (Cronbachs Alpha = .868), wiesen durchweg zufriedenstellende Werte auf, wenn auch leichte Modifizierungen aufgrund von statistischen Gütekriterien vorgenommen werden mussten. Berufliche Belastung umfasst nach Enzmann und Kleiber (1989) die Arbeitsunzufriedenheit, das Kontrolliertheitserleben und die Arbeitsüberforderung. Das Instrument ist ursprünglich bei Fachkräften in der psychosozialen Versorgung entwickelt und eingesetzt wurden.

4. Ergebnisse

Im folgenden Ergebnisteil steht das tätigkeitsbezogene Belastungs- und Beanspruchungserleben von Schulleitern und Schulleiterinnen im Vordergrund. Darüber hinaus wurden auf individueller Ebene Zusammenhänge zwischen Belastungserleben und weiteren Kontextmerkmalen wie z.B. der Schulform untersucht.

Zeitlicher Aufwand und Beanspruchungsempfinden

Den überwiegenden Teil ihrer Zeit verwenden Schulleitungen in den Tätigkeitsfeldern „Eigener Unterricht", „Unterrichtsbezogene Führungsarbeit" sowie „Verwaltungs- und Organisationsaufgaben", wobei die angegebenen Zeitanteile der Schulleiter und Schulleiterinnen, insbesondere was ihre eigene Unterrichtsverpflichtung angeht, stark streuen. Die Antwortdifferenzen verweisen auf die schulform- sowie länderspezifischen Regelungen zur Unterrichtsverpflichtung von Schulleitungen. Weniger Zeit widmen Schulleiter und Schulleiterinnen den Tätigkeitsfeldern „Organisationsführung und -entwicklung", „Schüler- und elternbezogene Arbeit", „Personalführung und -entwicklung" sowie Tätigkeiten, die der schulischen Außendarstellung dienen (vgl. Abb. 2).

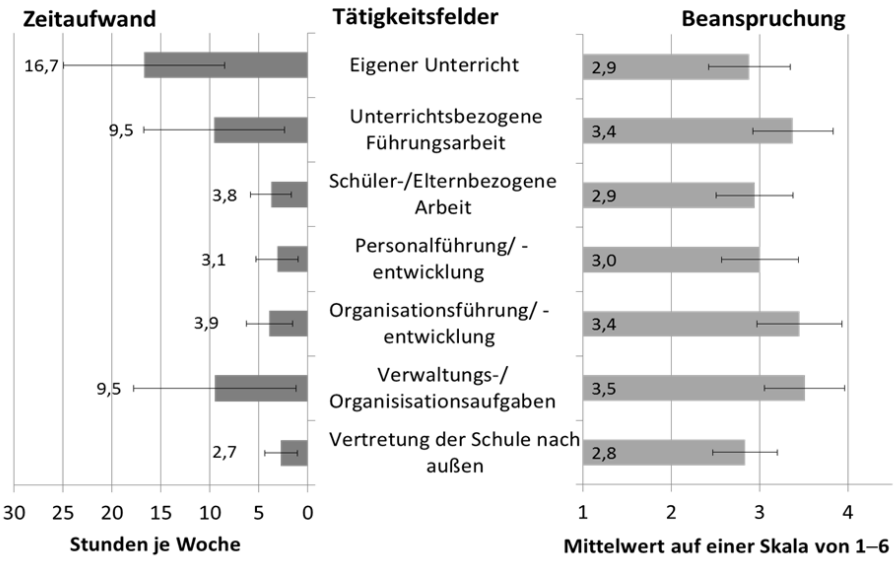
Generell kann festgestellt werden, dass sich der zeitliche Aufwand und das Beanspruchungsempfinden von Schulleitern und Schulleiterinnen in einzelnen Tätigkeitsfeldern deutlich voneinander unterscheiden. So sind Tätigkeitsfelder wie „Personalführung und -entwicklung" und „Organisationsführung und -entwicklung" durch starke Beanspruchung, aber eher geringe zeitliche Belastung gekennzeichnet.

Das Tätigkeitsfeld „Eigener Unterricht“ wiederum ist zwar durch einen hohen zeitlichen Aufwand, aber eine weniger stark ausgeprägte Beanspruchung charakterisiert.

Bei den Tätigkeitsfeldern „Verwaltungs- und Organisationsaufgaben“ und „Unterrichtsbezogene Führungsarbeit“ kann möglicherweise am ehesten von einer Deckungsgleichheit zwischen zeitlicher Belastung und empfundener Beanspruchung gesprochen werden. Schulleiter und Schulleiterinnen geben „Verwaltungs- und Organisationsaufgaben“ als zentrale Belastungsgröße an.

Tätigkeitsfeldübergreifend wird erkennbar, dass die insgesamt durchschnittlich aufgewendete Zeit 49,3 Stunden pro Woche (SD=43.84) beträgt, wobei die zeitlichen Angaben zwischen den einzelnen Schulleitungen z.T. jedoch beträchtlich variieren. Das Beanspruchungserleben von Schulleitern und Schulleiterinnen mit Blick auf das gesamte Tätigkeitsspektrum fällt mit einem Mittelwert von 3,2 (SD=.734) eher moderat aus.

Abb. 2: Zeitlicher Aufwand und Beanspruchungsempfinden



Quelle: Schulleiterdatensatz SHaRP, eigene Berechnungen

In weitergehenden, korrelativen Befunden (vgl. Tab. 3) zeigt sich zudem, dass in nur zwei der sieben Tätigkeitsfelder die Tätigkeitsausübung und die tätigkeitsspezifische Belastung in einem Zusammenhang zueinander stehen. Dies sind die Tätigkeitsfelder „Eigener Unterricht“ sowie „Vertretung der Schule nach außen“. Weiterhin ist dieser Tabelle zu entnehmen, dass Schulleitungen, die eine hohe zeitliche Belastung im

Tätigkeitsfeld „Eigener Unterricht“ aufweisen, sich auch in nahezu allen anderen Tätigkeitsfeldern stärker beansprucht fühlen. Schulleiter und Schulleiterinnen, die viel Zeit für die Außendarstellung ihrer Schule aufwenden, empfinden eine geringere Beanspruchung hinsichtlich der Tätigkeitsfelder „Eigener Unterricht“ und „Schüler- und elternbezogene Arbeit“.

Tab. 3: Zeitlicher Aufwand und Beanspruchungsempfinden

		erlebte Belastung						
		EU	UFA	SEA	PF/PE	OFOE	VOA	VS
aufgewendete Zeit	EU	**.318****	**.234****	**.195***	.093	.006	**.212***	**.180***
	UFA	.040	.132	.033	-.005	.029	.043	-.089
	SEA	-.037	.133	.130	.054	.084	.070	-.096
	PF/PE	-.140	-.027	-.115	.061	-.011	-.019	-.037
	OF/OE	-.147	-.138	-.146	.027	.022	-.019	.054
	VOA	-.048	.016	-.010	-.083	.122	.143	-.028
	VS	**-.168***	-.087	**-.183****	-.106	-.083	-.016	-.112

Anmerkungen. EU = eigener Unterricht, UFA = Unterrichtsbezogene Führungsarbeit, SEA = Schüler- und elternbezogene Arbeiten, PF/PE = Personalführung und -entwicklung, OF/OE = Organisationsführung und -entwicklung, VOA = Verwaltungs- und Organisationsaufgaben, VS = Vertretung der Schule nach außen.

* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Quelle: Schulleiterdatensatz SHaRP, eigene Berechnungen

Auch die korrelativen Befunde zur Tätigkeitsausübung und empfundenen beruflichen Belastung (vgl. Tab. 4; zum Konzept der beruflichen Belastung vgl. Enzmann und Kleiber 1989) zeigen, dass mit Ausnahme des Tätigkeitsfeldes „Verwaltungs- und Organisationsaufgaben“ keine statistisch signifikanten Zusammenhänge zwischen Tätigkeitsausübung und der beruflichen Belastung festzustellen sind. Weitergehend fällt auf, dass sich die zeitliche Belastung für Verwaltungs- und Organisationsaufgaben negativ auf die Wahrnehmung des selbstverantworteten Handelns auswirkt. Vielmehr fühlen sich die betreffenden Schulleiter und Schulleiterinnen stärker kontrolliert und weisen auch eine höhere Arbeitsunzufriedenheit auf.

Auch das verstärkte Kontrolliertheitserleben von Schulleitern und Schulleiterinnen innerhalb ihres eigenen Unterrichts kann als Indiz dafür verstanden werden, dass sie sich in ihrer Handlungsautonomie eingeschränkt fühlen.

Tab. 4: Zeitlicher Aufwand und berufliche Belastung

		Belastung	Arbeits-überforde-rung	Kontroll-iertheits-erleben	Arbeits-unzufrieden-heit
aufgewendete Zeit	EU	.113	.070	**.288****	.115
	UFA	-.119	-.135	-.027	-.080
	SEA	-.017	-.047	.106	.013
	PF/PE	-.108	-.107	-.048	-.059
	OF/OE	-.060	-.064	-.028	-.069
	VOA	**.161***	.060	**.238****	**.242****
	VS	-.056	-.032	-.012	-.058

Anmerkungen. EU = eigener Unterricht, UFA = Unterrichtsbezogene Führungsarbeit, SEA = Schüler- und elternbezogene Arbeiten, PF/PE = Personalführung und -entwicklung, OF/OE = Organisationsführung und -entwicklung, VOA = Verwaltungs- und Organisationsaufgaben, VS = Vertretung der Schule nach außen.

* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

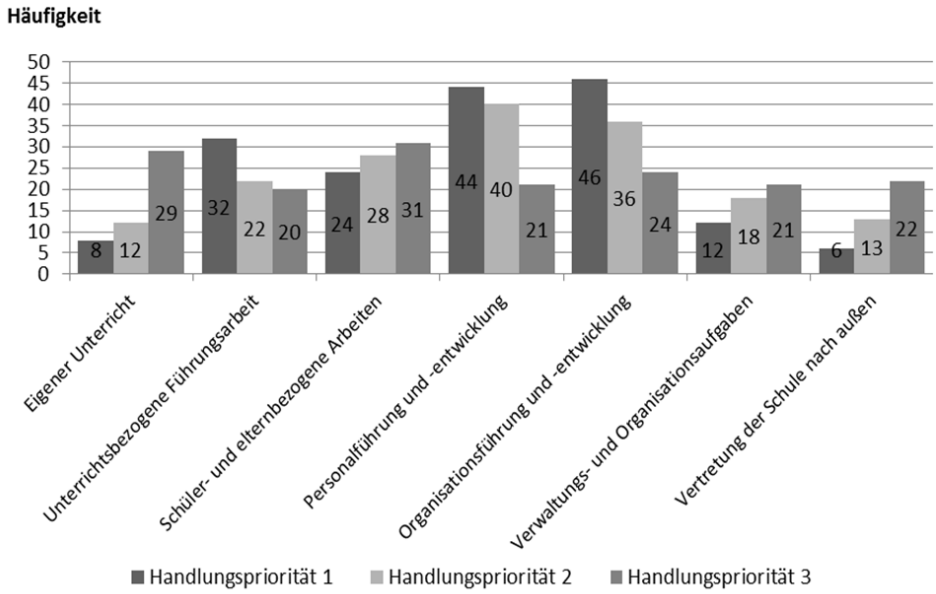
** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Quelle: Schulleiterdatensatz SHaRP, eigene Berechnungen

Anscheinend wird für die Bewertung der beruflichen Belastung mehr als nur die aufgewendete Zeit und die damit einhergehende Beanspruchung herangezogen. Das Belastungsempfinden im Tätigkeitsfeld „Verwaltungs- und Organisationsaufgaben" scheint mit globaleren Einflussgrößen und weniger mit spezifischen zeitlichen Tätigkeitsausübungen in Verbindung zu stehen.

Ferner wurden Schulleiter und Schulleiterinnen danach gefragt (vgl. Abb. 3), welchen Tätigkeitsfeldern sie für ihr Schulleitungshandeln die größte Bedeutung einräumen würden, und dementsprechend um eine Priorisierung gebeten. Die Schulleiter und Schulleiterinnen zählten zu ihren drei wichtigsten Handlungsprioritäten die Tätigkeitsfelder „Personalführung und -entwicklung", „Organisationsführung und -entwicklung" sowie die „Unterrichtsbezogene Führungsarbeit". Damit zeigt sich interessanterweise, dass bei Schulleitungen jene Tätigkeitsfelder die höchste Handlungspriorität genießen, in denen sie zum einen eher hohe Belastungswerte angeben, für die sie zum anderen (mit Ausnahme der unterrichtsbezogenen Führungstätigkeiten) aber auch weniger Zeit aufwenden.

Abb. 3: Handlungsprioritäten der Schulleiter und Schulleiterinnen bezogen auf die Tätigkeitsfelder



Quelle: Schulleiterdatensatz SHaRP, eigene Berechnungen

Aus der Gegenüberstellung von Abbildung 2 und 3 lässt sich erkennen, dass eine deutliche Diskrepanz zwischen den normativen Vorstellungen oder aber Wünschbarkeiten hinsichtlich eines Schulleiters bzw. einer Schulleiterin an einer stärker eigenverantwortlichen Schule und den tatsächlich vollzogenen Tätigkeiten an der eigenen Schule besteht.

Kontextmerkmale und Belastungserleben

In Anlehnung an das Belastungs- und Beanspruchungsmodell von Böhm-Kasper (2004) ist zu berücksichtigen, dass auch die inner- und außerschulischen Kontextbedingungen einen Einfluss auf die subjektiv erlebte Beanspruchung der Schulleitungen haben. Schulleiter und Schulleiterinnen wurden auf einer vierstufigen Antwortskala danach gefragt, wie stark sie durch bestimmte Merkmale inner- und außerschulischer Probleme belastet sind. Die zehn Probleme (getrennt nach Schulform), die Schulleitungen am stärksten belasten, verdeutlicht Tabelle 5.

Tab. 5: Schulische Probleme an Grundschulen und Gymnasien

Grundschulen			Gymnasien		
Problemlage	M	SD	Problemlage	M	SD
1. allgemeine Arbeitsbelastung der Lehrkräfte	4,63	1,09	1. allgemeine Arbeitsbelastung der Lehrkräfte	4,32	1,02
2. zunehmende Einflussnahme der Eltern	4,24	1,30	2. großer Anforderungsdruck von außen (Öffentlichkeit, Schulverwaltung, Partner usw.)	3,63	1,22
3. Erzieherische Problemlagen (Verhaltensauffälligkeiten)	4,18	1,38	3. Konkurrenz zu anderen Schulen	3,57	1,49
4. großer Anforderungsdruck von außen (Öffentlichkeit, Schulverwaltung, Partner usw.)	3,82	1,54	4. unzureichende Ausstattung	3,50	1,64
5. nicht bedarfsgerechte Personalversorgung	3,73	1,63	5. zunehmende Einflussnahme der Eltern	3,49	1,28
6. unterschiedliche Lernvoraussetzungen der Schüler/innen	3,53	1,66	6. nicht bedarfsgerechte Personalversorgung	3,44	1,41
7. unzureichende Lernförderung der Schüler/innen durch das Elternhaus	3,51	1,47	7. fehlende Innovationsbereitschaft von Lehrkräften	3,42	1,16
8. mangelnde Motivation bzw. Konzentration der Schüler/innen	3,41	1,59	8. Ärger mit Behörden bzw. Institutionen (z.B. Schulbehörde bzw. Jugendamt)	3,29	1,26
9. unzureichende Ausstattung	3,34	1,62	9. unzureichende Lernförderung der Schüler/innen durch das Elternhaus	3,23	1,07
10. Ärger mit Behörden bzw. Institutionen (z.B. Schulbehörde bzw. Jugendamt)	3,33	1,62	10. unterschiedliche Lernvoraussetzungen der Schüler/innen	3,23	0,98

Quelle: Schulleiterdatensatz SHaRP, eigene Berechnungen

In den Befunden kommen zum Teil strukturelle, vor allem schulextern zu verantwortende Defizite (z.B. bedarfsgerechte Unterrichtsversorgung, Arbeitsbelastung, Anforderungsdruck von außen), aber auch Problemlagen, die eher sozialen Ursachen zugeschrieben werden können (z.B. Drogen, Gewalt, unterschiedliche Lernvoraussetzungen, mangelnde Lernförderung durch die Eltern), zum Ausdruck. Die z.T. großen Streuwerte in den Antwortmustern der Schulleitungen bilden auch die differierenden einzelschulspezifischen Kontextbedingungen ab.

Die deutlichsten schulformspezifischen Abweichungen liegen zum einen in der Benennung der wahrgenommenen Probleme. Grundschulleitungen benennen im Gegensatz zu Gymnasialschulleitungen das Problemfeld der Lernschwierigkeiten

(Verhaltenssaufälligkeiten, mangelnde Motivation bzw. Konzentration der Schüler und Schülerinnen, unterschiedliche Lernvoraussetzungen der Schüler und Schülerinnen). Gymnasialschulleitungen wiederum betonen den Umgang mit den Ansprüchen außerschulischer Partner bzw. der außerschulischen Öffentlichkeit.

Zum anderen liegen die deutlichsten Unterschiede zwischen den beiden Schulformen in der Intensität des Belastungsempfindens. Selbst bei der deckungsgleichen Nennung von Problemen fallen die Einschätzungen der Grundschulleiter und -leiterinnen durchweg kritischer aus als die der Gymnasialschulleiter und -leiterinnen. Dies gilt auch für den größten Belastungsfaktor (die Arbeitssituation der Lehrkräfte). Eine Ausnahme bildet die Einschätzung der sachlich-materiellen Ausstattung der Schule, die von den Leitern und Leiterinnen beider Schulformen ähnlich bewertet wird.

Nachfolgend wird dem Einfluss schulischer Probleme auf das Belastungserleben von Schulleitern und -leiterinnen nochmals differenzierter nachgegangen, indem Zusammenhänge zwischen dem Problemfeld *unterschiedliche Lernausgangslagen der Schülerschaft* und tätigkeitsspezifischer Belastung dargestellt werden (vgl. Tab. 6). Es wird erkennbar, dass an Gymnasien mit stark unterschiedlich ausgeprägten Lernausgangslagen eine stärkere Beanspruchung in den Tätigkeitsfeldern „Unterrichtsbezogene Führung“ und „Eigener Unterricht“ zu beobachten ist.

Tab. 6: Zusammenhänge zwischen dem Problemfeld „Unterschiedliche Lernausgangslagen der Schülerschaft“ und dem Beanspruchungsempfinden

Tätigkeitsfeldbezogenes Beanspruchungsempfinden	Problemfeld „Unterschiedliche Lernausgangslagen der Schülerschaft“	
	Grundschule	Gymnasium
Eigener Unterricht	,097	**,296***
Unterrichtsbezogene Führung	,077	**,287***
Schüler- und elternbezogene Arbeit	,130	,231
Personalführung und -entwicklung	-,070	,141
Organisationsführung und -entwicklung	-,055	,115
Verwaltungs- und Organisationsaufgaben	,091	,193
Vertretung der Schule nach außen	,091	,078

Anmerkungen. * Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Quelle: Schulleiterdatensatz SHaRP, eigene Berechnungen

Möglicherweise stellt an Gymnasien mit einer z.T. heterogener zusammengesetzten Schülerschaft eine adaptiv zu betreibende Unterrichtsführung eine neue Anforderung für die jeweiligen Schulleitung dar, ob als Lehrkraft in seiner bzw. ihrer Klasse oder aber als der so oft neuerdings beschworene *instructional leader* seiner bzw. ihrer Organisation. Mit diesen unterschiedlichen Lernvoraussetzungen an einer Schule pro-

duktiv umzugehen, mag nur bedingt dem traditionellen beruflichen Selbstverständnis eines Gymnasialschulleiters bzw. einer Gymnasialschulleiterin entsprechen, der bzw. die gerade aufgrund seiner oder ihrer geringeren Unterrichtsverpflichtung auch andere Tätigkeiten ausüben sollte bzw. solche ausüben will.

An Grundschulen hingegen weisen Schulen mit einem höheren Problemdruck keinen Zusammenhang mit dem Beanspruchungsempfinden in den sieben Tätigkeitsfeldern auf. Es kann an dieser Stelle nicht abschließend geklärt werden, worauf dieser Unterschied zurückzuführen ist.

Zu vermuten ist, dass an Grundschulen die Auseinandersetzung mit Heterogenität in der alltäglichen Berufswelt stärker verankert ist und keine neue Herausforderung darstellt bzw. eine Form der Dissonanz erzeugen könnte. Vielmehr verfügen die Schulleiter und -leiterinnen über umfangreichere Erfahrungen, sei es durch extern oder intern organisierte Unterstützungsmaßnahmen. Möglicherweise federt auch das tradierte berufliche Selbstverständnis das potentielle Belastungserleben ab, indem die Übernahme von Erziehungsverantwortung für eine wesentlich jüngere Schülerschaft als etwas eher Selbstverständliches erachtet wird. Auch empfinden sie womöglich die sie umgebenden Rahmenbedingungen als etwas, dass ihrer Arbeit inhärent ist und folglich keine weitere systematisch betriebene Selbstreflexion lohnt.

5. Zusammenfassung und abschließende Bemerkungen

Die neuen Steuerungsansätze und -instrumente weisen den Schulleitungen ein erhöhtes Maß an Entscheidungskompetenzen zu. Deshalb verdienen Rahmenbedingungen und Rolle der Schulleitung besondere Aufmerksamkeit. Wichtige Aufschlüsse darüber liefern sowohl Angaben zu der zeitlichen Inanspruchnahme von Schulleitern und Schulleiterinnen in verschiedenen Tätigkeitsbereichen als auch ihre damit verbundene persönliche Belastungswahrnehmung. Festzuhalten ist, dass nicht nur zeitlicher Aufwand und persönliches Beanspruchungsempfinden in den einzelnen Tätigkeitsfeldern deutlich voneinander abweichen, sondern darüber hinaus auch nur in zwei Tätigkeitsfeldern Hinweise auf Zusammenhänge zwischen dem, was Schulleitungen tun, und dem, was sie dabei als Beanspruchung empfinden, festzustellen waren. Hervorzuheben ist, dass Schulleiter und Schulleiterinnen den weitaus größten Teil ihrer Arbeitszeit unterrichtsbezogenen sowie Verwaltungs- und Organisationstätigkeiten widmen.

Darüber hinaus konnte gezeigt werden, dass auch die Einschätzung einer Tätigkeit mitentscheidend dafür sein kann, welche Tätigkeiten als belastend wahrgenommen werden. Neben der Einschätzung und Wahrnehmung einer Tätigkeit liefert auch der schulspezifische Problemdruck, dem sich Schulleiter und Schulleiterinnen gegenüber-

sehen, einen eigenständigen Erklärungsbeitrag, wenn es um das Belastungsempfinden von Gymnasialschulleitungen in zwei Tätigkeitsfeldern geht. Diese Tätigkeitsfelder betreffen die Rolle der Schulleiter und Schulleiterinnen als autonom pädagogisch Handelnde, als Lehrkraft bzw. als pädagogische Führungskraft.

Anscheinend handelt es sich beim Belastungs- und Beanspruchungserleben um ein vorwiegend individuell verankertes und über verschiedenartig bewältigte oder aber nicht bewältigte Interaktionen vermitteltes Phänomen (vgl. Warwas 2013), das in nur sehr eingeschränkter Weise verallgemeinert werden kann. Eine eindeutige Trennung von Situation und Bewertung der Situation ist quantitativ anhand nur eines Messzeitpunktes schwer abzubilden, da vielmehr prozessuale Variablen in den Blick genommen werden müssten. So ist auch weiterhin davon auszugehen, dass subjektive Annahmen über und Vorstellungen von Belastung neben situationellen Kontextfaktoren das Verständnis, Empfinden und die persönliche Wahrnehmung von Belastung leiten. Somit kommt den subjektiven Überzeugungen von Schulleitern und Schulleiterinnen nicht nur eine handlungsleitende, sondern auch eine belastungssteuernde Funktion zu.

Vor diesem Hintergrund sind die angegebenen Belastungswerte ohne genauere Hintergrundinformationen nur sehr bedingt interpretierbar.

Der Aktionsradius von Schulen und Schulleitungen variiert mitunter stark in Abhängigkeit von jenen Kontextfaktoren, die der Schulsystemebene vorgelagert sind (vgl. OECD 2004). Hierzu zählen z.B. bislang zu wenig betrachtete historische Traditionen (vgl. Strutz/Nevermann 1985; Brown 2005) sowie soziale und wirtschaftliche Rahmenbedingungen, z.B. *globalisation, advances in science and technology, pressure on the environment, changes in demography* (vgl. Mulford 2008). Sie können sich hinsichtlich ihrer Ausprägungsgrade, ihres Entwicklungstempos und ihrer praktischen Bedeutsamkeit deutlich unterscheiden und als Ermöglichungs- oder aber Eingrenzungsfaktoren neue Anforderungen an die Schulen und damit verbundene Leitungsmodelle (vgl. Vadeboncoeur 2006; Gödde 2011) stellen.

Anhand der Befunde war es jedoch möglich, Zusammenhangsstrukturen zwischen bestimmten Tätigkeitsmerkmalen und Belastungs- bzw. Beanspruchungserleben zu prüfen und darzustellen.

Daher müsste mittels des Einsatzes qualitativer Forschungsmethoden künftig verstärkt danach gefragt werden, welche Erfahrungen die Schulleiter und Schulleiterinnen bisher mit Blick auf die mit den „Neuen" Steuerungsanasätzen verknüpften Erwartungen gemacht haben, vor allem im Hinblick auf Belastungsmomente. Zentrales Thema könnte dabei das schulische Innenverhältnis werden, insbesondere der Umgang mit der neudefinierten Rolle einer zunehmend ausdifferenzierten Schulleitung, der Mitwirkung von Lehrkräften, Schülerschaft und Eltern an schulischen und pädagogi-

schen Entscheidungsprozessen sowie der von Schulleitern und Schulleiterinnen spezifizierte Unterstützungs- bzw. Professionalisierungsbedarf.

Unter einer transferorientierten Perspektive könnte zudem die Schlussfolgerung gezogen werden, dass einzelschulische Problemlagen eben auch einzelfallspezifische oder zumindest flexible Lösungen und Unterstützungsmaßnahmen erfordern, auch bei der Systemsteuerung.

Literatur und Internetquellen

Altrichter, H./Maag Merki K. (2010): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS.

Avenarius, H./Brauckmann, S./Döbert, H./Isermann, K./Kimmig, T./Seeber, S. (2006): Durch größere Eigenverantwortlichkeit zu besseren Schulen. Ergebnisbericht der wissenschaftlichen Begleitung des „Modellvorhabens eigenverantwortliche Schule (MES)" im Land Berlin. Frankfurt a.M. u.a.: DIPF.

Avenarius, H./Döbert, H./Knauss, G./Weishaupt, H./Weiß, M. (2001): Stand und Perspektiven der Orientierungsstufe in Niedersachsen. Gutachten im Auftrag des Niedersächsischen Kultusministeriums. Frankfurt a.M.: DIPF.

Böhm-Kasper, O. (2004): Schulische Belastung und Beanspruchung. Eine Untersuchung von Schülern und Lehrern am Gymnasium. Münster u.a.: Waxmann.

Bonsen, M./Gathen, J. von der/Iglhaut, C./Pfeiffer, H. (2002): Die Wirksamkeit von Schulleitung. Empirische Annäherungen an ein Gesamtmodell schulischen Leitungshandelns. Weinheim: Juventa.

Bonsen, M./Iglhaut, C./Pfeiffer, H. (1999): Schulleitungshandeln aus Schulleitungssicht. Dortmund: IFS-Verlag.

Brauckmann, S. (2012): Schulleitungshandeln zwischen deconcentration, devolution und delegation (3D). Empirische Annäherungen aus internationaler Perspektive. In: Empirische Pädagogik 26, H. 1, S. 76-100.

Brauckmann, S./Herrmann, C. (2012): Schulleitungshandeln im Rahmen Neuer Steuerung. Belastung oder Chance? In: Zeitschrift für Bildungsverwaltung, H. 1, S. 87-98.

Brown, K. (2005): Pivotal Points. History, Development, and Promise of the Principalship. In: English, F. (Hrsg.): The Handbook of Educational Leadership. Thousand Oaks, CA: Sage, S. 109-141.

Day, C./Sammons, P./Hopkins, D./Harris, A./Leithwood, K./Gu, Q. u.a. (2009): The Impact of School Leadership on Pupil Outcomes. Final Report. University of Nottingham.

Dworking, A.G. (2009): Teacher Burnout and Teacher Resilience: Assessing the Impacts of the School Accountability Movement. In: Saha, L.J./Dworking, A.G. (Hrsg.): International Handbook of Research on Teachers and Teaching, Part 1. New York, NY: Springer Science + Business Media, S. 491-509.

Enzmann, D./Kleiber, D. (1989): Helfer-Leiden: Stress und Burnout in psychosozialen Berufen. Heidelberg: Asanger.

Füssel, H.-P. (2009): Schulaufsicht und Schulentwicklung. Die neue Rolle der Schulleitung in einer selbständigen Schule aus juristischer Sicht (überarbeitete Fassung Mai 2009). Vortrag, gehalten im Rahmen von: Fachanhörung des Bayrischen Lehrer- und Lehrerinnenverbandes „Schulleitung innovativ – Warum Schulleitungen Motor der Schulentwicklung sind". München, 28.04.2009. URL: http://www.bllv.de/filead

min/Dateien/Land-PDF/BLLV/Fuessel.Schulaufsicht_und_Schulentwicklung-1.pdf; Zugriffsdatum: 04.06.2012.

Gödde, B. (2011): Das Berufsfeld eines Schulleiters im Spannungsfeld von gesellschaftlichen Ansprüchen und systemimmanenten Bedingungen. In: Krüger, S. (Hrsg.): „Werde selbstständig und ich sage dir wie!" Bildung und die Anforderungen an Führungspersonal. Dokumentation der 31. DGBV-Jahrestagung vom 04.-06.11.2010 in Erfurt. Berlin: DGBV.

Goldring, E./Huff, J./May, H./Camburn, E. (2006): School Context and Individual Characteristics: What Influences Principal Practice? Vortrag, gehalten im Rahmen von: University Council of Educational Administration Annual Conference, San Antonio, Texas.

Goldring, E./Huff, J./May, H./Camburn, E. (2008): School Context and Individual Characteristics. What Influences Principal Practice? In: Journal of Educational Administration 46, H. 3, S. 332-352.

Häbler, H./Kunz, A. (1985): Qualität der Arbeit und Verkürzung der Arbeitszeit in Schule und Hochschule. Eine empirische Untersuchung im Auftrag der Max-Traeger-Stiftung. München: IMU-Institut.

Hanßen, K. (2012): Rechtliche Regelungen zu Tätigkeitsfeldern von Schulleiterinnen und Schulleitern bei erweiterter Eigenverantwortung von Schulen (unveröffentlichtes Manuskript).

Hasenbank, T. (2000): Führung und Leitung einer Schule als Dimension und Rahmenbedingung Berufsschulischer Entwicklung, München: Dissertation.

Horng, E.L./Klasik, D./Loeb, S. (2009): Principal Time-Use and School Effectiveness. School Leadership Research Report No. 09-3. Stanford University, Institute for Research on Education Policy & Practice., Stanford, CA.

Katzenburg, M. (2008): Die Relevanz von sozialer Kompetenz für den Lehrerberuf. Unveröffentlichte Diplomarbeit, Westfälische Wilhelms-Universität Münster.

Landert und Partner (2009): Arbeitssituation der Schulleitungen im Kanton Thurgau. Evaluation im Auftrag des Departments Erziehung und Kultur des Kantons Thurgau. Zürich: Amt für Volksschule.

Leithwood, K.A./Menzies, T./Jantzi, D./Leithwood, J. (1999): Teacher Burnout: A Critical Challenge for Leaders of Restructuring Schools. In Vandenberghe, R./Huberman, A.M. (Hrsg.): Understanding and Preventing Teacher Burnout. A Sourcebook of International Research and Practice. Cambridge: Cambridge University Press, S. 85-114.

Lienert, G.A./Raatz, R. ([6]1998): Testaufbau und Testanalyse. Weinheim: Psychologie Verlags Union.

Louis, K.S./Leithwood, K./Wahlstrom, K.L./Anderson, S.E. (2010): Learning from Leadership: Investigating the Links to Improved Student Learning. Final Report of Research to the Wallace Foundation. University of Minnesota, University of Toronto.

Marzano, R./Waters, T./McNulty, B.A. (2005): School Leadership That Works. From Research to Results. Alexandria, VA/Aurora, CO: Association for Supervision and Curriculum Development, Mid-continent Research for Education and Learning.

May, H./Goldring, E.B./Hiff, J. (2008): A Longitudinal Study of Principals Activities and Student Performance. Vortrag, gehalten im Rahmen von: Annual Meeting of the American Education Research Association, San Diego, CA, 16.04.2008.

Mulford, B. (2008): The Leadership Challenge. Improving Learning in Schools. Camberwell, Victoria, Australia: ACER Press.

OECD (2004): Bildung auf einen Blick 2004. OECD Indikatoren. Paris: OECD.

OECD (2008): Improving School Leadership Policy and Practice. Paris: OECD.

OECD (2009): Leading to Learn: School Leadership and Management Styles. In: OECD (Hrsg.): Creating Effective Teaching and Learning Environments: First Results from TALIS. Paris: OECD, S. 190-218.

Robinson, V.M.J./Lloyd C.A./Rowe K.J. (2008): The Impact of Leadership on Students Outcomes: an Analysis of the Differential Effects of Leadership Types. In: Education Administration Quarterly 44, S. 635-674.

Rosenbusch, H.S./Braun-Bau, S./Warwas, J. (2007): Die Entdeckung der Schulleitung und deren zukünftige Entwicklung. Ergebnisse eines Forschungsprojekts zur Schulleitungstätigkeit an Grund-, Haupt- und Realschulen. In: SchulVerwaltung (Bayern) 30, H. 4, S. 98-104.

Rosenbusch, H.S./Wissinger, J. (Hrsg.) (1989): Schulleiter zwischen Administration und Innovation. Braunschweig: SL Verlag.

Rudow, B. (1994): Die Arbeit des Lehrers. Zur Psychologie der Lehrertätigkeit, Lehrerbelastung und Lehrergesundheit. Bern u.a.: Hans Huber.

Rürup, M. (2007): Innovationswege im deutschen Bildungssystem. Die Verbreitung der Idee „Schulautonomie" im Ländervergleich. Wiesbaden: VS.

Schubert, R. (1983): Feldstudie zur Phänomenologie von psychopathogen relevanten Belastungssituationen. Leipzig: unveröff. Diplomarbeit der Sektion Psychologie.

Senkbeil, M./Drechsel, B./Rolff, H.-G/Bonsen, M./Zimmer, K./Lehmann, R.H./Neumann, A. (2004): Merkmale und Wahrnehmungen von Schule und Unterricht. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs. Münster u.a.: Waxmann, S. 296-313.

Sroka, W./Avenarius, H./Kampa, N./Döbert, H./Isermann, K./Seeber, S. (2007): Die Entwicklung zu selbstständigen Schulen im Land Brandenburg. Erfahrungen und Ergebnisse. Ergebnisbericht der wissenschaftlichen Begleitung des Modellvorhabens „Stärkung der Selbstständigkeit von Schulen" (MoSeS) im Land Brandenburg. Frankfurt a.M.: DIPF.

Strutz, S./Nevermann, K. (1985): Schulleitung – historisch gesehen. Braunschweig: Westermann.

Vadeboncoeur, J.A. (2006): Chapter 7: Engaging Young People: Learning in Informal Contexts. In: Green, J./Luke, A. (Hrsg.): Review of Research in Education. Washington, DC: American Educational Research Association, S. 239-278.

Van Dick, R. ([2]2006): Stress und Arbeitszufriedenheit bei Lehrerinnen und Lehrern. Zwischen „Horrorjob" und Erfüllung. Leicht veränderte Aufl. Marburg: Tectum.

Warwas, J. (2013): Einflüsse von professionellen Bewältigungsstrategien und Teamstrukturen auf das Belastungserleben schulischer Führungskräfte. In: Keller-Schneider, M./Albisser, S./Wissinger, J. (Hrsg.): Professionalität und Kooperation in Schulen. Beiträge zur Diskussion über Schulqualität. Bad Heilbrunn: Klinkhardt, S. 209-226.

Winkler, C. (2009): Outputsteuerung und Unterstützungssysteme für Lehrkräfte im belasteten Schulalltag. In: Zeitschrift für Bildungsverwaltung, H. 2, S. 29-42.

Wissinger, J. (2011): Schulleitung und Schulleitungshandeln. In: Terhart, E./Bennewitz, H./Rothland, M. (Hrsg.): Handbuch der Forschung zum Lehrerberuf. Münster u.a.: Waxmann, S. 98-115.

Witziers, B./Bosker, R.J./Kruger, M.L. (2003): Educational Leadership and Student Achievement. The Elusive Search for an Association. In: Educational Leadership Quarterly 39, H. 3, S. 398-425.

Stefan Brauckmann, Dr., geb. 1973, Projektleiter SHaRP, DIPF Berlin, AE Steuerung und Finanzierung des Bildungswesens.
E-Mail: brauckmann@dipf.de

Christoph Herrmann, Dipl. Päd., geb. 1983, wissenschaftlicher Mitarbeiter, DIPF Berlin, AE Steuerung und Finanzierung des Bildungswesens.
E-Mail: herrmann@dipf.de

Anschrift: DIPF, Warschauer Str. 36, 10243 Berlin

DDS – Die Deutsche Schule
Beiheft 12, S. 197-225
© 2013 Waxmann

Katja Thillmann/Anabel Bach/Christina Gerl/Felicitas Thiel

Schulische Selbststeuerung durch interne Evaluation

Eine Analyse von Evaluationsprogrammen Berliner und Brandenburger Gymnasien

Zusammenfassung

Auf der Basis einer Analyse von Evaluationsprogrammen und einer Schulleiterbefragung wird untersucht, ob interne Evaluation als systematisches Instrument schulischer Selbststeuerung genutzt wird, welche Unterschiede sich zwischen den untersuchten Bundesländern zeigen und von welchen organisationalen Voraussetzungen die Programmqualität abhängt. Die Ergebnisse zeigen große Bundeslandunterschiede und verweisen auf den hohen Stellenwert der Schulleitung für die Programmarbeit.
Schlüsselwörter: Schulinterne Evaluation, Schulprogramme, data-based decision making, Steuerung

Schools' Self-Governance by Using Internal Evaluation

An Analysis of Evaluation Programs at Schools in Berlin and Brandenburg

Summary

Based on an analysis of evaluation plans and an online survey of principals, this study examines if internal evaluation is used as a systematic instrument of self-governance in schools. Furthermore, we investigate if evaluation plans differ by federal states and if organizational variables at the school level influence program quality. Results show great differences between states and point to the key role principals play for program quality.
Keywords: internal evaluation, school improvement plans, data-based decision making, governance

Einleitung

Data-based decision making (vgl. Schildkamp/Ehren/Lai 2012) kann als ein Grundmotiv der Einführung eines neuen Steuerungsmodells im Bildungssystem betrachtet werden. Stärker als der Begriff der *Accountability* verweist dieses Konzept auf den systematischen Zusammenhang von Rechenschaftslegung und der Erweiterung von Gestaltungspielräumen für Schulen. Selbststeuerung von Schulen im Rahmen definierter Gestaltungspielräume vollzieht sich in der Perspektive dieser Steuerungslogik idealtypisch als *„retroactive planning"* (vgl. Scheerens/Demeuse 2005, S. 378). Planung umfasst dabei die Konzeption und Implementation von Maßnahmen auf der Basis belastbarer Daten zu Bedarfslagen, verfügbaren Ressourcen und Leistungen einer Schule sowie die Überprüfung der Wirksamkeit der implementierten Maßnahmen und ggf. ihre Fortschreibung. Grundlage des *data-based decision making* an Schulen können unterschiedliche Datentypen sein: Daten aus Vergleichsarbeiten, aus zentralen Abschlussprüfungen, aus Schulinspektionen, zum Unterricht aus Schüler-, Eltern- oder Lehrkräftebefragungen sowie regelmäßig erhobene Kenndaten wie Unterrichtsausfall oder Schulabsentismus. Im Rahmen der internen Evaluation haben Daten die Funktionen a) der Bestandsaufnahme von Bedarfslagen und Ressourcen, b) der Überprüfung der Leistungen einer Organisation und c) ihrer gezielten Entwicklung. Maßgeblich für interne Evaluation ist, dass die Auswertung vorhandener Daten systematisch im Hinblick auf strategische Ziele der Schule erfolgt und dass eigene Datenerhebungen passgenau auf die Situation der eigenen Schule zugeschnitten werden.

Zwei grundlegende Voraussetzungen müssen erfüllt sein, damit Daten aus interner Evaluation im Sinne der gezielten Planung und effektiven Selbststeuerung Verwendung finden können. Daten müssen einerseits methodischen Gütekriterien genügen. Dies gilt für die Entwicklung von erhebungsökonomisch realisierbaren Designs ebenso wie für die statistische Analyse der vorhandenen und erhobenen Daten. Die Generierung und Auswertung von Daten muss andererseits unmittelbar mit dem Planungsprozess verknüpft werden, weil Daten nicht wie in wissenschaftlichen Untersuchungen mit dem Ziel der Prüfung einer Hypothese oder der Exploration eines Sachverhalts erhoben werden, sondern praktischen Nutzungsimperativen unterworfen sind. Das gilt einerseits für die Konzeption und Implementation von Maßnahmen, andererseits für ihre Überprüfung und Adaption.

Schulinterne Evaluation als Instrument schulischer Selbststeuerung ist mit Schulprogrammarbeit unmittelbar verknüpft. Insofern ist es folgerichtig, dass die administrativen Vorgaben zur internen Evaluation in der Regel im Zusammenhang mit den Vorgaben zur Schulprogrammentwicklung in entsprechenden Ausführungsvorschriften definiert werden.

Wenngleich empirische Befunde darauf hinweisen, dass die schulinterne Evaluation im Vergleich zu Formen der externen Evaluation in Schulen eine große Akzeptanz erfährt und auch hinsichtlich ihrer Wirkungen positiver beurteilt wird als externe Evaluation (vgl. Hofman/Dijkstra/Hofman 2009, S. 50), ist sie an deutschen Schulen bislang wenig systematisch etabliert. Hierauf lassen insbesondere Untersuchungen zur Schulprogrammarbeit (vgl. Burkard 2004; Haenisch/Burkard 2002; Holtappels/Müller 2002; Kanders 2004; MBJS 2004; Mohr 2006; Schlömerkemper 2004) wie auch die Ergebnisse der Schulinspektion der Länder schließen (vgl. z.B. SenBWF 2011). Die meisten deutschsprachigen Untersuchungen zur internen Evaluation bzw. zur Schulprogrammarbeit sind in der frühen Phase der Einführung der Instrumente entstanden und liegen daher bereits ein Jahrzehnt zurück. Nachdem Ende der 1990er-Jahre ein Wandel der Bedeutung von Schulprogrammarbeit für die Einzelschule eingesetzt hat und Schulprogramme „von einem schulbezogenen Entwicklungsinstrument […] in ein systemisches Steuerungsinstrument transformiert" (Heinrich/Kussau 2010, S. 172) wurden, bedarf es aktuellerer Untersuchungen, die der Frage nachgehen, ob diese steuerungspolitische Akzentverschiebung auch in einer veränderten Handhabung der systematischen Schulprogrammarbeit, und insbesondere der internen Evaluation, durch die Schulen ihren Niederschlag gefunden hat.

Der vorliegende Beitrag präsentiert Ergebnisse einer Analyse von Evaluationsprogrammen Berliner und Brandenburger Gymnasien, wie sie in Schulprogrammen dokumentiert sind. Berlin und Brandenburg wurden ausgewählt, weil beide Bundesländer der internen Evaluation im Rahmen der Schulprogrammarbeit einen hohen Stellenwert einräumen, sich aber hinsichtlich zentraler Aspekte der Regulierung und Unterstützung der internen Evaluation unterscheiden (vgl. Kap. 3). Um mögliche systembedingte Unterschiede in der Qualität von Evaluationsprogrammen ausmachen zu können, wurde die Schulform konstant gehalten. Da Gymnasien in ihrer Schulorganisation traditionell eher durch eine lose Koppelung der Lehrkräfte gekennzeichnet sind, zugleich aber in besonderer Weise über Möglichkeiten der internen Spezialisierung verfügen,[1] ist die Untersuchung schulischer Selbststeuerungsprozesse und ihrer organisationalen Bedingungen hier von besonderem Interesse.

Der Darstellung des Forschungsstandes zur internen Evaluation an Schulen (1) folgen die Beschreibung der Funktionen und organisationalen Voraussetzungen des Einsatzes von interner Evaluation als Instrument schulischer Selbststeuerung (2) sowie die Beschreibung der unterschiedlichen Implementationsbedingungen in den bei-

1 So verfügen weiterführende Schulen über eine sehr viel differenziertere Leitungsfunktion bzw. über deutlich mehr Funktionsstellen als Grundschulen. Gemäß des Neuen Funktionsstellenmodells können die Leitungen weiterführender Schulen u.a. Funktionsstellen zur „Koordinierung schulfachlicher Aufgaben (Schulqualität/Schulprogramm/Schulentwicklung/Evaluation – Qualitätsbeauftragter/Qualitätsbeauftragte)" besetzen und haben mehr Spielraum, spezielle Funktionen (z.B. für Qualitätssicherung, Schulentwicklung, Schulprogrammarbeit), für die die Lehrkräfte Anrechnungsstunden erhalten, einzurichten (vgl. SenBWF 2010b).

den untersuchten Ländern (3). Nach Darstellung des Untersuchungsdesigns (4) werden die Ergebnisse präsentiert (5) sowie abschließend diskutiert (6).

1. Forschungsstand zur schulinternen Evaluation

In der Zusammenschau internationaler Befunde zur schulinternen Evaluation ergibt sich ein heterogenes Bild. Dies betrifft sowohl den inhaltlichen Fokus (z.B. Programme, Prozesse, Wirkungen) als auch die methodische Herangehensweise der Studien (ausführlicher vgl. Thiel/Thillmann 2012).

Einschlägige quantitative Untersuchungen, die teilweise auch die Wirkungen von interner Evaluation auf die Entwicklung von Schülerleistungen in den Blick nehmen, wurden in den Niederlanden und in Belgien durchgeführt. Signifikante Effekte der Nutzung von Instrumenten der internen Evaluation bzw. unterschiedlicher Nutzungsformen auf Schülerleistungen konnten die Autoren jedoch nicht nachweisen (vgl. Schildkamp/Visscher/Luyten 2009; Hofman/Dijkstra/Hofman 2009). Was die Frage nach der Nutzung von Daten aus interner Evaluation für schulische Selbststeuerungsprozesse betrifft, kommen die Studien ebenfalls zu wenig optimistischen Befunden: Demnach sind Schulen zwar in der Lage, Daten zu erheben und auszuwerten; die Überführung von Evaluationsergebnissen in Schulentwicklungsprozesse scheint ihnen jedoch vielfach große Schwierigkeiten zu bereiten. Im Umkehrschluss stehen schulische Entwicklungsvorhaben häufig nicht mit den Evaluationsergebnissen in Verbindung (vgl. Schildkamp/Visscher/Luyten 2009; Devos/Verhoeven 2003).

Studien zur schulinternen Evaluation aus dem deutschsprachigen Raum hingegen sind vielfach im Zusammenhang mit Pilotprojekten entstanden, haben überwiegend qualitativen Charakter und geben Einblicke in die subjektive Bewertung interner Evaluation durch Lehrkräfte. Demnach verstärkt die Einführung von interner Evaluation die Selbstwirksamkeit der Lehrkräfte, intensiviert die Kommunikation und verbessert das Schulklima (vgl. Altrichter/Posch 1999; Haenisch/Kindervater 1999; Müller 2002). Neben diesen positiven Einschätzungen verweisen Studien vereinzelt aber auch auf wahrgenommene negative Effekte, wie etwa eine erhöhte Arbeitsbelastung oder ein wahrgenommenes Bedrohungsgefühl (vgl. Boller 2009, S. 59ff.).

Interne Evaluation wurde darüber hinaus vor allem im Rahmen von Schulprogrammanalysen untersucht. Unter den zahlreichen Studien zur Schulprogrammarbeit finden sich neben Dokumentenanalysen (vgl. Holtappels/Müller 2002; MBJS 2004; Mohr 2006) auch quantitative und qualitative Befragungsstudien, die vielfach im Zusammenhang mit Evaluationsstudien entstanden sind (vgl. Burkard 2004; Haenisch/Burkard 2002; Kanders 2004; MBJS 2004; Schlömerkemper 2004). Das

Spektrum der Untersuchungsgegenstände reicht von Prozessen der Schulprogrammarbeit über die Inhalte der Programme bis hin zu eingeschätzten Wirkungen und Gelingensbedingungen. Im Folgenden werden zunächst Befunde zur inhaltlichen Ausgestaltung von Schulprogrammen bzw. von evaluationsrelevanten Aspekten der Schulprogrammarbeit dargestellt; anschließend werden die vorliegenden Befunde zu organisationalen Voraussetzungen von interner Evaluation berichtet.

Die Inhaltsanalysen von Schulprogrammen (vgl. Holtappels/Müller 2002; Mohr 2006; MBJS 2004[2]) kommen weitgehend einstimmig zu dem Ergebnis, dass die Mehrheit der Programme zwar Angaben zur Bestandsanalyse, zu Entwicklungszielen und -maßnahmen sowie zur Evaluationsplanung machen, die Einzelelemente untereinander jedoch größtenteils wenig konsistent sind. Den Ergebnissen von Mohr zufolge stehen „nur in knapp der Hälfte aller Programme die zunächst dargestellte Bestandsanalyse nebst entsprechender Schlussfolgerungen und die später gesetzten Schwerpunkte in erkennbarer logischer Konsequenz zueinander" (2006, S. 87). Die Bestandsaufnahmen haben außerdem überwiegend beschreibenden Charakter; eine Bewertung des festgestellten Ist-Zustandes findet eher selten statt (vgl. ebd.; MBJS 2004, S. 22). Auch was die Evaluationsplanung betrifft, verweisen die Ergebnisse aller drei Untersuchungen auf deutliche Defizite. Demnach beinhalten nur knapp die Hälfte der untersuchten Schulprogramme Angaben zur Evaluationsplanung (vgl. Holtappels/Müller 2002, S. 221; Mohr 2006, S. 89; MBJS 2004, S. 23).

Die schriftlichen Befragungen von Schulleitungen bestätigen die Befunde weitgehend, wenn auch in deutlich abgemilderter Form: Demnach gelingt es „nicht an allen Schulen […], stringent von der einfachen Schulbeschreibung zur Ableitung von pädagogischen Handlungsaufforderungen fortzuschreiten" (MBJS 2004, S. 82). Auch was die Ziel- und Umsetzungsorientierung der Programme sowie die Überprüfbarkeit der Arbeitsergebnisse betrifft, zeigen sich einige Schulleitungen skeptisch (vgl. ebd.).

Die Angaben von Schulaufsichtsbeamten und -beamtinnen im Rahmen der wissenschaftlichen Evaluation der Schulprogrammarbeit in Nordrhein-Westfalen, die im Anschluss an Dialoggespräche mit Schulen erhoben wurden, verweisen ebenfalls darauf, dass die Schulprogramme zum Erhebungszeitpunkt deutliche Defizite aufweisen (vgl. Burkard 2004). Demnach wird der Bereich der Bestandsaufnahme von 44 Prozent der untersuchten Schulen umfassend behandelt, Vereinbarungen für schulische Arbeitsfelder finden sich nur in rund 30 Prozent der Schulprogramme, Entwicklungsziele und ein Arbeitsplan in rund einem Fünftel der Programme. Die Evaluation der schulischen Arbeit wird nur in rund acht Prozent der Programme umfassend behandelt (vgl. ebd., S. 142f.).

2 Wiedergegeben sind hier nur die Befunde der Studie zu Schulprogrammen aus Ganztagsschulen. Untersucht wurden in dieser Studie auch Schulprogramme aus kleinen Grundschulen. Hinsichtlich der beschriebenen Kriterien stellen sich die Ergebnisse für diese Schulform ähnlich dar (vgl. ebd., S. 26 ff.; MBJS 2000).

Neben Untersuchungen zur Schulprogrammarbeit liefert auch die Schulinspektion wichtige Befunde zum Stand schulinterner Evaluation (vgl. Berkemeyer/Müller 2010, S. 209). Wenngleich etwa in Nordrhein-Westfalen der Qualitätsaspekt „Selbstevaluation" an über der Hälfte der Schulen mit „eher schwach als stark" bewertet wurde (vgl. ebd.), sprechen Befunde im Längsschnitt aber dafür, dass sich die Qualität der schulinternen Evaluation seit der Einführung des Verfahrens bis heute stark verbessert hat.[3] Haben etwa in Berlin im Schuljahr 2005/2006 weniger als zehn Prozent der beteiligten Schulen die Bestbewertung im Bereich „schulinterne Evaluation" erhalten (vgl. SenBWF 2007), waren es im Berichtszeitraum 2009/2010 über 60 Prozent (vgl. SenBWF 2011).

Auch was die Frage nach den organisationsspezifischen Voraussetzungen schulinterner Evaluation betrifft, können Studien zur Schulprogrammarbeit wichtige Hinweise geben. Durch das Zusammenspielen der Ergebnisse der Schulprogrammanalysen und Hintergrundinformationen der IGLU-Studie kommt Mohr zu den folgenden Ergebnissen: Die Leitungen von Schulen mit einem besonders gelungenen Schulprogramm[4] schätzen sowohl die Arbeitszufriedenheit der Lehrkräfte als auch den Kollegiumszusammenhalt und die Zusammenarbeit zwischen Schulleitung und Lehrkräften deutlich höher ein als die Schulleitungen der Vergleichsgruppe (vgl. Mohr 2006, S. 109). Den Ergebnissen von Holtappels und Müller (2002) zufolge zeigen sich je nach Qualität der Schulprogramme[5] signifikante Unterschiede hinsichtlich der in den Programmen beschriebenen Arbeitsorganisation: So haben Schulen mit einer hohen Programmqualität häufiger eine Steuergruppe und Arbeitsgruppen eingerichtet, und die Partizipation des Gesamtkollegiums ist stärker ausgeprägt (vgl. ebd., S. 229).

Haenisch und Burkhard (2002) ermitteln mit einer qualitativen Studie an 14 Schulen eine Vielzahl an Gelingensbedingungen für Schulprogrammarbeit. Als relevante Voraussetzungen auf der Ebene des Kollegiums können demnach gemeinsa-

3 Dies bezieht sich auf die Ergebnisse der Schulinspektion in Berlin. Zugrunde gelegt wurden hierbei die folgenden Kriterien: „Es gibt eine aktuelle Bestandsanalyse der Schule (Ausgangslage in Bezug auf Einzugsgebiet, soziale Struktur, unterrichtliche Prozesse etc.)", „Die Schule nutzt die Ergebnisse der Bestandsanalyse zur Ableitung von Stärken und Schwächen für die Schulprogrammarbeit", „Es liegt ein Evaluationsbericht vor", „In der Schule gibt es für die Evaluation Verantwortliche", „Es gibt ein vom Kollegium akzeptiertes Konzept zur schulinternen Evaluation", „Für die Evaluationsvorhaben sind Indikatoren bestimmt, die in der Schule bekannt sind", „Die schulinterne Evaluation erfolgt mit dem Ziel der Qualitätsverbesserung, insbesondere des Unterrichts" und „Die erzielten Evaluationsergebnisse werden in der Schule kommuniziert und zur Weiterentwicklung der Arbeit genutzt" (SenBWF 2009, S. 43).

4 Mohr bildet aus den Einschätzungen zur Schlüssigkeit, Konsistenz und Relevanz, zur didaktischen und theoretischen Fundierung und zum Konkretisierungsgrad der Schulprogramme einen Qualitätsindikator (vgl. Mohr 2006, S. 94).

5 Die Autoren legen die folgenden Qualitätsindikatoren zugrunde: „Klarheit und Verständlichkeit der textlichen Darstellung", „Differenzierungs- und Konkretisierungsgrad", „konzeptionelle Fundierung" und „Interne Konsistenz" (Holtappels/Müller 2002, S. 223).

me Ziele, wechselseitige Offenheit, wechselseitiges Vertrauen in die Professionalität, Innovationsbereitschaft, eine unterstützende Schulleitung, eine breite Beteiligung und tragfähige Arbeitsstrukturen gelten (vgl. ebd., S. 175ff.). Laut Auskunft der befragten Schulaufsichtsbeamten und -beamtinnen im Rahmen der oben erwähnten Evaluationsstudie zur Schulprogrammarbeit in Nordrhein-Westfalen (vgl. Burkard 2004) sind die folgenden Bedingungen besonders förderlich für die Schulprogrammarbeit: Innovationsbereitschaft im Kollegium, breite Beteiligung des Kollegiums, aktive Rolle der Schulleitung, eine klare Strukturierung des Arbeitsprozesses und die Einrichtung einer Steuergruppe (vgl. ebd., S. 151).

Zu ganz ähnlichen Ergebnissen kommen auch Untersuchungen, die sich direkt mit dem Gegenstand schulinterner Evaluation befassen. Förderliche organisationale Bedingungen der internen Evaluation sind demnach unter anderem die Innovationsbereitschaft und Zusammenarbeit im Kollegium, eine unterstützende Schulleitung, eine strukturierende und motivierende Steuergruppe, die zur Verfügung gestellten zeitlichen und materiellen Ressourcen sowie die Inanspruchnahme von Fortbildungen und Unterstützungsangeboten (vgl. Haenisch/Kindervater 1999, S. 79ff.; Müller 2002, S. 147ff.; Schildkamp/Visscher 2009, S. 157f.; Meuret/Morlaix 2003, S. 64f.).

2. Funktionen und Voraussetzungen interner Evaluation als Instrument schulischer Selbststeuerung

Wie eingangs bereits deutlich gemacht wurde, beziehen wir den Begriff der internen Evaluation – in Abgrenzung zum Begriff der Selbstevaluation des Unterrichts durch Lehrkräfte – auf die Ebene der Organisation (vgl. Thiel/Thillmann 2012). Interne Evaluation ist ein Instrument der schulischen Selbststeuerung und steht deshalb mit der Schulprogrammarbeit in unmittelbarem Zusammenhang. Die Verankerung der schulinternen Evaluation im Rahmen der Schulprogrammarbeit, wie sie in den Schulgesetzen bzw. Ausführungsverordnungen der meisten Bundesländer dokumentiert ist, trägt diesem Zusammenhang Rechnung (siehe Kap. 3). Schulinterne Evaluation kann demnach auch verstanden werden als eine Form schulischen Qualitätsmanagements (vgl. Berkemeyer/Müller 2010).

Interne Evaluation hat drei Funktionen (vgl. genauer Thiel/Thillmann 2012): Im Rahmen einer Bestandsaufnahme dient sie der *Beschreibung* einer Ausgangssituation im Hinblick auf spezifische Bedarfslagen und verfügbare Ressourcen einer Schule. Im Hinblick auf die Überprüfung der Leistungen einer Schule hat interne Evaluation die Funktion der *Beurteilung*. Für die Bestandsaufnahme und die Beurteilung können regelmäßig erhobene Kenndaten sowie Daten aus Vergleichsarbeiten und externer Evaluation Verwendung finden. Was die Funktion der Beschreibung betrifft, ist

von zentraler Bedeutung, dass Perspektive und Fokus und damit die Auswahl relevanter Daten gut begründet werden. Was die Beurteilung betrifft, ist entscheidend, ob Normen und Standards für eine Bewertung der Leistungen der Schule festgelegt wurden. Auf der Grundlage der entsprechenden Daten aus der Bestandsaufnahme und der Beurteilung der Leistungen erfolgen unter den Vorzeichen der *Optimierung* eine Zieldefinition und die Ableitung konkreter Entwicklungsmaßnahmen.[6]

Der Einsatz von interner Evaluation als Instrument schulischer Selbststeuerung setzt voraus, dass Daten nicht nur mit dem Ziel der Beschreibung und Beurteilung ausgewertet bzw. erhoben werden, sondern darüber hinaus Daten für eine kontinuierliche Optimierung bzw. Neukonzeption von Maßnahmen mit Bezug auf die schuleigenen Entwicklungsziele generiert werden, oder anders formuliert: dass alle drei Funktionen gleichermaßen zum Tragen kommen (vgl. Thiel/Thillmann 2012, S. 29ff.). Zu diesem Zweck definieren Schulen ein Evaluationsprogramm, das idealtypisch, neben strategischen Zielen, Standards und Indikatoren zur Überprüfung der Zielerreichung sowie Evaluationsverfahren und -instrumente festlegt. Dieses Programm ist in der Regel impliziter Bestandteil des Schulprogramms.

Es ist davon auszugehen, dass die Entwicklung von Evaluationsprogrammen im Sinne der beschriebenen Steuerungslogik von bestimmten organisationalen Voraussetzungen abhängt (vgl. zum Folgenden Thiel 2008):

- Eine entscheidungsfähige *strategische Spitze* (Schulleitung), die alle Planungsaktivitäten der Schule im Hinblick auf gemeinsame strategische Ziele koordiniert, stellt eine Voraussetzung dafür dar, dass Aktivitäten der internen Evaluation in engem Zusammenhang mit der strategischen Planung stehen.
- Weil die Generierung belastbarer Daten von spezifischen methodischen Kenntnissen abhängig ist, liegt es nahe, eine *funktionale Spezialisierung* dieser Aufgabenbearbeitung vorzusehen, d.h. Zuständigkeiten für interne Evaluation auszudifferenzieren und spezifische Qualifikationsangebote bereitzustellen.
- Eine datenbasierte Entwicklung effektiver Maßnahmen ist wiederum in hohem Maße von einem funktionsfähigen *mittleren Management* abhängig: einerseits von Steuergruppen, die die Entwicklung fachübergreifender Maßnahmen koordinieren, und andererseits von Fachkonferenzen, die für die koordinierte Entwicklung von fachbezogenen Maßnahmen verantwortlich sind.
- Weil es sich bei Schulen um spezifische professionelle Organisationen handelt, können organisationale Abstimmungsprozesse nur teilweise standardisiert werden. *Kollegiale Abstimmung* und das wechselseitige Zugeständnis von Autonomieansprüchen sind in professionellen Organisationen von zentraler Bedeutung. Beides kann, gerade wenn es um Fragen der Evaluation geht, in Konflikt ge-

6 Im Zyklus der Überarbeitung des Schulprogramms muss die Anpassung der Ziele und Maßnahmen – zusätzlich zu den Ergebnissen einer erneuten Bestandsaufnahme – dann natürlich maßgeblich auf den Ergebnissen der Evaluation der eigenen Maßnahmen basieren.

raten mit den vorausgehend beschriebenen Verfahren und Instrumenten der Standardisierung.
- Zudem besteht dort, wo strategische Planung individuelle professionelle Gestaltungsspielräume betrifft, ein besonderer Legitimationsbedarf von Entscheidungen. Mehr als in anderen Organisationen hängt auch die Legitimität von Entscheidungen, die Evaluation und Controlling betreffen, von *gemeinsamen professionellen Werten* und Überzeugungen ab.

Neben organisationsspezifischen Bedingungen sind Unterstützungssysteme in der Umwelt der Schule in Rechnung zu stellen. Weil Kompetenzen der internen Evaluation in der Ausbildung von Lehrkräften keine Rolle spielen, sind für den effektiven Einsatz interner Evaluation die Unterstützung der Schulen durch die staatliche Schulaufsicht sowie entsprechende Qualifizierungsangebote maßgeblich.

Im Mittelpunkt der hier präsentierten Analyse stehen Evaluationsprogramme, wie sie in den Schulprogrammen dokumentiert sind. Wir gehen von der Annahme aus, dass Programmen gerade in Prozessen der Qualitätsentwicklung und -sicherung eine entscheidende Bedeutung zukommt. Zielgerichtete Steuerung und eine aktive Adaption an dynamische Umweltbedingungen erfordern eine längerfristige strategische Planung, deren Dokumentation und kontinuierliche datenbasierte Fortschreibung.

Folgende Fragen leiten die Analyse der Dokumente:
1) Wie sind die Evaluationsprogramme hinsichtlich ihrer methodischen Fundierung und ihrer Kohärenz zu beurteilen?
2) Inwiefern wird interne Evaluation als Instrument strategischer Planung genutzt?
   - Erfolgt eine Bestandsaufnahme im Hinblick auf die für Schulentwicklung relevanten Bereiche umfassend?
   - Werden Leistungen der Schule auf der Basis konkreter Bewertungsmaßstäbe bzw. Standards beurteilt?
   - Erfolgt die im Evaluationsprogramm dokumentierte Definition von Entwicklungszielen auf der Basis von Daten der Bestandsaufnahme bzw. der Evaluation der schuleigenen Entwicklungsmaßnahmen?

Dass eine Programmanalyse eine Untersuchung der Evaluationspraxis nicht ersetzt, versteht sich von selbst: „the documents themselves will not unveil layers of lived culture and practice“ (Mintrop/MacLellan/Quintero 2001, S. 198). Gleichwohl ist eine Analyse von Programmen aufgrund ihrer oben beschriebenen Funktion für die Selbststeuerung von Organisationen von hohem Interesse. Programme haben für Organisationen die Funktion „to facilitate rational planning and efficiency“ (ebd.). Pläne oder Programme sind „management tools“ (ebd., S. 199). Sie stellen zwar keine hinreichende, aber eine notwendige Bedingung für die zielgerichtete, strategische Entwicklung einer Organisation dar (vgl. Mintzberg 1983). Dies gilt für operati-

ve Programme in Unternehmen und Evaluationsprogramme in Schulen in derselben Weise.

In der diesem Beitrag zugrunde liegenden Untersuchung werden Evaluationsprogramme im Rahmen von Schulprogrammen untersucht. Es ist davon auszugehen, dass die Qualität der Evaluationsprogramme mit anderen organisationsspezifischen Merkmalen, etwa mit der strategischen Spitze, der Spezialisierung und Qualifizierung von Evaluationsverantwortlichen oder der Arbeit von Steuergruppen, im Zusammenhang steht.

Außerdem gehen wir davon aus, dass Implementationsbedingungen wie administrative Rahmenvorgaben, systematische Qualifizierungs- und Unterstützungsangebote die Qualität von Evaluationsprogrammen maßgeblich beeinflussen, und vergleichen deshalb zwei Bundesländer, die sich im Hinblick auf Rahmenvorgaben und Unterstützung systematisch unterscheiden.

3. Interne Evaluation in Berlin und Brandenburg – Regulierung, Qualifizierung und Unterstützung

Mit Ausnahme des Saarlandes und Schleswig-Holsteins werden in den Schulgesetzen aller deutschen Bundesländer Vorgaben zur schulinternen Evaluation gemacht, wenngleich hinsichtlich der eingeräumten Gestaltungsfreiheit in Bezug auf die Planung und Durchführung der Evaluationsvorhaben und hinsichtlich des Umgangs mit den Ergebnissen zwischen den einzelnen Ländern teilweise große Unterschiede bestehen (vgl. Vereinigung der Bayerischen Wirtschaft e.V. 2010, S. 11ff.). Berlin und Brandenburg sind zwei Bundesländer, die das Instrument der internen Evaluation bereits zu einem frühen Zeitpunkt systematisch eingeführt und inzwischen schulgesetzlich verankert haben. In beiden Ländern wird interner Evaluation als schulischem Selbststeuerungsinstrument ein hoher Stellenwert eingeräumt. Was die Implementation dieses Instruments betrifft, unterscheiden sich die beiden Bundesländer allerdings in einigen zentralen Punkten. Ein systematischer Vergleich der Wirksamkeit unterschiedlicher Implementationsbedingungen liegt deshalb nahe. Im Folgenden werden die Vorgaben beider Länder zur Durchführung der schulinternen Evaluation, die im Rahmen von Schulprogrammarbeit stattfindet, ausführlicher dargestellt.

Während in den 1990er-Jahren sowohl in Berlin als auch in Brandenburg zunächst nur ausgewählte Schulen ein Schulprogramm entwickeln mussten, sind seit Einführung der neuen Schulgesetze (2002 in Brandenburg, 2004 in Berlin) alle Schulen zur Erstellung von Schulprogrammen verpflichtet. Hinsichtlich einiger zentraler formaler Vorgaben bestehen zwischen den beiden Ländern deut-

liche Unterschiede: In Berlin müssen die Schulprogramme, und damit auch die Evaluationskonzepte, aller Schulen von der Schulaufsichtsbehörde genehmigt werden. Brandenburger Schulen müssen ihre Schulprogramme ausschließlich in einem Gespräch mit der Schulaufsicht erörtern,[7] die „in geeigneter Weise Rückmeldungen zum Schulprogramm und den darin festgelegten pädagogischen Zielen" (MBJS 2009, S. 2) gibt.

Auch bezüglich der internen Evaluation liegen in Berlin und Brandenburg grundsätzlich ähnliche schulgesetzliche Vorgaben vor: Alle Schulen werden im jeweiligen Schulgesetz dazu verpflichtet, über interne Evaluationsvorhaben „den Erfolg ihrer pädagogischen Arbeit" (§ 8 Abs. 5 SchulG: SenBWF 2010a) bzw. das „Erreichen ihrer pädagogischen Ziele und die Umsetzung ihrer verabredeten Arbeitsschwerpunkte oder ihres Schulprogramms" (§ 7 Abs. 2 Satz 2 BbgSchulG: MBJS 2007) zu überprüfen. In Berlin soll eine derartige Überprüfung in allen Schulen spätestens nach drei Jahren geschehen,[8] während in Brandenburg nur für Ganztagsschulen seit 2011 (vgl. MBJS 2011) ein festgelegter Rhythmus vorgeschrieben wird.

Während in Brandenburg ausschließlich die oben erwähnten Erörterungsgespräche des Schulprogramms erfolgen, werden in Berliner Schulen explizite Rückmeldegespräche zur internen Evaluation durchgeführt (vgl. SenBWF 2008), in denen „konkrete Maßnahmen zur Qualitätsverbesserung der schulischen Arbeit" (ebd., S. 3) vereinbart werden sollen. Darüber hinaus müssen Berliner Schulen laut § 9 Abs. 2 Satz 5 SchulG (vgl. SenBWF 2010a) der Schulkonferenz sowie der Schulaufsicht einen schriftlichen Evaluationsbericht vorlegen. In Brandenburg wird indes nur von Ganztagsschulen gefordert, „die Ergebnisse der schulinternen Evaluation [zu] dokumentier[en]" (MBJS 2011, S. 5); eine Vorlage des Berichts bei der Schulaufsicht wird in diesem Zusammenhang nicht erwähnt.

Im Hinblick auf die inhaltlichen Anforderungen an den Prozess der schulinternen Evaluation liegen für Berlin und Brandenburg vergleichbare Vorgaben vor: Ausgehend von einer Bestandsaufnahme soll im Rahmen der Schulprogrammarbeit der oben beschriebene Qualitätszyklus angestoßen werden und nach Auswertung der Ergebnisse aus der internen Evaluation schließlich in eine Fortschreibung des Schulprogramms münden. Bei der Durchführung des Evaluationsprozesses stehen Berliner Schulen allerdings umfangreichere und detaillierte Materialien zur Verfügung: Neben einem Leitfaden zur Schulprogrammentwicklung (vgl. LISUM

7 Ausschließlich Schulen mit besonderer Prägung – beispielsweise sportbetonte Gesamtschulen, die besondere Aufnahmebedingungen haben und die Schülerinnen und Schüler in angeschlossenen Wohnheimen unterbringen können (vgl. MBJS 2003, S. 15) – sind zur Genehmigung ihrer Schulprogramme verpflichtet.

8 Bis 2011 waren Berliner Schulen sogar dazu verpflichtet, der Schulaufsicht innerhalb der festen Frist von drei Jahren ein überarbeitetes Schulprogramm vorzulegen.

2003[9]) können Berliner Schulen auf den „Leitfaden zur internen Evaluation für Schulen in Berlin“ (vgl. LISUM 2007) zurückgreifen. Abgesehen von den für Berlin und Brandenburg vergleichbaren Erläuterungen zur Schulprogrammarbeit bzw. zum Evaluationsprozess (Schulgesetze, Rundschreiben, Ausführungsvorschriften) gibt es für Brandenburger Schulen keine systematisch zusammengestellten Dokumente zur Unterstützung bei der Durchführung des gesamten Evaluationsprozesses. Ein zentraler Unterschied zwischen den beiden Ländern liegt außerdem in der Qualifizierung für interne Evaluation. In Berlin sind seit dem Schuljahr 2004/2005 an allen Schulen jeweils zwei Lehrkräfte zu Evaluationsberatern bzw. -beraterinnen ausgebildet worden (vgl. Ulber/Buchholz 2007), die im Idealfall, gemeinsam mit der Schulleitung und der Steuergruppe, den Prozess der schulinternen Evaluation „initiieren, steuern, durchführen und begleiten“ (LISUM 2007, S. 5) sollen.

Zusammenfassend kann festgehalten werden, dass sich die Implementationsbedingungen interner Evaluation in Berlin und Brandenburg in folgenden Punkten systematisch unterscheiden:

- Genehmigung der Schulprogramme und Evaluationskonzepte durch die Schulaufsicht;
- Überprüfung der datenbasierten Fortschreibung der Schulprogramme durch die Schulaufsicht;
- Qualifizierung von Evaluationsverantwortlichen;
- Materialien zur Schulprogrammarbeit und zur Durchführung des gesamten Evaluationsprozesses (Leitfäden).

4. Untersuchungsdesign und Methoden

Die vorliegende Untersuchung zur schulinternen Evaluation wurde als ein Teilvorhaben im Rahmen des vom Bundesministerium für Bildung und Forschung geförderten Projektes „Schulen als Steuerungsakteure im Bildungssystem (StABil)“ durchgeführt. Aufgrund der beschriebenen Systemunterschiede zwischen den Bundesländern Berlin und Brandenburg wurden in der hier berichteten Studie die Evaluationsprogramme von Gymnasien aus beiden Bundesländern vergleichend analysiert. Neben der Beurteilung der methodischen Fundierung und der Kohärenz der Programme sollte außerdem der Frage nachgegangen werden, inwiefern interne Evaluation von Schulen als Instrument strategischer Planung genutzt wird (vgl. Kap. 2). Durch Verknüpfung der Analyseergebnisse mit den Daten einer Onlinebefragung

9 Dieser vom Berliner Landesinstitut für Schule und Medien herausgegebene Leitfaden basiert auf den Ergebnissen eines Pilotprojekts zur „Schulprogrammentwicklung und Evaluation“ an Berliner Schulen. Auf den Internetseiten des Bildungsservers Berlin wird dieser Leitfaden nur im Zusammenhang mit den Erläuterungen zur Schulprogrammarbeit für Berliner Schulen erwähnt.

von Schulleitungen wurde anschließend untersucht, inwiefern die Qualität der Evaluationsprogramme von bestimmten organisationalen Voraussetzungen abhängt.

4.1 Beschreibung der Stichprobe

Die vorliegende Analyse beschränkt sich auf Programme von Gymnasien aus den beiden Bundesländern Berlin und Brandenburg. Um die Qualität der Evaluationsprogramme mit verschiedenen Kontext- und Organisationsmerkmalen in Zusammenhang setzen zu können, wurden nur Programme von Schulen analysiert, deren Leitungen sich an der zeitlich vorangegangenen Onlinebefragung beteiligt hatten. Aufgrund der leichteren Zugänglichkeit erfolgte der Zugang zu den Programmen zunächst über das Internet: Knapp zwei Drittel (60 Prozent) der an der Onlinebefragung beteiligten Gymnasien haben ihr Schulprogramm online gestellt. Die so gewonnene Stichprobe wurde im Hinblick auf zentrale relevante Merkmale (regionaler Standort, Schulgröße, Durchschnittsalter der Lehrkräfte, soziale Zusammensetzung der Schülerschaft) sowohl für Berlin als auch für Brandenburg mit der Population der Gymnasien verglichen. Da die Stichprobe in zentralen Merkmalen der Population entspricht, konnte auf eine Ausweitung der Stichprobe verzichtet werden, sodass sich die Zahl der analysierten Programme auf insgesamt N = 31 (18 Prozent aller Gymnasien der Population) beläuft. Darunter befinden sich 20 Programme von Berliner Gymnasien und 11 Programme von Brandenburger Gymnasien.

4.2 Analyse der Evaluationsprogramme

4.2.1 Beschreibung des Analyserasters

Auf der Grundlage der theoretischen Vorüberlegungen (vgl. Kapitel 2.1) und unter Rückgriff auf bestehende Arbeiten zur Programmanalyse (vgl. Mohr 2006; Thiel u.a. 2010) sowie auf die Berliner „Ausführungsvorschriften zur Erstellung der Schulprogramme und zur internen Evaluation“ (SenBWF 2008) wurde ein Analyseraster für ein Rating der Evaluationsprogramme entwickelt, das sich in seinen Kernelementen wie folgt beschreiben lässt:

Zur Beurteilung der Qualität der internen Evaluation werden die Programme zunächst im Hinblick auf zwei übergeordnete Qualitätskriterien analysiert: die methodische Fundierung sowie die Kohärenz des Evaluationsprogramms. Die *methodische Fundierung* bemisst sich daran, inwieweit für die benannten Evaluationsvorhaben Indikatoren zur Überprüfung der Zielerreichung festgelegt werden und der geplante Einsatz von Evaluationsverfahren und -instrumenten beschrieben wird bzw. wie konkret die Indikatoren bzw. die Evaluationsverfahren und -instrumente be-

schrieben werden.[10] Auf der Basis dieser Kriterien wird ein Index gebildet, der die Ausprägungen 0 (überhaupt nicht fundiert) bis 8 (ausgesprochen fundiert) annehmen kann. Das Kriterium der *Kohärenz* ergibt sich als Index aus dem Ausmaß der Zusammenhänge zwischen der Bestandsaufnahme, den Entwicklungszielen, den Entwicklungsmaßnahmen und den Evaluationsverfahren.[11] Dieser Index kann die Ausprägungen 0 (gar nicht kohärent) bis 6 (durchweg kohärent) annehmen.[12] Die beiden Indices der „methodischen Fundierung" und der „Kohärenz" werden anschließend aufsummiert und als Qualitätsindikator in den Datensatz der Onlinebefragung integriert. Der aufsummierte Qualitätsindikator kann damit die Ausprägungen 0 (keine Qualität) bis 14 (hervorragende Qualität) annehmen.

Darüber hinaus werden die Evaluationsprogramme entlang der beschriebenen Funktionen von interner Evaluation („Beschreibung", „Beurteilung" und „Optimierung") hinsichtlich ihrer Eignung als Instrumente der strategischen Planung untersucht. Der Aspekt der *Beschreibung* bezieht sich auf die Frage, ob bzw. wie umfassend Schulen eine Bestandsaufnahme ihrer schulischen Arbeit vornehmen. Inwieweit die Bestandsaufnahme als auf relevante Bereiche schulischer Entwicklung fokussiert bewertet wird, bemisst sich daran, ob Aspekte aus relevanten Bereichen (z.B. schulische Daten zu Lernergebnissen, Daten zum Unterrichtsausfall, Ist-Zustand der Unterrichtsentwicklung, Ist-Zustand der Organisationsentwicklung) erfasst werden. Der Aspekt der *Beurteilung* bezieht sich auf die Frage, inwiefern die Ergebnisse der Bestandsaufnahme auf der Basis ausgewiesener Normen und Standards bewertet werden[13] und ob konkrete Bewertungsmaßstäbe benannt werden. Hinsichtlich der Frage der *Optimierung* wird untersucht, ob konkrete Entwicklungsziele aus der Bestandsaufnahme (oder den Ergebnissen einer bereits erfolgten Evaluation) abgeleitet werden. Hier wird, wie oben bereits im Zusammenhang mit der Programmkohärenz beschrieben, detailliert dargestellt, inwieweit Entwicklungsziele mit den Evaluationsergebnissen im Zusammenhang stehen.

10 Beide Aspekte können jeweils drei Ausprägungen annehmen (0 = keine Indikatoren bzw. Evaluationsverfahren/-instrumente; 1 = nur teilweise; 2 = umfassend) und werden zusätzlich hinsichtlich ihres Konkretisierungsgrades bewertet (0 = keine Indikatoren bzw. Evaluationsverfahren/-instrumente; 1 = insgesamt allgemein; 2 = insgesamt konkret).

11 0 = Entwicklungsziele überhaupt nicht im Zusammenhang mit Bestandsaufnahme bzw. früheren Evaluationsergebnissen – Maßnahmen überhaupt nicht im Zusammenhang mit Entwicklungszielen – Evaluationsverfahren überhaupt nicht im Zusammenhang mit Entwicklungszielen; 1 = teilweise im Zusammenhang; 2 = voll und ganz im Zusammenhang; 9 = keine Entwicklungsziele – keine Maßnahmen – keine Evaluationsverfahren.

12 Ausprägungen, die mit 9 geratet wurden, werden bei der Berechnung des Index als fehlende Werte behandelt.

13 0 = nein, gar nicht; 1 = interpretative Bewertung des Ist-Zustandes; 2 = teilweise Gegenüberstellung von Ist-Soll-Zustand; 3 = konsequente Gegenüberstellung von Ist-Soll-Zustand.

4.2.2 Erläuterung des Ratingverfahrens

Alle verfügbaren Evaluationsprogramme der in der Onlinebefragung beteiligten Gymnasien wurden von einer Verfasserin des Artikels anhand des beschriebenen Analyserasters analysiert. Um die Objektivität des Analyserasters zu überprüfen, wurde ein Drittel (N = 10) dieser Evaluationsprogramme von einer zweiten Raterin analysiert, nachdem zunächst beide Beurteilerinnen eine ausführliche Schulung zu Inhalt und Handhabung des Analyserasters erhalten hatten.[14] Nachdem beide Raterinnen zunächst zwei der zufällig ausgewählten zehn Evaluationsprogramme analysiert hatten, lag die durchschnittliche prozentuale Übereinstimmung bei 68 Prozent. In einem anschließenden Gespräch wurden alle aufgetretenen Unklarheiten erörtert, und die während des Analyseprozesses aufgetretenen offenen Fragen konnten anhand konkreter Beispiele geklärt werden. Für die folgenden acht Programme konnte schließlich eine zufriedenstellende Übereinstimmungsrate von durchschnittlich 88 Prozent erreicht werden. Jegliche Abweichungen beliefen sich auf nur +/– 1 Punkt auf der jeweiligen Skala; aufgrund der großen Heterogenität der untersuchten Evaluationsprogramme kann bei diesem Ergebnis davon ausgegangen werden, dass eine objektive und reliable Auswertung gewährleistet werden konnte.

4.3 Onlinebefragung von Schulleitungen

4.3.1 Datenerhebung

Die Onlinebefragung der Leitungen aller allgemein bildenden Schulen in Berlin und Brandenburg wurde von November 2011 bis Februar 2012 durchgeführt. Ziel der Befragung war unter anderem die umfassende, theoriegeleitete Erfassung schulischer Organisationsmerkmale. Schulleitungen wurden dabei als Experten befragt, die über einen privilegierten Zugang zu Informationen über die Strukturen und Abläufe der Schule verfügen (vgl. Meuser/Nagel 2009, S. 470).

4.3.2 Auswahl relevanter Untersuchungsmerkmale

Auf der Basis der theoretischen Vorüberlegungen zu organisationalen Voraussetzungen schulischer Selbststeuerung durch interne Evaluation (Kap. 2) werden Merkmale zur Rolle der Schulleitung als *strategische Spitze*, zur Steuergruppe als *mittleres Management* sowie zur *funktionalen Spezialisierung* in die Untersuchung auf-

14 Die zweite Raterin ist ebenfalls eine Verfasserin des vorliegenden Artikels. Aufgrund der vergleichbaren Vorkenntnisse der beiden Beurteilerinnen und aufgrund der zu Beginn des Analyseprozesses durchgeführten gemeinsamen Schulung kann von in etwa „gleichem Wissensstand" (Wirtz/Casper 2002, S. 15) ausgegangen werden, welcher als Vorbedingung für eine reliable Beurteilung betrachtet werden kann (vgl. ebd., S. 15, 24).

genommen. Die Aspekte der *kollegialen Abstimmung* und *gemeinsamen Werte* im Kollegium wurden in der Schulleiterbefragung nicht erhoben und können daher nicht berücksichtigt werden.

Die Merkmale wurden wie folgt operationalisiert (vgl. Tabelle 1):

- Die Frage, ob die Schulen über eine entscheidungsfähige *strategische Spitze* (Schulleitung) verfügen, wurde zum einen über das Selbstverständnis der Schulleitungen operationalisiert. In Abgrenzung zu einem eher traditionellen Schulleiterverständnis als „primus inter pares" wurde erfragt, inwieweit sich die Schulleitungen selbst als Vorgesetzte der Lehrkräfte betrachten. Zum anderen wurde erhoben, ob die Schulleitungen eine spezielle Aus- oder Fortbildung für die Schulleitungstätigkeit absolviert haben.
- Die Einrichtung eines funktionierenden *mittleren Managements* in Form einer Steuergruppe wurde zum einen über die Existenz einer Steuergruppe und zum anderen über deren Funktion (koordinierend vs. inhaltlich arbeitend) operationalisiert.
- Als relevantes Merkmal der *funktionalen Spezialisierung* wurde erhoben, ob die Schulen eine formale Zuständigkeit für Evaluation eingerichtet haben.

Tab. 1: Operationalisierung schulischer Organisationsmerkmale

Merkmal	*Operationalisierung*
Schulleiterselbstverständnis als Vorgesetzte/r	Ich verstehe mich vor allem als Vorgesetzte/r. [a]
Spezielle Schulleiterqualifikation	Haben Sie eine Aus- oder Fortbildung für die Schulleitungstätigkeit absolviert bzw. entsprechende Qualifikationen erworben? [b]
Etablierung einer Steuergruppe	Welche der folgenden festen Zuständigkeiten wurden an Ihrer Schule etabliert? – Steuergruppe [b]
Koordinierungsfunktion der Steuergruppe	Wie würden Sie die Funktion der Steuergruppe Ihrer Schule eher charakterisieren? [c]
Etablierung einer/eines Evaluationsverantwortlichen	Welche der folgenden festen Zuständigkeiten wurden an Ihrer Schule etabliert? – Evaluationsverantwortliche/r [b]

Anmerkungen.

[a] 1 = „stimme überhaupt nicht zu" bis 6 = „stimme vollständig zu".

[b] 0 = „nein"; 1 = „ja".

[c] 0 = „Die Steuergruppe führt die schulischen Entwicklungsvorhaben im Wesentlichen selbst durch";
1 = „Die Steuergruppe hat bei uns vorwiegend eine Koordinierungsfunktion".

Quelle: eigene Darstellung

4.4 Auswertung der Daten

Die Datenanalyse erfolgte in zwei Schritten. Zunächst wurden die vorliegenden Evaluationsprogramme anhand des beschriebenen Analyserasters ausgewertet. Die Kodierung und die deskriptiven Analysen erfolgten unter Verwendung von SPSS 21. In einem zweiten Schritt wurden die Kodierungen in den Datensatz der Onlinebefragung integriert und hinsichtlich möglicher Zusammenhänge mit schulischen Kontext- und Organisationsmerkmalen analysiert. Zur statistischen Überprüfung möglicher Bundeslandunterschiede werden t-Tests für unabhängige Stichproben und zur Überprüfung der vom Bundesland bereinigten Zusammenhänge partielle Korrelationen berechnet. Aufgrund der vorliegenden kleinen Stichprobe werden zur Bewertung der praktischen Relevanz (vgl. Lind 2012) darüber hinaus Effektgrößen (d und r)[15] angegeben.

5. Ergebnisse

In einem ersten Schritt werden die Ergebnisse der Programmanalyse vergleichend für die beiden Untersuchungsländer dargestellt. Dabei wird zum einen auf die übergeordneten Qualitätsmerkmale der methodischen Fundierung und der Kohärenz der Programme eingegangen. Zum anderen wird entlang der beschriebenen Funktionen von Evaluation („Beschreibung", „Beurteilung" und „Optimierung") dargestellt, inwiefern interne Evaluation von Schulen als Instrument strategischer Planung genutzt wird. In einem zweiten Schritt wird dargestellt, inwiefern die Qualität der Evaluationsprogramme von bestimmten organisationalen Voraussetzungen und externen Unterstützungssystemen abhängt.

5.1 Ergebnisse der Programmanalyse

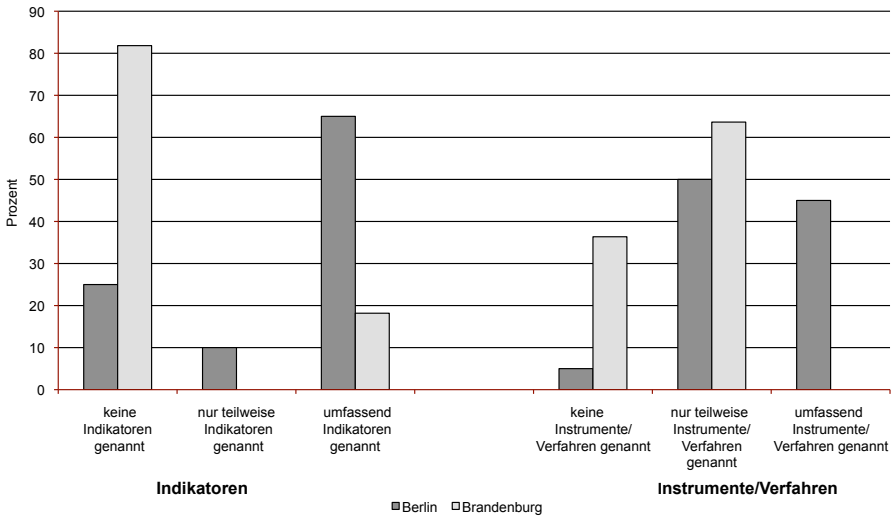
5.1.1 Methodische Fundierung der Evaluationsprogramme

Die Ergebnisse der Programmanalyse verweisen auf große Unterschiede zwischen den Bundesländern hinsichtlich der methodischen Fundierung der Evaluationsprogramme. Während in 65 Prozent (N = 13) der Evaluationsprogramme von Berliner Gymnasien für die benannten Evaluationsvorhaben umfassend Indikatoren für die Zielerreichung benannt sind, weist der Großteil der untersuchten Programme von Brandenburger Gymnasien überhaupt keine Indikatoren auf (82 Prozent, N = 9). Doch auch die in den Berliner Programmen benannten Indikatoren sind größten-

15 Cohen (1988) bezeichnet einen Effekt von d = 0.20 bzw. r = 0.10 als klein, einen Effekt von d = 0.50 bzw. r = 0.30 als mittleren und einen Effekt von d = 0.80 bzw. r = 0.50 als großen Effekt.

teils als wenig konkret zu beschreiben (67 Prozent, N = 10), etwa in dem Sinne, dass Standards zur Überprüfung der Zielerreichung benannt wären. Was die Beschreibung des geplanten Einsatzes von Evaluationsverfahren und -instrumenten betrifft, so werden in fast allen der untersuchten Programme von Berliner Gymnasien teilweise oder umfassend Verfahren und Instrumente benannt (95 Prozent, N = 19); die Beschreibung der Verfahren und Instrumente wird jedoch überwiegend eher allgemein gehalten (58 Prozent, N = 11). So wird z.B. häufig nur die Art der Erhebungsinstrumente benannt (z.B. Fragebögen), ohne näher zu beschreiben, wer konkret zu welchen Aspekten befragt werden soll. Deutlich seltener und insgesamt wenig konkret werden hingegen in den Programmen der Brandenburger Gymnasien Evaluationsinstrumente und -verfahren benannt (64 Prozent, N = 7) (vgl. Abb. 1).

Abb. 1: Benennung von Indikatoren und Evaluationsverfahren



Quelle: eigene Darstellung

Insgesamt – über die verschiedenen Kriterien hinweg – weisen die Evaluationsprogramme der Berliner Gymnasien eine deutlich bessere methodische Fundierung (M = 5,2; SD = 2,3) auf als die Programme der Brandenburger Gymnasien (M = 1,9; SD = 1,1) (d = 1,73, p = ,000). Der Mittelwert der Brandenburger Programme liegt deutlich unterhalb des semantischen Mittels des Index[16] (vgl. Abb. 2).

16 Das semantische Mittel des Index „Methodische Fundierung“ liegt bei 4.

5.1.2 Kohärenz der Evaluationsprogramme

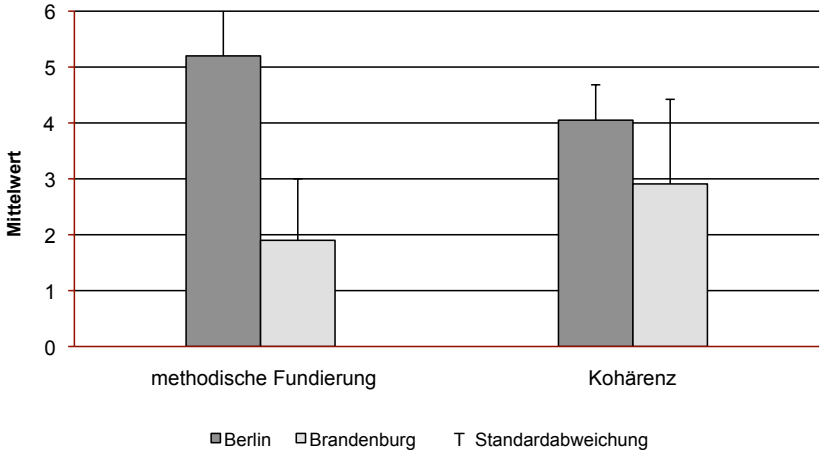
Was die *Kohärenz der Evaluationsprogramme* betrifft, zeigen alle untersuchten Programme von Berliner Gymnasien, dass die Entwicklungsziele „voll und ganz" (45 Prozent, N = 9) oder „teilweise" (55 Prozent, N = 11) im Zusammenhang mit der Bestandsaufnahme oder den Ergebnissen einer früheren Evaluation stehen. Auch die Analyse der Brandenburger Programme zeigt nur in wenigen Fällen überhaupt keine Zusammenhänge (18 Prozent, N = 2); in der überwiegenden Mehrheit stehen auch hier die formulierten Ziele mit der Bestandsaufnahme oder einer früheren Evaluation „teilweise" (55 Prozent, N = 6) oder „voll und ganz" (27 Prozent, N = 3) im Zusammenhang.

Auch mit den beschriebenen Entwicklungsmaßnahmen wird in den Programmen aus Schulen beider Bundesländer überwiegend auf die genannten Entwicklungsziele Bezug genommen: In 60 Prozent (N = 12) der Evaluationsprogramme der Berliner Gymnasien stehen die Entwicklungsziele „voll und ganz" im Zusammenhang, in 40 Prozent (N = 8) „teilweise". In den Programmen der Brandenburger Gymnasien hingegen stehen Maßnahmen und Ziele meist nur „teilweise" im Zusammenhang (67 Prozent, N = 6), in drei Programmen „voll und ganz" (33 Prozent).

Größere Schwierigkeiten scheinen die Schulen damit zu haben, auf ihre schuleigenen Ziele zugeschnittene Evaluationsverfahren zu entwickeln: In den Programmen beider Bundesländer beziehen sich die genannten Evaluationsverfahren mehrheitlich „überhaupt nicht" (Berlin: 16 Prozent; Brandenburg: 22 Prozent) oder nur „teilweise" (Berlin: 63 Prozent; Brandenburg: 67 Prozent) auf die schulischen Entwicklungsziele.

Insgesamt weisen die Berliner Evaluationsprogramme eine mittlere Kohärenz von 4,1 auf (SD = 1,1); der Kohärenzindex für die Brandenburger Programme hingegen liegt bei 2,9 (SD = 1,5). Die Programme beider Bundesländer unterscheiden sich auch hinsichtlich dieses übergeordneten Merkmals signifikant voneinander (d = ,94, p = .022) (vgl. Abb. 2).

Abb. 2: Übergreifende Qualitätsmerkmale der Evaluationsprogramme



Quelle: eigene Darstellung

Fasst man beide Aspekte der Kohärenz und der methodischen Fundierung zu einem übergreifenden Qualitätsindikator zusammen, so zeigt sich das folgende Bild: Während die untersuchten Programme von Berliner Gymnasien im Durchschnitt einen Mittelwert des Qualitätsindex von M = 9,2 (SD = 3,1) aufweisen und damit deutlich über dem semantischen Mittel des Index[17] liegen, nimmt die Qualität der Programme der Brandenburger Gymnasien einen durchschnittlichen Wert von M = 4,8 (SD = 2,0) an. Beide Länder unterscheiden sich damit hinsichtlich der Qualität der Evaluationsprogramme signifikant voneinander (d = 1.63, p = .000).

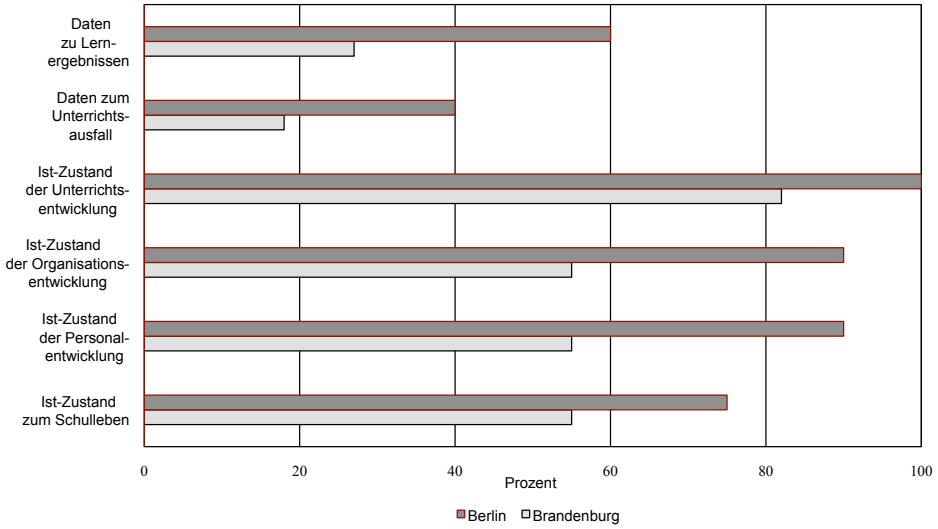
5.1.3 Evaluationsprogramme als Instrumente der strategischen Planung

Auch im Hinblick auf die Eignung der Evaluationsprogramme als Instrumente der strategischen Planung zeigen sich signifikante Unterschiede zwischen den Bundesländern. Zwar ist eine *Beschreibung* des schulischen Entwicklungsstandes in allen Programmen vorhanden. Während aber in den Berliner Evaluationsprogrammen die durchschnittliche Anzahl der berücksichtigten Bereiche bei 4,6 (SD = 1,40) liegt, wird der Entwicklungsstand in den Programmen der Brandenburger Gymnasien durchschnittlich nur in knapp drei Bereichen (SD = 1,04) abgebildet (d = 1.32, p = .002). Am häufigsten erfolgt in den Programmen von Schulen aus beiden Bundesländern eine Bestandsaufnahme zum Ist-Zustand der Unterrichtsentwicklung (z.B. Formen der Unterrichtsorganisation und des Medieneinsatzes), der Personalentwicklung (z.B. Fortbildungen, Studientage) und der Orga-

17 Der semantische Mittelwert des Qualitätsindikators liegt bei 7.

nisationsentwicklung (z.B. Teambildung im Kollegium). Daten zu Lernergebnissen oder zum Unterrichtsausfall werden hingegen seltener erfasst (vgl. Abb. 3).

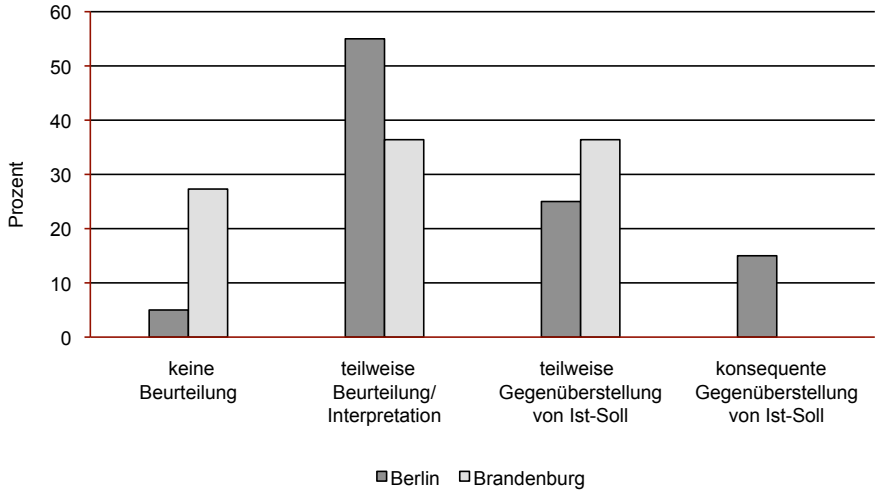
Abb. 3: Berücksichtigte Bereiche der Bestandsaufnahme



Quelle: eigene Darstellung

Während in 95 Prozent (N = 19) aller Berliner Programme zusätzlich eine *Beurteilung* der Bestandsaufnahme vorgenommen wird, ist dies nur in 73 Prozent (N = 8) aller Programme der Brandenburger Gymnasien der Fall. Meist handelt es sich dabei jedoch nur um eine interpretative Bewertung des Ist-Zustandes, ohne dass diesem konsequent ein Soll-Zustand gegenübergestellt wird (vgl. Abb. 4). Konkrete Bewertungsmaßstäbe (z.B. ein Vergleich mit anderen Schulen oder früheren Jahrgängen) werden in 45 Prozent der Berliner Programme (N = 9) und 27 Prozent der Brandenburger Programme (N = 3) benannt. Meist wird hierbei auf soziale Vergleichsmaßstäbe verwiesen (z.B. Vergleich zu anderen Schulen der Stadt oder des Bezirks); vereinzelt werden aber auch innerschulische Entwicklungsmaßstäbe angelegt (z.B. Vergleich zu früheren Jahrgängen).

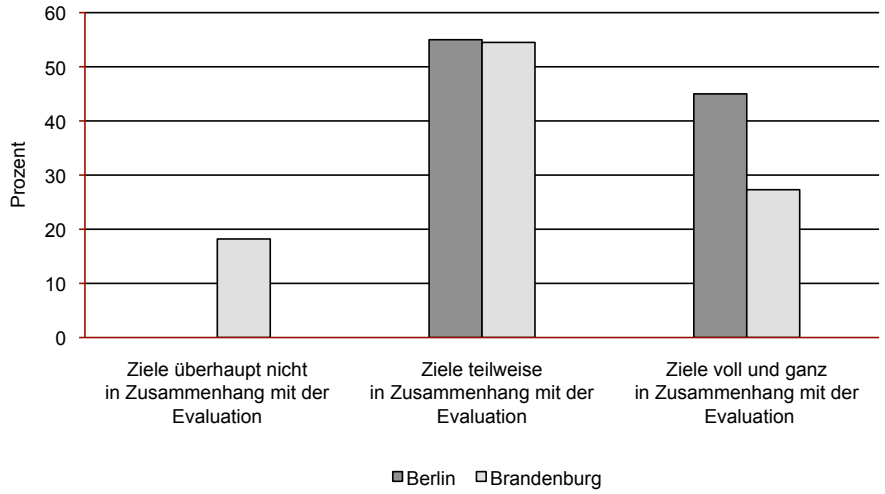
Abb. 4: Beurteilung der Bestandsaufnahme



Quelle: eigene Darstellung

Hinsichtlich der Frage der *Optimierung* zeigen die Ergebnisse zunächst, dass sich in allen analysierten Programmen Entwicklungsziele und – mit Ausnahme von zwei Programmen von Brandenburger Schulen – auch konkrete Entwicklungsmaßnahmen finden. Wie oben im Zusammenhang mit der Programmkohärenz bereits beschrieben, scheinen die in den meisten Evaluationsprogrammen dokumentierten Entwicklungsziele, zumindest teilweise, aus der Bestandsaufnahme bzw. den Evaluationsergebnissen abgeleitet worden zu sein. „Voll und ganz" im Zusammenhang mit den Evaluationsergebnissen stehen die Entwicklungsziele aber nur in 45 Prozent der Berliner Programme (N = 9) und in 27 Prozent der Brandenburger Programme (N = 3) (vgl. Abb. 5).

Abb. 5: Ableitung der Entwicklungsziele aus der Evaluation



Quelle: eigene Darstellung

Zusammenfassend kann also festgehalten werden, dass es den wenigsten Schulen gelingt, ein Evaluationsprogramm zu entwickeln, das als Instrument der strategischen Planung gelten kann. Eine breite Bestandsaufnahme, die unterschiedliche für Schulentwicklung relevante Bereiche in den Blick nimmt, eine konsequente Beurteilung der Bestandsaufnahme in Hinblick auf einen definierten Soll-Zustand und eine entsprechende Ableitung von Zielen erfolgen nur in drei der untersuchten Programme. Dabei handelt es sich bei allen drei Fällen um Evaluationsprogramme von Berliner Schulen.

5.2 Der Stellenwert organisationaler Bedingungen

Was die Zusammenhänge der Programmqualität mit schulorganisationalen Merkmalen anbelangt, zeigen sich – unter Kontrolle des Bundeslandes – deutliche Effekte für die Schulleiterqualifizierung ($r = ,408$; $p = ,028$, $N = 30$). Das heißt, die Evaluationsprogramme der Schulen, deren Leitungen eine Aus- oder Fortbildung für die Schulleitungstätigkeit absolviert bzw. entsprechende Qualifikationen erworben haben, weisen tendenziell eine höhere Programmqualität auf. Ebenso zeigen sich kleine positive Zusammenhänge mit dem Selbstverständnis der Schulleitungen als Vorgesetzte ($r = ,217$, n.s., $N = 26$).

Auch das Merkmal der funktionalen Spezialisierung korreliert schwach mit der Qualität der Evaluationsprogramme: Schulen, die Evaluationsverantwortliche eta-

bliert haben, weisen tendenziell etwas bessere Evaluationsprogramme auf (r = ,204, n.s., N = 27).

Hinsichtlich der Bedeutsamkeit eines funktionierenden Mittleren Managements in Form einer Steuergruppe für die Qualität der Evaluationsprogramme zeichnen die Ergebnisse das folgende Bild: Während die Einrichtung einer Steuergruppe alleine die Qualität der Programme nicht zu begünstigen scheint, zeigen sich kleine positive Zusammenhänge, wenn die Steuergruppe eine Koordinierungsfunktion im Kollegium einnimmt (r = ,127, n.s., N = 22).

6. Diskussion

Wenngleich es sich bei den analysierten Evaluationsprogrammen um eine selbstselektierte Stichprobe[18] handelt, weist sie im Hinblick auf zentrale Merkmale eine hohe Ähnlichkeit mit der Population auf. Es kann also davon ausgegangen werden, dass die Gymnasien beider Bundesländer gut repräsentiert sind und die Ergebnisse der Studie damit Rückschlüsse auf die Gesamtheit der Evaluationsprogramme an Berliner und Brandenburger Gymnasien zulassen.

Zwar ist die vorliegende Untersuchung sowohl im Hinblick auf die Stichprobe (Bundesländer, Schulformen) als auch auf die zugrunde gelegten Analysekategorien nur begrenzt mit älteren Studien zur Schulprogrammarbeit (vgl. Holtappels/Müller 2002; Mohr 2006) vergleichbar, doch weisen die untersuchten Evaluationsprogramme auch rund zehn Jahre nach der Einführung des Instrumentes noch ähnliche Defizite auf. Dies betrifft insbesondere die methodische Fundierung, aber auch die Kohärenz der analysierten Programme. Zugleich zeigen die Ergebnisse aber große Qualitätsunterschiede zwischen den Programmen beider Bundesländer, die mit großer Wahrscheinlichkeit den unterschiedlichen Rahmenvorgaben bzw. Implementationsbedingungen in beiden Ländern geschuldet sind. Dies deutet darauf hin, dass eine klare Steuerungsstrategie, die auf externe Steuerungsimpulse in Form von konkreten Vorgaben, gezielter Unterstützung und regelmäßiger Rechenschaftspflicht setzt, die Qualität der Evaluationsprogramme maßgeblich beeinflussen kann.

Neben der Bedeutsamkeit unterschiedlicher Implementationsbedingungen verweisen die Ergebnisse – über beide Bundesländer hinweg – auch auf den Stellenwert der organisationalen Voraussetzungen der Schulen für die Qualität der Evaluationsprogramme. Eine zentrale Rolle kommt dabei der Schulleitung zu. Zwar ist angesichts der allgemeinen Operationalisierung des Merkmals der Schulleiterqualifizierung nicht eindeutig zu klären, ob tatsachlich eine spezifische Leitungs-

18 Dies gilt sowohl einerseits für die Teilnahme an der freiwilligen Onlinebefragung als auch andererseits für die Veröffentlichung der Schulprogramme im Internet.

qualifizierung der Schulleitungen für die Qualität der Evaluationsprogramme ausschlaggebend ist oder ob sich hinter dem Befund nicht vielmehr eine allgemeine Weiterbildungs- und damit auch Innovationsbereitschaft von Schulleitungen verbirgt. Deutlich wird aber, dass Schulleitungen – sei es aufgrund ihrer spezifischen Qualifikationen oder ihrer innovationsorientierten Haltung – maßgeblichen Einfluss auf die Qualität der Evaluationsprogramme nehmen können. Genauer wäre in Anschlussstudien zu untersuchen, welche konkrete Rolle den Schulleitungen im Prozess der Programmentwicklung zufällt, das heißt, ob sie eher eine konzeptuelle oder eine moderierende Funktion übernehmen, inwieweit die Programmarbeit unter Einbeziehung der Lehrkräfte erfolgt bzw. von den Schulleitungen „monopolisiert" wird (vgl. Heinrich/Kussau 2010, S. 178) und ob – je nach Partizipationsgrad – die Akzeptanz von Evaluationsprogrammen im Kollegium variiert. Die oben berichteten Forschungsbefunde zumindest sprechen für die Bedeutsamkeit der Partizipation des Kollegiums bzw. einer unterstützenden und ermutigenden Schulleitung für das Gelingen der internen Evaluation bzw. die Qualität der Schulprogramme (vgl. Burkard 2004; Haenisch/Burkhard 2002; Haenisch/Kindervater 1999; Holtappels/Müller 2002; Schildkamp/Visscher 2009).

Auf den Stellenwert einer breiten Beteiligung der Lehrkräfte weisen auch unsere Befunde zur Einrichtung von Steuergruppen hin: Steuergruppen haben eher Einfluss auf die Qualität der Evaluationsprogramme, wenn sie eine koordinierende Funktion übernehmen, die inhaltliche Arbeit aber in den Händen des Gesamtkollegiums bleibt.

Neben dem Einfluss der Schulleitungen und der Steuergruppen auf die Qualität der Programme können die Ergebnisse auch als Bestätigung der Wirksamkeit der Berliner „Qualifizierungsoffensive" gewertet werden: Die schulinterne Benennung von Evaluationsverantwortlichen begünstigt die Qualität der Evaluationsprogramme. Dieser bundeslandübergreifende Befund weist aber zugleich darauf hin, dass neben einer spezifischen Qualifizierung bereits die eindeutige Festlegung der Zuständigkeit eine wichtige Gelingensbedingung schulinterner Evaluation darstellt.

Insgesamt geben die Ergebnisse der Untersuchung wichtige Hinweise darauf, dass sowohl landesspezifische Implementationsbedingungen als auch innerorganisationale Bedingungen die Qualität von Evaluationsprogrammen und damit auch ihre Eignung als Instrument schulischer Selbststeuerung beeinflussen. Inwieweit die Programme in der Praxis für strategische Entwicklung genutzt werden, kann auf der Grundlage dieser Untersuchung nicht beantwortet werden. Auf jeden Fall aber haben Schulen mit einem qualitativ hochwertigen Evaluationsprogramm eine gute Grundlage für die Etablierung einer wirksamen Evaluationspraxis im Sinne einer datenbasierten schulischen Selbststeuerung.

Literatur und Internetquellen

Altrichter, H./Posch, P. (1999): Wege zur Schulqualität. Studien über den Aufbau von qualitätssichernden und qualitätsentwickelnden Systemen in berufsbildenden Schulen. Innsbruck: Studien Verlag.

Berkemeyer, N./Müller, S. (2010): Schulinterne Evaluation – nur ein Instrument zur Selbststeuerung von Schulen? In: Altrichter, H./Maag Merki, K. (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS, S. 195-218.

Bildungsserver Berlin-Brandenburg. URL: http://bildungsserver.berlin-brandenburg.de/schulformen_schularten.html; Zugriffsdatum: 10.05.2013.

Boller, S. (2009): Kooperation in der Schulentwicklung: Interdisziplinäre Zusammenarbeit in Evaluationsprojekten. Wiesbaden: VS.

Burkard, C. (2004): Funktionen und Schwerpunkte von Schulprogrammen aus Sicht der Schulaufsicht. In: Holtappels, H.G. (Hrsg.): Schulprogramme – Instrumente der Schulentwicklung. Weinheim: Juventa, S. 137-154.

Cohen, J. (1988): Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Devos, G./Verhoeven, J.C. (2003): School Self-Evaluation – Conditions and Caveats. The Case of Secondary Schools. In: Educational Management & Administration 31, H. 4, S. 403-420.

Haenisch, H./Burkard, C. (2002): Schulprogrammarbeit erfolgreich gestalten. Ergebnisse einer qualitativen Studie zu den Gelingensbedingungen der Entwicklung und Umsetzung des Schulprogramms. In: Ministerium für Schule, Wissenschaft und Forschung/Landesinstitut für Schule und Weiterbildung NRW (Hrsg.): Schulprogrammarbeit in Nordrhein-Westfalen. Ergebnisse der wissenschaftlichen Evaluationsstudien. Bönen: Kettler, S. 123-197.

Haenisch, H./Kindervater, C. (1999): Evaluation der Qualität von Schule und Unterricht. EU-Pilotprojekt zur Selbstevaluation: Ergebnisse der deutschen Projektschulen. Bönen: Verlag für Schule und Weiterbildung.

Heinrich, M./Kussau, J. (2010): Das Schulprogramm zwischen schulischer Selbstregelung und externer Steuerung. In: Altrichter, H./Maag Merki, K. (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS, S. 171-194.

Hofman, R.H./Dijkstra, N.J./Hofman, W.H.A. (2009): School Self-Evaluation and Student Achievement. In: School Effectiveness and School Improvement 20, H. 1, S. 47-68.

Holtappels, H.G./Müller, S. (2002): Inhalte und Struktur von Schulprogrammen. Inhaltsanalyse der Schulprogrammtexte Hamburger Schulen. In: Holtappels, H.G./Schulz-Zander, R./Pfeiffer, H./Klemm, K./Rolff, H.-G. (Hrsg.): Jahrbuch der Schulentwicklung, Bd. 12. Daten, Beispiele und Perspektiven. Weinheim/München: Juventa, S. 209-231.

Kanders, M. (2004): Schulprogrammarbeit in NRW. In: Holtappels, H.G. (Hrsg.): Schulprogramme – Instrumente der Schulentwicklung. Weinheim: Juventa, S. 117-135.

Lind, G. (2012): Effektstärken: Statistische, praktische und theoretische Bedeutsamkeit von empirischen Studien. Überarbeiteter Vortrag. KOPS, Universität Konstanz.

LISUM (Landesinstitut für Schule und Medien) (2003): Stationen auf dem Weg der Schulprogrammentwicklung. URL: http://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/schulqualitaet/schulentwicklung/schulprogrammentwicklung/Leitfaden.pdf; Zugriffsdatum: 10.05.2013.

LISUM (Landesinstitut für Schule und Medien) (2007): Leitfaden zur internen Evaluation für Schulen in Berlin. URL: http://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/schulqualitaet/qualitaetssicherung/pdf/Leitfaden.pdf; Zugriffsdatum: 10.05.2013.

MBJS (Ministerium für Bildung, Jugend und Sport) (2000): Die Schule soll im Dorf bleiben: Das Modell der kleinen Grundschule. URL: http://www.mbjs.brandenburg.de/sixcms/media.php/bb2.a.5813.de/Modell_kleine_Grundschule.pdf; Zugriffsdatum: 10.05.2013.

MBJS (Ministerium für Bildung, Jugend und Sport) (2003): Anwerbung, berufliche Entwicklung und Verbleib von qualifizierten Lehrerinnen und Lehrern. URL: http://www.oecd.org/germany/31078073.pdf; Zugriffsdatum: 10.05.2013.

MBJS (Ministerium für Bildung, Jugend und Sport) (2004): Schulprogramme in Brandenburg – Selbsteinschätzungen und Fremdbeurteilung. URL: http://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/schule/schulformen_und_schularten/ganztagsschulen/gestaltung/Schulprogramme_BB.pdf; Zugriffsdatum: 10.05.2013.

MBJS (Ministerium für Bildung, Jugend und Sport) (2007): Brandenburgisches Schulgesetz. URL: http://www.mbjs.brandenburg.de/sixcms/media.php/5527/broschuere_schulgesetz_2007.pdf; Zugriffsdatum: 10.05.2013.

MBJS (Ministerium für Bildung, Jugend und Sport) (2009): Rundschreiben 8/09 vom 27.07.2009. Fortschreibung und Evaluation von Schulprogrammen an Schulen in öffentlicher Trägerschaft im Land Brandenburg. URL: http://www.mbjs.brandenburg.de/sixcms/media.php/bb1.c.173333.de; Zugriffsdatum: 10.05.2013.

MBJS (Ministerium für Bildung, Jugend und Sport) (2011): Verwaltungsvorschriften über Ganztagsangebote an allgemeinbildenden Schulen (VV-Ganztag). URL: http://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/aus_und_fortbildung/fortbildung/News_2012/Verwaltungsvorschriften_GT_19.03.12.pdf; Zugriffsdatum: 10.05.2013.

Meuret, D./Morlaix, S. (2003): Conditions of Success of a School's Self-Evaluation. Some Lessons of an European Experience. In: School Effectiveness and School Improvement 14, H. 1, S. 53-71.

Meuser, M./Nagel, U. (2009): The Expert Interview and Changes in Knowledge Production. In: Bogner, A./Littig B./Menz, W. (Hrsg.): Interviewing Experts. Hampshire: Palgrave Macmillan, S. 17-42.

Mintrop, H./MacLellan, A.M./Quintero, M.F. (2001): School Improvement Plans in Schools on Probation: A Comparative Content Analysis across Three Accountability Systems. In: Educational Administration Quarterly 37, H. 2, S. 197-218.

Mintzberg, H. (1983): Structure in Fives. Designing Effective Organizations. Englewood Cliffs, NJ: Prentice-Hall.

Mohr, I. (2006): Analyse von Schulprogrammen. Eine Arbeit im Rahmen der Internationalen Grundschul-Lese-Untersuchung (IGLU). Münster u.a.: Waxmann.

Müller, S. (2002): Schulinterne Evaluation – Gelingensbedingungen und Wirkungen. Dortmund: IFS-Verlag.

Scheerens, J./Demeuse, M. (2005): The Theoretical Basis of the Effective School Improvement Model (ESI). In: School Effectiveness and School Improvement 16, H. 4, S. 373-385.

Schildkamp, K./Ehren, M./Lai, M.K. (2012): Editorial Article for the Special Issue on Data-based Decision Making around the World: from Policy to Practice to Results. In: School Effectiveness and School Improvement 23, H. 2, S. 123-131.

Schildkamp, K./Visscher, A. (2009): Factors Influencing the Utilisation of a School Self-Evaluation Instrument. In: Studies in Educational Evaluation 35, H. 4, S. 150-159.

Schildkamp, K./Visscher, A./Luyten, H. (2009): The Effects of the Use of a School Self-Evaluation Instrument. In: School Effectiveness and School Improvement 20, H. 1, S. 69-88.

Schlömerkemper, J. (2004): Einstellungen und Erwartungen zum Schulprogramm. In: Holtappels, H.G. (Hrsg.): Schulprogramm als Schulentwicklungsinstrument. Weinheim: Juventa, S. 61-78.

SEIS Deutschland. URL: http://www.seis-deutschland.de/; Zugriffsdatum: 10.05.2013.

SenBJS (Senatsverwaltung für Bildung, Jugend und Sport) (2004): Schulqualität in Berlin (SQIB) 1999-2004. URL: http://www.berlin.de/imperia/md/content/sen-bildung/schulqualitaet/sqib_bericht_102004.pdf?start&ts=1155494303&file=sqib_bericht_102004.pdf; Zugriffsdatum: 10.05.2013.

SenBWF (Senatsverwaltung für Bildung, Wissenschaft und Forschung) (2007): Schulinspektionen im Schuljahr 2005/2006. URL: http://www.berlin.de/imperia/md/content/sen-bildung/schulqualitaet/schulinspektion/schulinspektionen_2005_2006.pdf?start&ts=1178886155&file=schulinspektionen_2005_2006.pdf; Zugriffsdatum: 10.05.2013.

SenBWF (Senatsverwaltung für Bildung, Wissenschaft und Forschung) (2008): Ausführungsvorschriften zur Erstellung der Schulprogramme und zur internen Evaluation (AV Schulprogramm). URL: http://www.berlin.de/imperia/md/content/sen-bildung/rechtsvorschriften/av_schulprogramm.pdf; Zugriffsdatum: 10.05.2013.

SenBWF (Senatsverwaltung für Bildung, Wissenschaft und Forschung) (Hrsg.) (2009): Handbuch Schulinspektion. Berlin. URL: http://www.fosbe.de/fileadmin/daten/Schulentwicklung/Schulinspektion/handbuch_schulinspektion.pdf; Zugriffsdatum: 10.05.2013.

SenBWF (Senatsverwaltung für Bildung, Wissenschaft und Forschung) (2010a): Schulgesetz für Berlin. URL: http://www.berlin.de/imperia/md/content/sen-bildung/rechtsvorschriften/schulgesetz.pdf?start&ts=1303211502&file=schulgesetz.pdf; Zugriffsdatum: 10.05.2013.

SenBWF (Senatsverwaltung für Bildung, Wissenschaft und Forschung) (2010b): Verwaltungsvorschrift über die einheitliche Gestaltung und Zuordnung von Aufgabenbereichen an öffentlichen Schulen des Landes Berlin (VV Zuordnung). URL: http://www.pr-ts.de/pdf/vv_zuord.pdf; Zugriffsdatum: 10.05.2013.

SenBWF (Senatsverwaltung für Bildung, Wissenschaft und Forschung) (2011): Schulinspektionen im Schuljahr 2009/2010. URL: http://www.berlin.de/imperia/md/content/sen-bildung/schulqualitaet/schulinspektion/schulinspektion_2009_2010.pdf?start&ts=1319707825&file=schulinspektion_2009_2010.pdf; Zugriffsdatum: 10.05.2013.

Thiel, F. (2008): Organisationssoziologische Vorarbeiten zu einer Theorie der Schulentwicklung. In: Journal für Schulentwicklung 12, H. 2, S. 31-39.

Thiel, F./Thillmann, K. (2012): Interne Evaluation als Instrument der Selbststeuerung von Schulen. In: Wacker, A./Maier, U./Wissinger, J. (Hrsg.): Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung – Empirische Befunde und forschungsmethodische Implikationen. (Reihe „Educational Governance", Bd. 9.) Wiesbaden: VS, S. 35-55.

Thiel, F./Thillmann, K./Vogel, S./Stanat, P./Wendt, W. (2010): Abschlussbericht „Bestandsaufnahme der Sprachfördermaßnahmen an Berliner Grundschulen" (unveröffentlichter Bericht).

Ulber, D./Buchholz, N. (2007): Qualifizierung von Berliner Lehrerinnen und Lehrern zu Evaluationsberaterinnen und -beratern. In: Krämer, M./Preiser, S./Brusdeylins, K. (Hrsg.): Psychologiedidaktik und Evaluation VI. Göttingen: Vandenhoeck & Ruprecht Unipress, S. 63-69.

Vereinigung der Bayerischen Wirtschaft e.V. (2010): Bildungsautonomie: Zwischen Regulierung und Eigenverantwortung. Jahresgutachten 2010. URL: http://www.aktionsrat-bildung.de/fileadmin/Dokumente/Jahresgutachten_2010.pdf; Zugriffsdatum: 10.05.2013.

Wirtz M./Caspar F. (2002): Beurteilerübereinstimmung und Beurteilerreliabilität. Göttingen: Hogrefe.

Katja Thillmann, Dr., geb. 1978, wissenschaftliche Mitarbeiterin an der Freien Universität Berlin.
E-Mail: k.thillmann@fu-berlin.de

Anabel Bach, Dipl.-Psych., geb. 1982, wissenschaftliche Mitarbeiterin an der Freien Universität Berlin.
E-Mail: anabel.bach@fu-berlin.de

Christina Gerl, geb. 1988, studentische Mitarbeiterin an der Freien Universität Berlin.
E-Mail: c.gerl@fu-berlin.de

Felicitas Thiel, Prof. Dr., geb. 1961, Leiterin des Arbeitsbereichs Schulpädagogik/Schulentwicklungsforschung an der Freien Universität Berlin.
E-Mail: felicitas.thiel@fu-berlin.de

Anschrift: Arbeitsbereich Schulpädagogik/Schulentwicklungsforschung, Freie Universität Berlin, Habelschwerdter Allee, 45 14195 Berlin

STEUERUNG UND UMSTRUKTURIERUNG VON SCHULEN IN HERAUSFORDERNDEN LAGEN

DDS – Die Deutsche Schule
Beiheft 12, S. 226-254
© 2013 Waxmann

Kathrin Racherbäumer/Christina Funke/Isabell van Ackeren/Marten Clausen

Datennutzung und Schulleitungshandeln an Schulen in weniger begünstigter Lage

Empirische Befunde zu ausgewählten Aspekten der Qualitätsentwicklung

Zusammenfassung
Schulen, die sich an Standorten in schwierigen sozialen Kontexten befinden („Brennpunktschulen"), sind im deutschen Sprachraum bislang deutlich weniger beforscht, als dies international der Fall ist. Dabei ist die Arbeit in diesen Schulen besonders herausfordernd und im Hinblick auf die Förderung leistungsschwacher und benachteiligter Schülerinnen und Schüler zugleich besonders relevant, um reibungslose Übergänge in Ausbildung und Beruf zu fördern. Die spezifischen Herausforderungen werden im Beitrag vor dem Hintergrund des Forschungsstandes skizziert. Zu den besonders relevanten Aspekten der datengestützten Schulentwicklung und des darauf bezogenen Schulleitungshandelns werden Befunde aus einem Forschungsprojekt im Fallstudiendesign auf der Grundlage quantitativer und qualitativer Daten präsentiert. Dabei werden Gemeinsamkeiten und Unterschiede zwischen so genannten erwartungswidrig guten und erwartungskonform schwachen Schulen beschrieben.
Schlüsselwörter: Schulen in schwieriger Lage, Evidenzbasierung, Datengestützte Schulentwicklung, Schulleitung

Quality Development of Schools in Challenging Circumstances

Empirical Findings on Data Use and Quality Development

Summary
Much less research has been conducted on schools in challenging circumstances ("focus schools") in the German-speaking area than internationally. The work in these schools

is particularly challenging, and in order to encourage low-achieving and disadvantaged students it is particularly relevant to promote smooth transitions in vocational training and employment. The specific challenges are outlined in the article in the light of current research. Particularly relevant aspects are the data-based school development and school management. In this context, we present findings from a research project in case study design, based on quantitative and qualitative data. Similarities and differences between surprisingly well performing and unsurprisingly weak performing schools are described.
Keywords: schools in challenging circumstances, evidence based / data-driven school development, school management

1. Schulen in schwieriger Lage: Stand der Forschung

In den letzten Jahren sind so genannte Brennpunktschulen zunehmend in den Fokus der Öffentlichkeit gerückt. Hierbei handelt es sich um Schulen, die aufgrund ihrer Lage in sozial segregierten Stadtteilen mit einer Kumulation von negativen Kompositionsmerkmalen konfrontiert sind (vgl. Baumert/Stanat/Watermann 2006). Die Schulstandorte sind durch einen hohen Anteil an sozial benachteiligten Schülerinnen und Schülern aus bildungsfernen Elternhäusern mit und ohne Migrationshintergrund charakterisiert. Auf Einzelschulebene wirken sich diese Kontextbedingungen insofern ungünstig aus, als diese Schulen zum Beispiel in nationalen Vergleichsarbeiten wie VERA-8 in der Regel unterdurchschnittliche Resultate erzielen (vgl. Bonsen u.a. 2010, S. 27ff.). Darüber hinaus lässt sich der geringere Bildungserfolg an höheren Anteilen von Schülerinnen und Schülern ohne Schulabschluss bzw. niedrigeren Übergangszahlen in das berufliche Übergangssystem festmachen.

Diese ungünstigen Rahmenbedingungen und schwachen Lernergebnisse sowie Probleme beim Übergang in die berufliche Ausbildung können insgesamt zu einem negativen Schulimage beitragen, das es der Schule erschwert, in schulscharfen Ausschreibungsverfahren gute Lehrkräfte zu gewinnen bzw. gute Lehrkräfte längerfristig an ihre Schulen zu binden. An Standorten in schulischer Konkurrenzlage dürften diese Schulen darüber hinaus eher Probleme haben, ihre Schülerzahlen zu halten.

Auf der Unterrichtsebene sehen sich die Lehrkräfte mit Schülerinnen und Schülern konfrontiert, die von ihren Eltern zumeist wenig Unterstützung erfahren. Dies äußert sich durch fehlende Schulmaterialien oder durch häufiges Zuspätkommen bis hin zu Schulabstinenz. Der Unterricht ist oftmals durch ein hohes Maß an Disziplinproblemen geprägt, wodurch die für diese Schülerinnen und Schüler so wichtige Lernzeit weiter minimiert wird (vgl. Racherbäumer/Funke/van Ackeren 2013). Das Zusammenspiel dieser unterschiedlichen Faktoren wirkt sich negativ auf die Leistungen der Schülerinnen und Schüler aus und kann zudem das Belastungserleben der Lehrkräfte verstärken (vgl. Maaz/Baumert 2011).

1.1 Qualitätsmerkmale erwartungswidrig guter Schulen in schwieriger Lage

Durch die Einführung zentral administrierter Tests und Prüfungen (VERA-3 bzw. VERA-8, zentrale Abschlussprüfungen am Ende der Sekundarstufe I, Zentralabitur) ist es möglich, hinsichtlich des kognitiven Lernertrags unterschiedlich erfolgreiche Schulen zu identifizieren und zu systematisieren. Unterschieden werden können so genannte erwartungskonform schwache Schulen, die auch unter Berücksichtigung ihres Schulstandortes (über Sozialindices, vgl. Forschungsdesign unten) schlechte Schülerleistungen erzielen**,** ferner *Turnaround*-Schulen, bei denen eine positive Entwicklung mit Blick auf die Schülerleistung zu beobachten ist, und erwartungswidrig gute Schulen, denen es kontinuierlich gelingt, trotz erschwerter Kontextbedingungen gute Schülerleistungen zu befördern.[1] International sind die Schul- und Unterrichtsentwicklungsprozesse von so genannten *Improving Schools in Challenging Circumstances* seit geraumer Zeit Gegenstand wissenschaftlicher Betrachtung. Eine Reihe von Studien konnte verschiedene Qualitätsmerkmale dieser Schulen identifizieren.

Merkmale der Organisationskultur und Zusammenarbeit der Organisationsmitglieder. Insgesamt wird eine positive Schulkultur als ein zentrales Merkmal von erwartungswidrig guten Schulen beschrieben (vgl. Muijs u.a. 2004). In diesem Zusammenhang ist ein freundliches, wertschätzendes Schul- und Lernklima entscheidend, um den Lernerfolg der Schülerinnen und Schüler zu befördern.[2] Ein weiterer Aspekt der Schulkultur ist die Zusammenarbeit im Kollegium – einerseits zur Herstellung einer positiven Schul- und Unterrichtskultur[3] und andererseits um Arbeitserleichterungen für die Lehrkräfte selbst zu erzielen. So wurde die Bedeutung der Anzahl und Art von Kontakten zwischen Organisationsmitgliedern für die Effektivität von Schulen

1 Anzumerken ist an dieser Stelle, dass die Differenzierung auf der Grundlage zentraler Tests und Prüfungen einem eindimensionalen, outputorientierten Qualitätsverständnis folgt. Weitere, eher prozessbezogene Aspekte der Schulqualität, wie z.B. das Schul- oder Unterrichtsklima, bleiben als abhängige Variablen bei der Bewertung unberücksichtigt. Obgleich es einen kritischen Diskurs zu diesem Ansatz gibt, werden in den USA und England Schulen nach diesem Verfahren mit wiederum kontrovers diskutierten (Neben-)Wirkungen kategorisiert (vgl. hierzu z.B. Hermann 2013; Quesel/Husfeldt 2013).

2 Schulklima kann als Teil von Schulkultur aufgefasst werden. Das Schulklima betrifft dabei eher die Gefühlsatmosphäre, während die Schulkultur darüber hinaus – neben gemeinsamen Werten und Zielorientierungen – konkrete Organisations- und Handlungsstrukturen meint.

3 Stoll und Fink (1996, S. 25) haben kulturelle Einflussgrößen identifiziert, die sie als „norms of improving schools" wie folgt zusammenfassen (die *catch-phrases* in Klammern artikulieren typische Kernaussagen): shared goals (we know where we're going), responsibility for success (we must succeed), collegiality (we're working on this together), continuous improvement (we can get better), lifelong learning (learning is for everyone), risk taking (we learn by trying something new), support (there's always someone there to help), mutual respect (everyone has something to offer), openness (we can discuss our differences), celebration and humor (we feel good about ourselves).

vielfach formuliert (vgl. zusammenfassend Bauer 2002; Kullmann 2010). Dies lässt sich mit der Teilhabe an gemeinsamem Wissen, der Erhöhung der Produktivität, der Entwicklung von Expertise, verbesserten Beziehungen und der Förderung des Selbstvertrauens erklären (vgl. Cavanagh 1997). Diesen Aspekten kommt insbesondere an Schulen in schwierigen Kontexten vor dem Hintergrund der erhöhten Belastung der Lehrpersonen besondere Bedeutung zu. Die Bedeutung des Aspektes der kollegialen Zusammenarbeit für die Schuleffektivität unterstreichen z.B. Scheerens und Bosker (1997) in ihrer Metastudie. In der deutschsprachigen Forschung werden Praktiken der Lehrerkooperation aktuell ebenfalls im Zusammenhang mit Ansätzen der schulischen Qualitätssicherung und -entwicklung diskutiert und in ihren Formen und Wirkungen überprüft (vgl. z.B. Fussangel/Gräsel 2012; Steinert u.a. 2006; Huber/Ahlgrimm/Hader-Popp 2012; West 2012).

Merkmale des Unterrichts / der Unterrichtsentwicklung. Ein weiterer Fokus liegt auf dem Kerngeschäft der Schule, dem Unterricht. Hier wird international ein instruktionaler Fokus auf den Unterricht als zentrales Merkmal effektiver Schulen beschrieben (vgl. z.B. Reynolds u.a. 2001). Anknüpfend an Ergebnisse der Lehr- und Lernforschung zeigt sich, dass lernschwache Schülerinnen und Schüler stärker strukturierter Lernprozesse und eines geringer portionierten Curriculums in Verbindung mit einem unmittelbaren Feedback bedürfen (vgl. Ledoux/Overmaat 2001). Zudem gibt es Hinweise darauf, dass erwartungswidrig gute Schulen eine begrenztere Anzahl an Lernzielen und auch kurzfristigere Lernziele formuliert haben (vgl. Teddlie u.a. 1989). Zugleich gilt es, Unterforderung zu vermeiden und anspruchsvolles Lernen zu fördern. Verschiedene Studien konnten darüber hinaus zeigen, dass Schülerinnen und Schüler durchaus in der Lage sind, höhere Kompetenzstufen angesichts ihrer kognitiven Grundfähigkeiten zu erreichen, als vorher angenommen wurde (vgl. Leithwood/Steinbach 2002).

Merkmale der Personalentwicklung. Weiterhin gibt es Hinweise, dass *Improving Schools* mehr Zeit und Anstrengungen in die professionelle Entwicklung der Lehrkräfte investiert haben, insbesondere durch unterrichtsbezogene Fortbildung mit starkem Praxisbezug bei gleichzeitiger theoretischer Fundierung und individuellem Feedback (vgl. Freeman 1997; Henchey 2001; Reynolds u.a. 2001).

Merkmale der Vernetzung im Umfeld. Darüber hinaus benötigen gerade Schulen in benachteiligter Lage zur Bewältigung der je spezifischen Problemstellungen Unterstützung des lokalen Umfelds als erweitertes Sozialkapital der Schule; dies bietet Zugang zu den Ressourcen des sozialen Lebens, wie Wissen, Unterstützung, Hilfestellung, aber auch Anerkennung. Hinsichtlich der Interaktion zwischen der Schule und ihrer Umgebung erscheint dabei die Intensität von Außenkontakten besonders bedeutsam. Diese trägt dazu bei, dass Schülerinnen und Schülern inhaltlich erweiterte Lernmöglichkeiten geboten werden. In einer Interview-Studie von Seeley, Niemeyer und Greenspan (1990) benannten die befragten Schulleitungen erwar-

tungswidrig guter Schulen besonders häufig die Einbindung der Eltern in ihre Arbeit. Neben der Zusammenarbeit mit den Eltern können aber auch Netzwerke mit lokalen Unternehmen und Betrieben sowie die Kooperation mit Partnern, etwa aus der Sozial- und Jugendarbeit, förderlich für das Lernen sowie die Motivationsentwicklung sein.

Die Herausforderungen einer solchen institutionellen Einbindung von Schulen in sozialräumliche Kooperations- und Entwicklungsprozesse sind dabei auch empirisch beschrieben (vgl. Burchardt 2008). Als günstig hat sich dabei neben einer kommunalpolitischen Priorisierung des Themas *Schule und Bildung* eine die Arbeits- und Kommunikationsstrukturen institutionalisierende und auf administrativer Ebene gefestigte Netzwerkstruktur erwiesen (vgl. z.B. Solzbacher/Minderhop 2007; Tippelt u.a. 2009). Schließlich wird auch die Einbindung von Schulen in Schul-Netzwerke, in denen im Sinne professioneller Lerngemeinschaften Ideen und Konzepte für die Praxis ausgetauscht werden, als qualitätsfördernd beschrieben (vgl. Hopkins/Reynolds 2002; Berkemeyer u.a. 2008).

1.2 Zur besonderen Bedeutung interner und externer Daten für die Schul- und Unterrichtsentwicklung an Schulen in schwieriger Lage

Instrumente wie Zentralprüfungen, Vergleichsarbeiten und Schulinspektionen bieten Einzelschulen datengestützte Anknüpfungspunkte für schulische Qualitätsentwicklung und ermöglichen ihrer Intention nach die Generierung und Nutzbarmachung von Steuerungs- und Handlungswissen auf der Ebene der Einzelschule. Die systematische Steuerung und Entwicklung einer Schule setzt jedoch zunächst die gezielte Kenntnisnahme und Auswahl von verfügbaren, rückgemeldeten Daten bzw. gegebenenfalls die möglichst passgenaue und differenzierte Datengenerierung voraus, um daran anknüpfend Ist- und Soll-Zustände zu vergleichen und Entwicklungen transparent abzubilden.

Dabei ist zwischen sehr unterschiedlichen Datenquellen zu unterscheiden. Sie reichen von landesweit angelegten Erhebungen, wie z.B. VERA-8, bis hin zu schulintern erhobenen Daten, wie z.B. aus unterrichtsbezogenen Feedbacks von Schülerinnen und Schülern. Reynolds u.a. (2001) und Hopkins (2001) beschreiben ausgehend von internationalen Ergebnissen einen Zusammenhang zwischen dem Datenreichtum und dem Erfolg von Schulen, die zudem unter schwierigen Bedingungen arbeiten. Auch zeichnen sich Schulen, die sich vergleichsweise stark verbessern konnten, dadurch aus, dass sie relevante Expertise und Beratung im Umgang mit Daten und ihrer schwierigen Situation mobilisieren konnten.

Forschungsergebnisse zur tatsächlichen Nutzung von Daten für Schul- und Unterrichtsprozesse in Deutschland zeigen hier ein unterschiedliches Rezeptionsverhal-

ten der Lehrkräfte (vgl. Kühle/Peek 2007; Nachtigall/Jantowski 2007; Schrader/Helmke 2003; Groß Ophoff/Hosenfeld/Koch 2007; Maier 2008; Hartung-Beck 2009; Schulze 2012). Als förderlich für die Datennutzung werden z.B. die Bereitschaft zur Veränderung als Bestandteil der Lehrerprofession bzw. fachdidaktisches und fachwissenschaftliches Wissen zur Unterrichtsentwicklung beschrieben (vgl. Dedering 2011). Darüber hinaus muss es Lehrkräften in den entsprechenden Fachkonferenzen gelingen, fachdidaktisch orientierte Neuerungen zu sammeln und systematisch mit Implikationen externer Leistungsindikatoren zu verknüpfen (vgl. Maier/Kuper 2012). Dabei ist ein positives Innovationsklima von hoher Bedeutung (vgl. ebd.).

Diese Ergebnisse deuten darauf hin, dass an erfolgreichen Schulen in schwieriger Lage besonders kompetente, innovationsbereite Schulleiterinnen und Schulleiter und Lehrkräfte agieren. Hier ist jedoch anzumerken, dass der Beleg für einen entsprechenden Zusammenhang zwischen der Kompetenz der Lehrkräfte bzw. der Schulleitungen zur datengestützten Schulentwicklung auf der einen und dem Schulerfolg von Schulen in herausfordernder Lage auf der anderen Seite noch aussteht.

1.3 Zur besonderen Bedeutung des Leitungshandelns an Schulen in schwieriger Lage

Die Darstellung dieser unterschiedlichen Befunde zu erfolgreichen Schulen in schwieriger Lage macht deutlich, dass der Schulleitung innerhalb des schulischen Mehrebenensystems eine erhöhte Bedeutung beigemessen werden muss. Sie kann in besonderem Maße auf die beschriebenen Qualitätsdimensionen Einfluss nehmen (vgl. z.B. Bonsen 2010; Harris u.a. 2006; Huber/Muijs 2010; Reynolds u.a. 2001; Schratz/Hartmann/Schley 2010). Zudem zeigt die (inter-)nationale Forschung, dass Schulleitungen innerhalb neuer, veränderter bildungspolitischer Rahmenbedingungen, die einerseits durch eine erhöhte, datengestützte Rechenschaftspflicht und andererseits durch mehr Eigenständigkeit gekennzeichnet sind, mit einem breiteren Aufgabenspektrum konfrontiert sind (vgl. Pont u.a. 2008; Brauckmann 2012). Als Hauptaufgabe benennt Wissinger (2011) „die Verbesserung der Schülerleistungen als Zielpunkt schulischen Handelns und als Antwort auf spezifische soziale und ethische Kontexte des Lehrens und Lernens“ (S. 104), womit auch das adaptive Handeln der Schulleitungen den schulischen Kontexten entsprechend impliziert ist. Dieses Handeln soll – im Sinne neuer Steuerungsinstrumente und damit einhergehender empirisch generierter Befunde – möglichst datengestützt erfolgen (vgl. dazu den Beitrag von van Ackeren u.a. zur evidenzbasierten Schulentwicklung in diesem Band). Die Prozesse der Rezeption und Nutzbarmachung verfügbarer Informationen zum schulischen Umfeld bzw. den spezifischen Lernvoraussetzungen der Schülerinnen und Schüler sowie zu den schulischen Lernerträgen erfordern Anleitung und Koordination, so dass der Schulleitung eine Schlüsselposition zukommt. Sie muss beispielsweise wissen, welche Befunde für den individuellen Schulstandort Relevanz haben,

welche Daten die Schule daher braucht (und welche nicht), und sie muss insbesondere durch die Schaffung entsprechender Rahmenbedingungen (z.B. Ermöglichung von Zeiträumen, entsprechende Fortbildungsplanung etc.) sicherstellen, dass schulinterne Analyse- und Austauschprozesse auf der Ebene der Lehrkräfte stattfinden können (vgl. Rolff 2010; Rolff 2013; Kühn/Racherbäumer 2011).

Führungsstile von Schulleitungen

Allgemeine Befunde. Im Rahmen der international vergleichenden Studie LISA *(Leadership Improvement for Student Achievement)* konnten durch die Befragung von Lehrkräften und Schulleitungen fünf unterschiedliche Führungsstile identifiziert werden. Hierbei handelt es sich um (a) den *unterrichtsbezogenen Führungsstil*, der u.a. durch die Bereitstellung von Lehr- und Lernressourcen sowie ein gezieltes Monitoring von Lehr- und Lernstandards gekennzeichnet ist, und (b) den *personalen Führungsstil*, der sich durch die Anerkennung und Belohnung von Lehrkräften sowie die Ermunterung zur fachlichen Weiterbildung auszeichnet. Ferner gibt es (c) den *partizipativen Führungsstil*, der sich u.a. durch die Förderung offener Kommunikation und Entscheidungsfindung und die Gewährung unterrichtsbezogener Autonomie auszeichnet. Darüber hinaus konnten (d) der *unternehmerische Führungsstil*, der sich durch die Förderung von Beziehungen zwischen der Kommune und der Schule sowie den Eltern und einer positiven Außendarstellung auszeichnet, sowie (e) der *strukturierende Führungsstil*, der durch ein hohes Maß an Transparenz über Funktionen und Aufgaben von Regeln und eine geordnete Arbeitsatmosphäre charakterisiert ist, identifiziert werden (vgl. Brauckmann 2012). Die Ergebnisse zur quantitativen Bedeutsamkeit der Führungsstile zeigen, dass in allen Staaten eine Kombination aller fünf Führungsstile in unterschiedlicher Ausprägung vorkommt (vgl. ebd.). Für Deutschland konnte der unternehmerische Führungsstil als dominantes Handlungsmuster identifiziert werden. Der unterrichtsbezogene Führungsstil, der wie beschrieben auch ein hohes Maß an Datennutzung impliziert, kommt deutlich seltener vor. Ebenso konnte der partizipative Führungsstil, der mit den beschriebenen Eigenschaften einen wesentlichen Beitrag zur Herstellung einer positiven Schulkultur bieten kann, weniger oft identifiziert werden.

Befunde zu Schulen in schwierigen Kontexten. Das Schulleitungshandeln wurde mit Blick auf den praktizierten Führungsstil in verschiedenen internationalen Studien zu erwartungswidrig guten Schulen näher betrachtet. Keys u.a. (2003) kommen zu dem Ergebnis, dass ein kooperativer/partizipativer Führungsstil, der zudem durch eine hohe Erwartungshaltung an die Lehrkräfte sowie die Schülerinnen und Schüler gekennzeichnet ist, vorteilhaft ist. Die Arbeitsweise der Schulleitungen sollte sich ferner durch ein hohes Maß an Transparenz auszeichnen, das durch demokratische Entscheidungsprozesse aller schulischen Akteursgruppen flankiert wird (vgl. McDougall u.a. 2006; Flintham 2006). In diesem Kontext wird mit Bezug auf den

sogenannten distributiven Führungsstil[4] das Delegieren von Aufgaben an spezielle Arbeitsgruppen positiv hervorgehoben (vgl. Harris 2004), um Teilhabe zu befördern.

Zudem legen auf der Grundlage vorliegender Befunde Schulleitungen an erfolgreichen Schulen in schwieriger Lage einen deutlichen Schwerpunkt ihrer Arbeit auf die Entwicklung einer positiven Schulkultur sowie auf das Unterrichtsgeschehen, während administrative Aufgaben weniger im Mittelpunkt stehen. Zur schulischen Qualitätssicherung und Qualitätsentwicklung greifen diese Schulleiterinnen und Schulleiter systematisch auf schulinterne und -externe Daten zurück (vgl. Hopkins 2001; Reynolds u.a. 2001; Leithwood/Harris/Strauss 2010).

Diese internationalen Befunde zu *Improving Schools in Challenging Circumstances* sind jedoch nicht unmittelbar auf Deutschland übertragbar, da sich die bildungspolitischen Kontextbedingungen stark unterscheiden und differenzielle Wirkungen erwarten lassen. So werden Schulen im Zuge des so genannten *High-Stakes Monitoring* in England oder den USA sanktioniert, sofern gesetzte Lernziele nicht erreicht werden. Die bildungspolitischen Maßnahmen reichen dabei von Schulfusionen über die Entlassung einzelner Lehrkräfte und Schulleitungen bis hin zur Einstellung speziell geschulter Schulleiterinnen und Schulleiter. Darüber hinaus erhalten die betroffenen Schulen mehr materielle und personelle Ressourcen, um die Leistungen ihrer Schülerinnen und Schüler zu verbessern (differenzierte Mittelallokation). Im Vergleich dazu hat Deutschland durch die Implementierung zentraler Test und Prüfungen zwar grundlegende bildungspolitische Veränderungen von einer input- zu einer stärker outputorientierten Steuerung vorgenommen; gleichwohl sind gezielte Sanktionen als Folge eines wiederholt schlechten Abschneidens, wie sie oben mit Blick auf die USA und England beschrieben wurden, nicht vorgesehen.

2. Fragestellung

Vor dem Hintergrund der beschriebenen Qualitätsmerkmale erwartungswidrig effektiver Schulen an besonders herausfordernden Standorten fokussiert das im Weiteren vorgestellte Forschungsprojekt „*Strategien der Qualitätsentwicklung von Schulen in schwieriger Lage*“[5] unter anderem die hervorgehobenen Aspekte des Schulleitungshandelns im Allgemeinen bzw. des datengestützten Handelns dieser Schulen im Speziellen im Vergleich zu gemäß der schulischen Ausgangslage erwartungskonform (schwach) abschneidenden Schulen.

4 Der distributive Führungsstil wird unter Benennung ähnlicher Merkmale auch als partizipatorischer Führungsstil bezeichnet.

5 Gefördert vom Bundesministerium für Bildung und Forschung, Förderlinie „Steuerung im Bildungssystem“ (SteBis), Förderkennzeichen 01JG1008.

Im Untersuchungszusammenhang wurde auf der Ebene der Lehrkräfte danach gefragt, inwiefern gezielt Informationen bezüglich der Lern- und Leistungsentwicklung der Schülerinnen und Schüler aufgegriffen werden, um Maßnahmen der Förderung individuell anzupassen. Darüber hinaus wurde eine Einschätzung der Lehrkräfte zum entsprechenden Schulleitungshandeln eruiert. Zudem wurden die Schulleitungen nach ihren Strategien befragt, um schulische und unterrichtliche Prozesse systematisch und datengestützt zu evaluieren und weiterzuentwickeln. Vertiefend wurde der Frage nachgegangen, inwieweit die Rezeption verschiedener datengenerierender Steuerungsinstrumente der strategischen Testvorbereitung bzw. dem gezielten Testcoaching dient und dazu beiträgt, positive Testergebnisse zu erzielen.

Zusammenfassend ergeben sich daraus die folgenden Forschungsfragen, die vertiefend auf Unterschiede zwischen erwartungsgemäß bzw. erwartungswidrig gut abschneidenden Schulen hin untersucht wurden:

1) Wird die Nutzung von Daten formal von den Schulen benannt (z.B. in Schulprogrammen)? (nominelle Nutzung)
2) Welche internen und externen Daten stehen den befragten Lehrkräften zur Verfügung und tragen aus ihrer Sicht zur Unterrichtsentwicklung bei? (Gelegenheitsstrukturen / wahrgenommene Nützlichkeit)
3) Welche Strategien zur Nutzung interner bzw. externer Daten werden von Schulleitungen beschrieben? (strategisches Datennutzungshandeln)
4) Wie schätzen die Lehrkräfte in diesem Zusammenhang die Bedeutung des Schulleitungshandelns an Schulen in segregierter Lage insgesamt ein? (Bedeutsamkeit Schulleitungshandeln allgemein)
5) Wie schätzen Lehrkräfte das tatsächliche Schulleitungshandeln ein? (wahrgenommenes allgemeines Schulleitungshandeln)
6) Welche Führungsstrategien beschreiben Schulleitungen selbst? (selbstberichtetes strategisches Schulleitungshandeln)

3. Forschungsdesign

Dem Projekt „Strategien der Qualitätsentwicklung von Schulen in schwieriger Lage“ liegt ein Fallstudienansatz zugrunde, durch den im Rahmen einer „umfassenden, systematischen Darstellung der einzelschulischen Wirklichkeit“ (vgl. Horstkemper/Tillmann 2008, S. 298) Prozessverläufe differenzierter beschrieben werden können. Dabei ist ausgehend von einem theoretischen Sampling, innerhalb dessen sich neben erwartungskonform schwachen auch erwartungswidrig gute Schulen befinden, ein fallkontrastiver Zugang gewählt worden. Auf der Grundlage von VERA-8 (in Nordrhein-Westfalen: Lernstand 8) wurden acht Fallschulen identifiziert und akquiriert, die dem ungünstigen Standorttyp der Stufe 5 angehören (vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen o.J.).

In Lernstand 8 werden den Schulen in Abhängigkeit von ihrem Standort sowie ihrer Schulform Erwartungswerte hinsichtlich der Leistungen der Schülerinnen und Schüler in den Lernstanderhebungen zugeschrieben. Hierdurch ist es möglich, die Ergebnisse der einzelnen Schule mit den Schulen in Nordrhein-Westfalen insgesamt sowie mit Schulen ähnlicher Standortvoraussetzungen zu vergleichen (vgl. ebd.). Davon ausgehend wurden im Sample zwei Gruppen von Schulen in sozioökonomisch benachteiligter Lage gebildet:

- Vier Schulen, die vom Erwartungswert nicht abweichen, bilden die Gruppe der erwartungsgemäßen (leistungsschwachen) Schulen (eine Hauptschule, zwei Gesamtschulen, ein Gymnasium).
- Vier Schulen, die von diesem Wert signifikant positiv abweichen, bilden die Gruppe der erwartungswidrig guten Schulen (zwei Hauptschulen, eine Gesamtschule, ein Gymnasium).

Ein solches Sample ermöglicht einen kontrastierenden Vergleich von Merkmalen und Prozessen an Schulen, die allesamt mit ungünstigen Bedingungen starten, sich jedoch hinsichtlich ihrer Ergebnisse, dem Erfolg der Schülerinnen und Schüler bei den Lernstandserhebungen, unterscheiden. Die Datenerhebung erfolgte multimethodisch unter Einbezug mehrerer am Bildungsprozess beteiligter Akteure, um möglichst viele Perspektiven auf die einzelschulische Wirklichkeit zu erhalten.

Schuldokumentenanalyse. Mittels einer Daten- und Dokumentenanalyse wurden u.a. Schüler- und Lehrerzahlen, das Schulprogramm und Evaluationsdaten analysiert, um die (organisatorischen) Rahmenbedingungen der Schulen bzw. deren formale Benennung zu erfassen.

Fragebogenerhebung. Zudem wurde eine Fragebogenerhebung durchgeführt, in der die Lehrkräfte u.a. zu Strategien schulischer und unterrichtlicher Prozesse sowie zum Schulleitungshandeln befragt wurden. An der Befragung beteiligten sich insgesamt 198 Lehrkräfte, wobei 78 Lehrkräfte an erwartungswidrig guten Schulen unterrichteten. Die ungleiche Verteilung der Gesamtanzahl ist den unterschiedlich großen Kollegien der Schulen in den zwei Vergleichsgruppen geschuldet. Der Rücklauf an erwartungswidrig guten Schulen betrug im Durchschnitt insgesamt 42,9 Prozent, an erwartungskonform abschneidenden Schulen insgesamt 37,6 Prozent.

Im Kontext der Datennutzung kamen u.a. Instrumente zum Einsatz, die im Rahmen des SteBis-Projektes zur evidenzbasierten Schulentwicklung (EviS) entwickelt wurden (vgl. den Beitrag von van Ackeren u.a. in diesem Band). Der Bereich des Schulleitungshandelns wurde anknüpfend an vorliegende internationale Studien adaptiert bzw. selbst entwickelt. Die Daten wurden quantitativ ausgewertet. Zum Schulleitungshandeln wurden nach einer explorativen Faktorenanalyse unter weiterer Berücksichtigung theoretischer Vorannahmen Skalen gebildet, die sich als reliabel erwiesen haben (vgl. Tabelle 1).

Tab. 1: Gebildete Skalen zum Führungsstil der Schulleiterinnen und Schulleiter an Schulen in schwieriger Lage.

Skala	*Anzahl Items & Beispielitem*	*Cronbachs Alpha*
Partizipativer Führungsstil	5 Items „Wir arbeiten nach einem Konzept zur internen Kooperation.“	0,88
Strukturierender, organisationsbetonter Führungsstil	4 Items „Unsere Schulleitung fördert die Kooperation im Kollegium, in dem er/sie Ressourcen (Zeit/Räume) zur Verfügung stellt.“	0,67
Visionärer Führungsstil	3 Items „Unser Schulleiter schlägt neue Wege vor, wie Aufgaben bearbeitet werden können.“	0,84
Personaler Führungsstil	3 Items „Unser Schulleiter hilft mir, meine Stärken auszubauen.“	0,90

Quelle: eigene Berechnungen

Interviews. Weiterhin wurden semi-strukturierte Interviews mit je zwei Vertretern bzw. Vertreterinnen der Schulleitung geführt, um die Ergebnisse der schriftlichen Befragungen vertiefen und ergänzen zu können. Die Interviewleitfäden gliedern sich in folgende Bereiche:

a) Beschreibung der spezifischen Situation von Schulen in segregierter Lage;
b) Strategien der Unterrichtsgestaltung;
c) Strategien der Datennutzung zur Schul- und Unterrichtsentwicklung;
d) Strategien inner- und außerschulischer Vernetzung;
e) Strategien der Lehrkräfteführung (Motivation, Partizipation etc.) von Schulleitungen.

Die Interviews wurden inhaltsanalytisch (vgl. Mayring 2010) mit dem Programm *MAXqda* codiert und ausgewertet. Die Bildung der Kategorien ergibt sich theoriegeleitet mit Blick auf die Ergebnisse der internationalen Forschung zum Schulleitungshandeln an Schulen in herausfordernder Lage. Darüber hinaus wurde zum Teil an die bestehende systematische Klassifizierung des LISA-Projektes angeknüpft, die dann im Verlaufe der vertiefenden Analyse weiter ausdifferenziert bzw. verändert wurde (vgl. Tabelle 2).

Tab. 2: Kategorien der inhaltsanalytischen Auswertung der Schulleitungsinterviews

Kategorien
1. Datennutzung
1.1 landesweite Lernstandserhebungen
1.1.1 werden für alle Lehrkräfte dargestellt und thematisiert
1.1.2 werden mit dem Ziel der Unterrichtsentwicklung an Arbeitsgruppen delegiert
1.1.2.1 Schulleitung unterstützt den Prozess inhaltlich
1.1.2.2 Schulleitung begleitet den Prozess formal
1.1.3 werden im Schulleitungsteam besprochen
1.1.4 es werden personale Konsequenzen eingeleitet
1.1.5 es werden strukturelle unterrichtliche Rahmenbedingungen verändert bzw. geschaffen (z.B. weitere Förderstunden, Abschaffung der Differenzierung in Grund-/Erweiterungskurse)
1.2 schulbezogene Parallelarbeiten
(...) Subkategorien s.o.
1.3 innerhalb der Schule bzw. des Unterrichts eingesetzte Tests (z.B. Kompetenztests)
… weitere Datenformate vgl. Tabelle 3
2. Führungsstil
2.1 partizipativ
2.2 strukturierend, organisierend
2.3 visionär
2.4 personal
2.5 schülerorientiert

Quelle: eigene Berechnungen

Die Ergebnisse der qualitativen und quantitativen Daten wurden entsprechend den dargestellten Forschungsfragen systematisch aufeinander bezogen. Ziel ist es, durch die Ergänzung der quantitativ erhobenen Perspektive der Lehrkräfte um die qualitativ-subjektive Perspektive der Schulleitungen zu einer möglichst differenzierten Beschreibung der schulischen Situation zu gelangen, um der Komplexität des schulischen Mehrebenensystems auf den Ebenen der Einzelschule und des Unterrichts Rechnung zu tragen. Demnach kann das beschriebene Vorgehen als Methodentriangulation bezeichnet werden (vgl. Gorad/Taylor 2004).

4. Ergebnisse

Den Forschungsfragen folgend beginnt die Darstellung der Ergebnisse mit der Frage nach der Verfügbarkeit und Nutzung von Daten durch Lehrkräfte an Schulen in schwieriger Lage, die entweder erwartungskonform oder erwartungswidrig gut abgeschnitten haben. Daran anknüpfend werden die Nutzungsstrategien der Schulleitungen dargestellt, die sich aus den Schulleitungsinterviews ableiten ließen. Im zweiten Schritt wird das Schulleitungshandeln aus der Perspektive der Lehrkräfte beleuchtet, um anschließend die Sicht der Schulleitungen selbst vergleichend vorzustellen.

4.1 Verfügbarkeit und Nutzung von internen und externen Daten zur Schul- und Unterrichtsentwicklung

Ergebnisse der Daten- und Dokumentenanalyse. Die Analyse der Dokumente zeigt hinsichtlich der Datennutzung, dass alle Schulen angeben, insgesamt großen Wert auf eine datengestützte Schul- und Unterrichtsentwicklung zu legen. Die Schulen stellen dar, dass sie die Leistungsergebnisse der Schülerinnen und Schüler zum Anlass nehmen, den Fachunterricht zu reflektieren und ggf. Entwicklungsstrategien in den Fachkonferenzen einzuleiten. Darüber hinaus erheben Schulen gezielt die Lerneingangsvoraussetzungen ihrer Schülerinnen und Schüler zu Beginn der 5. Klasse durch (teil-)standardisierte Verfahren im Sinne einer Förderdiagnostik, verfolgen die Lernentwicklung kontinuierlich und leiten daraus Implikationen für organisatorische Veränderungen ab (z.B. Erhöhung der Klassenlehrerstunden, Stunden zur individuellen Förderung einzelner Schülerinnen und Schüler).

Verfügbarkeit und Nutzung interner und externer Daten durch Lehrkräfte an Schulen in schwieriger Lage

Zur Erfassung der Nutzung von Daten wurde auf Items der EviS-Studie (vgl. Demski u.a. 2012 sowie den Beitrag von van Ackeren u.a. in diesem Band) zurückgegriffen, mit denen zunächst erfragt wurde, ob entsprechende Instrumente bzw. Verfahren den Lehrkräften an der jeweiligen Schule zur Verfügung stehen, um danach vertiefend nach der unterrichtlichen Nutzung der Daten zu fragen.

Tab. 3: Verfügbare Instrumente und Daten an Schulen in segregierter Lage (N=198; Angaben in Prozent)

Die genannten Instrumente sind an meiner Schule verfügbar	**erwartungswidrig gut abschneidende Schule (n=78)**	**erwartungsgemäß abschneidende Schule (n=120)**	**gesamt (N=198)**	***gesamt EviS-Studie zum Vergleich* (N=1230)***
landesweite Lernstandserhebungen / Vergleichsarbeiten	98,66	99,06	98,9	*82*
schulbezogene Parallelarbeiten	84,5	89,1	87,2	*83,9*
innerhalb der Schule eingesetzte Tests	82,94	84,28	82	*68,3*
Schülerfeedback zum Unterricht	80,8	80,3	80,5	*82,1*
Auswertungen der Schulstatistik	78,87	87	83,6	*62,5*
Berichterstattung zu Schulleistungsvergleichen (z.B. PISA)	68,18	73,6	71,4	*56,1*
gemeinsame Unterrichtsentwicklungsmaßnahmen	68	50	57,55	*61*
Dokumentation der individuellen Lernentwicklung	55,1	44,9	48,75	-
Schulinspektion/ Qualitätsanalyse	44,9	31,68	37,05	*85,2*
selbst durchgeführte Befragungen	34,8	34,4	34,5	*44*

Anmerkung: * EviS = EviS-Studie zur Evidenzbasierten Schulentwicklung (vgl. den Beitrag von van Ackeren u.a. in diesem Band).

Quelle: eigene Berechnungen

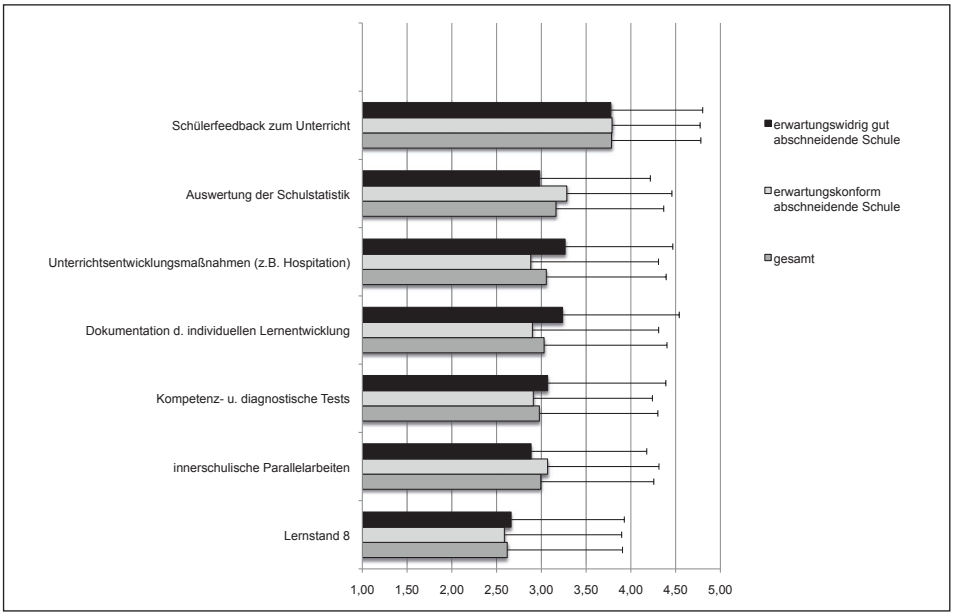
In der Gesamtschau zeigt Tabelle 2 zunächst, dass aus Sicht der Lehrkräfte die Verfügbarkeit der Daten unterschiedlich ausfällt. Während die Daten aus VERA-8/ Lernstand 8 fast allen Lehrkräften zur Verfügung stehen, sind Daten aus von der Schule selbst durchgeführten Befragungen deutlich seltener verfügbar.

Der Vergleich der für die Lehrkräfte zur Verfügung stehenden Daten mit der Erhebung der Lehrkräfte der EviS-Vergleichsgruppe (Rheinland-Pfalz) zeigt, dass die Verfügbarkeit der Daten insgesamt an Schulen in segregierter Lage in NRW aus Sicht der Lehrkräfte deutlich ausgeprägter ist. Fraglich ist jedoch, ob dieser Befund tatsächlich dem Schulstandort geschuldet ist oder ob die Verfügbarkeit von bzw. Aufmerksamkeit für Daten in NRW insgesamt ausgeprägter ist, so dass sich das Ergebnis unter Berücksichtigung des Bundeslandes relativieren würde.

Zwischen erwartungskonform und erwartungswidrig gut abschneidenden Schulen zeigen sich zum Teil deutliche Unterschiede, wobei ein signifikanter Unterschied nur hinsichtlich der Verfügbarkeit der Daten zu gemeinsamen Unterrichtsentwicklungsmaßnahmen zugunsten der erwartungswidrig guten Schulen zu konstatieren ist ($p \leq 0.05$).

In einem weiteren Schritt wurde nach der tatsächlichen Nutzung der Daten gefragt, indem die Lehrkräfte auf einer 5-stufigen Likert-Skala (1 = nie bis 5 = sehr oft) angeben konnten, in welchem Maße sie die Informationen zur Ausgestaltung und Weiterentwicklung ihrer Arbeit nutzen.

Abb. 1: Nutzung interner und externer Daten zur Unterrichtsgestaltung und Unterrichtsentwicklung durch Lehrkräfte (N=198; Mittelwerte und Standardabweichungen einer 5-stufigen Likert-Skala; höhere Werte entsprechen einer höheren Nutzung.)



Quelle: eigene Berechnungen

Es zeigt sich, dass die berichtete tatsächliche Nutzung der verschiedenen Datenformate sehr unterschiedlich ausgeprägt ist. So werden Berichte zu (inter-)nationalen Schulleistungsvergleichsstudien eher selten mit Blick auf das unterrichtliche Handeln der Lehrkräfte rezipiert. Die systematische Auswertung der Schulstatistik sowie auch Daten aus Dokumentationen zur individuellen Leistungsentwicklung der Schülerinnen und Schüler werden deutlich häufiger genutzt. Insgesamt unterscheiden sich Schulen in segregierter Lage hinsichtlich der Datennutzung damit nicht von der Datennutzung der innerhalb des EviS-Projektes befragten Lehrkräfte in Rheinland

Pfalz (vgl. Demski u.a. 2012 bzw. van Ackeren u.a. in diesem Band). Betrachtet man nunmehr erwartungskonform und erwartungswidrig gut abschneidende Schulen, zeigen sich keine statistisch signifikanten Unterschiede zwischen den beiden Gruppen. Die relativ hohe Standardabweichung verringert sich bei Betrachtung der Einzelschulebene nicht. Demnach divergiert das Nutzungsverhalten der einzelnen Lehrkräfte sehr stark, was sich vermutlich auf unterschiedliche Gründe, wie z.B. Fachzugehörigkeit, Berufserfahrung etc., zurückführen lässt.

Nutzung interner und externer Daten durch die Schulleitung zur Schul- und Unterrichtsentwicklung

Wie eingangs geschildert, ist das Handeln der Schulleitung mit Blick auf die Verfügbarmachung von Daten für die einzelne Lehrkraft bis hin zur Anleitung von datengestützten Unterrichts- und Schulentwicklungsprozessen entscheidend. Um einen tiefergehenden Einblick in die Strategien der Datennutzung aus Sicht der Schulleitungen zu erhalten, wurden diese im Rahmen der Leitfadeninterviews zu ihren Datennutzungsstrategien im Allgemeinen angesprochen bzw. vertiefend zu besonderen Datenformaten, die im Schulprogramm genannt werden.

Entgegen dem von den Lehrkräften berichteten Nutzungsverhalten wird das unterrichtliche Schülerinnen- und Schülerfeedback durch die Schulleitungen nicht näher thematisiert. Im Fokus der Schulleitungen stehen schulintern entwickelte Verfahren des Monitorings der individuellen Lernentwicklung, innerschulische Parallelarbeiten, diverse Verfahren der Kompetenzmessung nach dem Übergang von der Grundschule in die Sekundarstufe I sowie insbesondere die Daten aus Lernstand 8.

Im Folgenden wird der Fokus auf die Nutzung von Lernstand 8 aus Schulleitungssicht gerichtet, da hier unterschiedliche Nutzungsstrategien mit Blick auf ein einheitliches gängiges Datenformat verglichen werden können. Formal ist zunächst festzustellen, dass alle Schulleitungen berichten, die Ergebnisse im Rahmen von Lehrerkonferenzen zentral vorzustellen. Die weiteren Schritte unterscheiden sich jedoch deutlich. Während eine Schulleitung einer erwartungswidrig guten Schule den Ergebnissen im Vergleich zu den Beobachtungen der Lehrkräfte wenig Bedeutung schenkt und es bei der formalen Berichterstattung belässt, geben alle anderen Schulleitungen die Ergebnisse zur weiteren Bearbeitung an die Fachkonferenzen Deutsch, Mathematik und Englisch weiter. Hinsichtlich der weiteren Prozessbegleitung sind wiederum Unterschiede erkennbar; so erläutert eine Schulleitung einer *erwartungskonform abschneidenden Schule*:

> *Ja, die Auswertung von Lernstand ist also Aufgabe der Fachkonferenzen. Es geht ja um die drei zentralen Fächer Deutsch, Mathe und Englisch. In den Fachkonferenzen werden die Ergebnisse ausgewertet, und dann wird*

> *überlegt, an welchen Stellen kann verbessert werden. Was das jetzt im Einzelnen ist, kann ich nicht darstellen. Ich lege nur Wert darauf, dass das in den Fachschaften passiert, die müssen ja auch diese Ergebnisse in der Schulkonferenz vorstellen, da sind wir den Eltern ja dann auch Rechenschaft schuldig sozusagen, und das heißt also, meine Aufgabe ist es im Grunde, den Rahmen dafür zu schaffen, dass es überhaupt stattfinden kann.*

Es zeigt sich, dass die Schulleitung keine weiteren Schritte unternimmt, die Unterrichtsentwicklungsmaßnahmen, die möglicherweise in den Fachkonferenzen angestoßen werden, unmittelbar zu begleiten oder zu unterstützen. Die Schulleitung überträgt den Kolleginnen und Kollegen in den Fachkonferenzen die Verantwortung, entsprechende Entwicklungsmaßnamen daraus abzuleiten. Dies entspricht demnach dem so genannten distributiven/partizipativen Führungsstil.

Die Lernstand-Ergebnisse führen in einer anderen *erwartungsgemäß abschneidenden Schule* zu gezielten Überlegungen hinsichtlich des Personaleinsatzes:

> *Die Frage ist, sind vorher so unglückliche Lehrerkonstellationen in der Gruppe gewesen, dass die vom Lernstand so hinterherhinken, dass wir also gegebenenfalls einen sehr leistungsstarken Kollegen an der Stelle dann einsetzen, um Defizite darüber aufzufangen und nicht jetzt noch im neunten und zehnten Schuljahr dann auch einen schwachen Kollegen von der Unterrichtsverteilung rein nehmen mit dem Wissen, der wird eh nichts mehr in zwei Jahren retten können, was vier Jahre da versäumt wurde.*

Den Lernstandergebnissen wird auch mit Blick auf die weitere Lernentwicklung wesentliche Bedeutung beigemessen. Die Schulleitung sieht es als ihre Aufgabe an, dafür zu sorgen, den Schülerinnen und Schülern die bestmöglichen Leistungsergebnisse und damit auch Lebensperspektiven zu eröffnen. Zielführend erscheint dabei der gezielte Einsatz von Lehrkräften, die sich – so betont die Schulleitung im weiteren Interview – im Vergleich zu Kolleginnen und Kollegen in den Vorjahren hinsichtlich der Vorbereitung auf Lernstand 8 bereits „bewährt" haben. Dieser in gewisser Weise sanktionierende Umgang mit den Daten aus Lernstand 8 findet sich in abgeschwächter Form noch an einer weiteren erwartungskonform abschneidenden Schule.

Eine andere Strategie einer Schulleitung einer *erwartungswidrig guten Schule* wird im nachfolgenden Interviewauszug deutlich. Anzumerken ist hier, dass die Schulleitung die Ergebnisse aus Lernstand 8 nicht isoliert betrachtet, sondern mit den standardisiert erhobenen Leistungsdaten zu Beginn der Jahrgangsstufe 5 unter der Frage der Leistungsentwicklung vergleicht.

> *Also, ich habe da keine gezielten Gespräche mit Kolleginnen und Kollegen geführt, die in Klassen unterrichten, wo die Leistungen nicht so gut waren. Ich*

habe das in den Fällen, in denen es positiv war, also ich den Eindruck hatte, dass da wirklich gute Lernzuwächse stattgefunden haben, da habe ich das den Kollegen zurückgemeldet. Das ist vielleicht sehr einseitig. Aber ich gehe einfach davon aus, dass eine positive Verstärkung auf Dauer auch eine Veränderung in den Fachdiskussionen bringt, weil ich glaube, wenn ein Kollege eben eine positive Rückmeldung bekommen hat und gesehen hat, es wird auch wahrgenommen von Seiten der Schulleitung, dann wird das eher kommuniziert, als wenn ich jetzt sagen würde, dein Ergebnis ist ganz schlecht und du musst da was dran verändern.

Diese beschriebene positive Feedbackkultur führt aus Sicht der Schulleitung im weiteren Verlauf dazu, dass über erfolgreiche Unterrichtskonzepte kommuniziert wird und diese eher in die Breite getragen werden können. Um dies zu unterstützen, hat die Schulleitung eine Internetplattform etabliert, um besonders innovative Konzepte verfügbar zu machen, diese aber auch im Rahmen der Lehrerkonferenzen im Sinne einer Anerkennungskultur zu würdigen.

An dieser Stelle, aber auch insgesamt wird für die Gruppe der erwartungswidrig guten Schulen ein hohes persönliches inhaltliches Interesse von Seiten der Schulleitungen an fachlichen und/oder methodischen Unterrichtskonzepten deutlich.

Insgesamt offenbaren die Interviews unterschiedliche Strategien des Umgangs mit Daten bzw. auch die unterschiedliche Bedeutung, die externen und internen Daten zugemessen wird. Gleichwohl ließ sich keine für alle erwartungswidrig guten Schulen einheitliche Strategie rekonstruieren, wobei tendenziell ein Lehrkräfte unterstützendes Leitungshandeln abzuleiten ist. Das allgemeine Schulleitungshandeln an den untersuchten Schulen wird nachfolgend noch einmal in breiterer Perspektive in den Blick genommen.

4.2 Schulleitungshandeln an Schulen in schwieriger Lage

Anknüpfend an empirische Befunde der internationalen Forschung, die, wie dargestellt, das Schulleitungshandeln als ganz wesentlichen Qualitätsfaktor für gute Schulen in schwieriger Lage immer wieder herausgearbeitet hat, wurden die Lehrkräfte der beteiligten Schulen nach ihrer Einschätzung zur Bedeutung der Schulleitung befragt. Diese Frage ist auch vor dem Hintergrund unterschiedlicher bildungspolitischer Rahmenbedingungen, die Schulleitungen im internationalen Vergleich mit divergierender Entscheidungsmacht ausstatten, relevant.

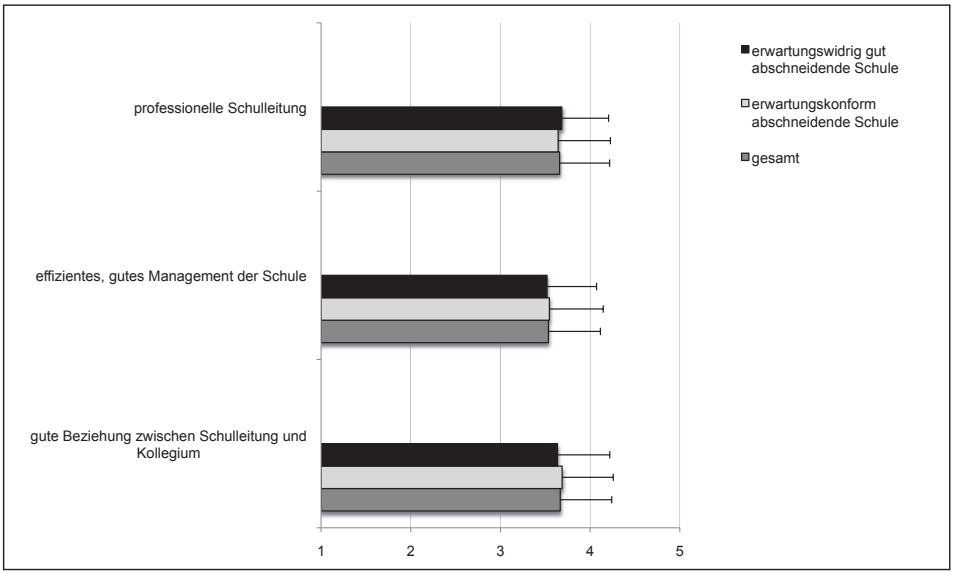
Die befragten Lehrkräfte konnten hierzu auf einer Skala von 1 (völlig unwichtig) bis 4 (sehr wichtig) die grundsätzliche Bedeutung des Schulleitungshandelns für erfolgreiche Schulen in segregierter Lage einschätzen. Es wird deutlich, dass die Bedeutung

des Schulleitungshandelns aus Sicht der befragten Lehrkräfte sehr hoch eingeschätzt wird. Ein Unterschied zwischen den befragten Lehrkräften erwartungskonform bzw. erwartungswidrig gut abschneidender Schulen ist nicht festzustellen.

Einschätzung des Führungsstils der Schulleitung aus Sicht der Lehrkräfte

Im nächsten Schritt wurden unterschiedliche Facetten des Schulleitungshandelns, die international als wesentliche Eigenschaften für Schulleitungen an erfolgreichen Schulen beschrieben wurden, erfragt. Die Lehrkräfte hatten hier die Möglichkeit, auf einer Skala von 1 bis 4 (Items zum partizipativen, strukturierenden und organisationsbetonten Führungsstil) und auf einer Skala von 1 bis 5 (Items zum visionären und personalen Führungsstil) anzugeben, inwieweit die entsprechenden Handlungsweisen ausgeprägt sind.

Abb. 2: Bedeutung des Schulleitungshandelns an Schulen in segregierter Lage aus Sicht der Lehrkräfte (N=198; Mittelwerte und Standardabweichungen; vierstufige Likert-Skala; höhere Werte entsprechen einer höheren Zustimmung.)



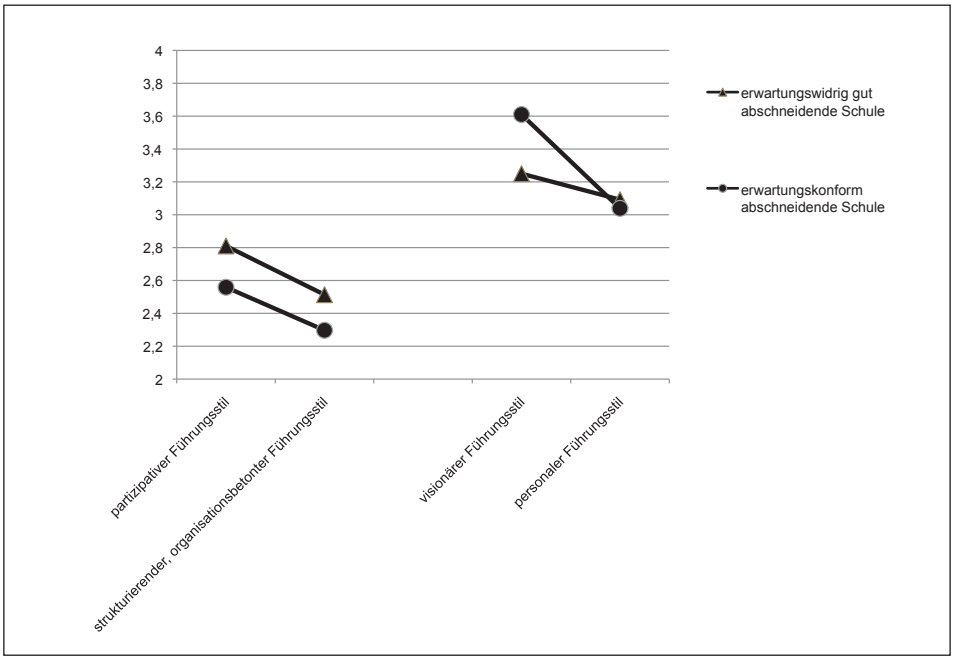
Quelle: eigene Berechnungen

Abbildung 3 veranschaulicht Unterschiede hinsichtlich der Ausprägungen der Führungsstile in den Angaben von Lehrkräften an erwartungskonform und erwartungswidrig gut abschneidenden Schulen. So sind sowohl Handlungsmerkmale des partizipativen als auch des organisationsorientierten Führungsstils an erwartungswidrig gut abschneidenden Schulen signifikant höher ausgeprägt ($p \leq 0.5$). Hinsichtlich ih-

rer personalen Unterstützung unterscheiden sich die Schulleiterinnen und Schulleiter aus Sicht der Lehrkräfte nicht voneinander.

Überraschend ist das Ergebnis zum visionären Führungsstil, da hier Lehrkräfte an den befragten erwartungskonform abschneidenden Schulen die Schulleitungen als visionärer wahrnehmen als an erwartungswidrig gut abschneidenden Schulen ($p \leq 0.5$). Es ist zu vermuten, dass sich Schulleitungen an erwartungswidrig gut abschneidenden Schulen eher in der Phase der Konsolidierung befinden, da sie das Ziel der guten Schülerinnen- und Schülerleistungen bereits erreicht haben. Erwartungskonform abschneidende Schulen hingegen sind möglicherweise aktuell mit der Entwicklung von Schul- und Unterrichtsentwicklungsmaßnahmen befasst, um die Leistungen ihrer Schülerinnen und Schüler zu verbessern.

Abb. 3: Durchschnittliche Ausprägung der wahrgenommenen Führungsstile an erwartungskonform bzw. erwartungswidrig gut abschneidenden Schulen in schwieriger Lage im Vergleich (N=198; Mittelwerte einer vier- bis fünfstufigen Likert-Skala. Höhere Werte entsprechen einer höheren Zustimmung.)



Quelle: eigene Berechnungen

Praktizierter Führungsstil aus Sicht der Schulleitung

Insgesamt beschreiben alle Schulleitungen ein transparentes Vorgehen, das darüber hinaus darauf abzielt, Verantwortlichkeiten gezielt an Subgruppen, wie z.B. die Fachgruppen, abzugeben. Dieses Vorgehen wurde bereits im Umgang mit Lernstand 8 sichtbar. Darüber hinaus wird aber bei allen erwartungswidrig guten Schulen deutlich, dass zentrale Ideen und Veränderungen von den Schulleitungen ausgehen. So berichtet zum Beispiel eine Schulleitung einer *erwartungswidrig gut abschneidenden Schule*:

> *Also, ich sag mal ganz klar, die Idee, eine vertikale Abteilungsstruktur zu schaffen, habe ich in die Schule gebracht. Die Vorstellung war bei mir im Kopf, dass man so eine große Schule auch überschaubarer machen kann. Und das habe ich dann irgendwann in die Schulleitung mit eingebracht als Idee und die Begeisterung bei den anderen sechs Schulleitungsmitgliedern war sehr groß. Wir waren auf einem sehr, sehr schwierigen Stand damals, rückläufige Anmeldezahlen, schlechte Abschlussergebnisse, ständig Disziplinprobleme und der Alltag war also wirklich sehr, sehr anstrengend, und dann haben wir gesagt, wir versuchen das mit dieser vertikalen Struktur. Das ist ein Schulentwicklungsprozess gewesen der nicht ganz einfach war, weil ich auch von Schulleitungsseite aus eklatante Fehler gemacht habe, was Beteiligungen und so was anbelangt. Da mussten wir wieder zurückrudern und erst mal sehen, dass wir das Kollegium wirklich mit ins Boot nehmen.*

Dieser Interviewauszug verdeutlicht unterschiedliche Aspekte. Zum einen bestätigt er die aus der internationalen Forschung stammende Annahme, nach der Schulleitungen letztlich die Funktion der *Change Agents* zukommt. Aus Sicht der Befragten befand sich die Schule damals in einer schwierigen Lage; möglicherweise wäre sie damals als erwartungskonform schwach abschneidende oder sogar als *Failing School* mit deutlich unterdurchschnittlichen Leistungen charakterisiert worden. Die Schulleitung hatte in dieser Situation die Idee, die Schule vertikal in Abteilungen aufzuteilen, um so z.B. die Lehrer-Schüler-Beziehungen, aber auch die Teamarbeit unter den Kolleginnen und Kollegen weiter intensivieren zu können.

In Kooperation mit dem restlichen Schulleitungsteam ist es ihr zwar gelungen, diese Idee zu etablieren; sie scheiterte aber auch insofern zunächst, als sie den Rest des Kollegiums nicht in ihre Entscheidungen einbezogen hatte. Vor dem Hintergrund dieser Erfahrung wird der insgesamt als visionär zu bezeichnende Führungsstil um Komponenten der Partizipation ergänzt, die im weiteren Verlauf des Interviews deutlich werden.

Schulleitungen erwartungskonform abschneidender Schulen benennen zum Teil unterschiedliche Handlungsziele und Schwerpunkte, die sie als entscheidend für die

Unterrichtsentwicklung ansehen. Hierzu gehört, wie oben geschildert, nicht zuletzt eine gezielte Nutzung verschiedener Datenformate. Dieses Ergebnis entspricht den Angaben der Lehrkräfte, die den Führungsstil der Schulleitungen erwartungskonform abschneidender Schulen als visionär einschätzen.

Gleichwohl wird in den Interviews deutlich, dass eine entscheidende Hürde in der Motivation und Motivierung der Lehrerinnen und Lehrer gesehen wird, diesen Weg mitzugehen bzw. auch einen erhöhten zeitlichen Mehraufwand zu leisten. Als ursächlich für die fehlende Motivation werden die erschwerten schulischen Rahmenbedingungen angesehen, die bildungspolitisch nicht durch eine entsprechende Ressourcenallokation – so die Wahrnehmung – kompensiert werden.

Obgleich der strukturierende und organisationsbetonte Führungsstil aus Sicht der Lehrkräfte weniger stark ausgeprägt ist, werden diese Handlungsmerkmale insbesondere von den Schulleitungen erwartungswidrig guter Schulen als entscheidende Merkmale benannt.

> *Ich habe immer Wert darauf gelegt, egal wie die Zahlen als Parameter vorgegeben waren, wir haben immer Kleinstgruppen, eigentlich immer. Ich habe immer gesagt, auch wenn das bei der Schulaufsicht nicht als glücklich aufgenommen wurde, nur die Erfolge haben mir das bestätigt, wir fahren lieber kürzer unsere Stundenzahl auf 28 und nicht auf 32, um dadurch Raum zu gewinnen für weitere Differenzierungsmöglichkeiten und das ist A und O. Das ist alles, also Schule kann effizient sein, wenn man die Schüler-Lehrerrelation verändert.*

Dieses Schulleitungsmitglied einer *erwartungswidrig guten Schule* beschreibt, dass die organisationalen Rahmenbedingungen gegen die Vorgaben verändert wurden, um aus seiner Sicht optimale Lehr-Lernbedingungen herzustellen. Dieser Fokus ist auch im oben dargestellten Interviewausschnitt sichtbar, der als Beispiel einer visionären Handlungsstrategie dient.

Darüber hinaus wird von Schulleitungen erwartungswidrig guter Schulen ein weiterer Aspekt als wesentlich erachtet. Dabei handelt es sich um ein persönliches Verhältnis zu den Schülerinnen und Schüler, wie der nachfolgende Auszug verdeutlicht:

> *Also ich denke mir, ich bin auch als Schulleitung ganz nah an den Kindern dran. Das heißt, ich kenne jeden mit Namen, und ich sorge auch dafür, dass ich im Unterricht, also dass ich nicht zu weit weg bin wirklich von der pädagogischen Front. Dass ich die Kinder auch kenne, persönlich kenne, dass ich mich beim Mittagessen mit da hinsetze, dass die bei mir am Tisch sitzen, und wir unterhalten uns dann über dies und jenes.*

Dieses persönliche Verhältnis wird von Schulleitungen erwartungsgemäß abschneidender Schulen nicht in dieser Weise benannt. Von zwei Schulleitungen wird eine Distanz sowohl zu den Lehrkräften als auch zu den Schülerinnen und Schüler berichtet, die jedoch auch nicht als produktiv wahrgenommen wird. Vielmehr scheint es sich um eine Distanz zu handeln, die mit der Rolle der Schulleitung einhergeht. Als ursächlich hierfür kann die Gesamtgröße der Schule angenommen werden, die es Schulleitungen mitunter nur eingeschränkt erlaubt, einen persönlichen Kontakt zu den Schülerinnen und Schüler zu wahren. Hier kann die Schaffung einer vertikalen Struktur, die die gesamte Schule in Abteilungen von Klasse 1-10 unterteilt, den persönlichen Kontakt zwischen Lehrkräften und Schülerinnen und Schülern intensivieren; für die Schulleitung gilt dies jedoch nicht, da sie für alle Abteilungen gleichermaßen verantwortlich ist.

5. Zusammenfassung der Ergebnisse und Diskussion

Die Ergebnisse der vorliegenden Untersuchung zeigen zunächst, dass die Nutzung von internen und externen Daten Einzug in Schulprogramme von Schulen in schwieriger Lage gefunden hat, was darauf hindeutet, dass allen Schulleitungen bewusst ist, dass die Nutzung evidenzbasierter Informationen grundsätzlich ein Ansatz für systematische Qualitätsentwicklung sein kann, da sich alle Schulen in Konkurrenzsituation zu anderen Schulen befinden. Vor diesem Hintergrund ergibt sich zunächst kein Unterschied zwischen erwartungswidrig gut bzw. erwartungskonform abschneidenden Schulen.

Betrachtet man nun die Verfügbarkeit auf der einen und die tatsächliche Nutzung unterschiedlicher Datenformate auf der anderen Seite auf der Ebene der Lehrkräfte, wird deutlich, dass verschiedenste Datenquellen zwar verfügbar sind, aber zur Unterrichtsentwicklung kaum genutzt werden. Auf der Ebene der Schulleitungen werden z.B. die Ergebnisse aus Lernstand 8 zumeist unter der Perspektive von Unterrichtsentwicklungsstrategien rezipiert und zur vertiefenden inhaltlichen Bearbeitung an die entsprechenden Fachkonferenzen weitergeleitet. Gleichwohl ergeben sich im Anschluss daran sehr divergierende Umgangsstrategien. Unterscheiden sich die Lernergebnisse der Schülerinnen und Schüler zwischen den einzelnen Klassen, führt dies bei einer befragten Schulleitung einer erwartungskonform abschneidenden Schule dazu, den Lehrkräfteeinsatz zu verändern. So soll der Einsatz „leistungsstärkerer“ Lehrkräfte die Schülerleistungen bis zu den zentralen Abschlussprüfungen deutlich verbessern, so dass sich keine gravierenden innerschulischen Unterschiede mehr zeigen. Eine andere Schulleitung geht implizit davon aus, dass Lehrkräfte von Klassen, die deutlich besser abschneiden als ihre Parallelklassen, besonders gute Unterrichtskonzepte implementieren konnten. Im Sinne innerschulischer Good-

Practice-Beispiele versucht sie, diese Verfahren für alle Lehrkräfte verfügbar zu machen und breiter zu etablieren.

Eine vergleichende Analyse erwartungskonform und erwartungswidrig guter Schulen zeigt hinsichtlich der Datennutzung, dass Schulleitungen erwartungswidrig guter Schulen ein deutliches fachliches Interesse an einer datengestützten Unterrichtsentwicklung haben. Schulleitungen an erwartungskonform abschneidenden Schulen initiieren die Auseinandersetzung mit den Daten aus Lernstand 8 formal in den Fachkonferenzen und beobachten sie bzw. leiten personale Konsequenzen bei Lehrkräften ein, die vermeintlich schlechtere Ergebnisse bei den Schülerinnen und Schülern erzielen. Ohne die systematische Berücksichtigung weiterer Daten, z.B. konkreter Unterrichtsbeobachtungen, erscheint letzteres Vorgehen jedoch problematisch und das Qualitätsverständnis verkürzt. Im Vergleich dazu deutet sich in den Studien von Hanushek und Raymond (2004) sowie West, Ainscow und Stanford (2005) an erwartungswidrig guten Schulen an, dass eine professionelle, kooperative Zusammenarbeit zwischen Schulleitung und Lehrkräften zu einer produktiven, kontextabhängigen Schul- und Unterrichtsentwicklung beitragen kann.

Zusammenfassend fällt auf, dass Schulleitungen ausgehend von den Lernstand-8-Daten Rückschlüsse auf die Unterrichtsqualität ziehen. Ein solches Vorgehen sollte jedoch ohne Berücksichtigung weiterer Kontextmerkmale wie z.B. der Klassenzusammensetzung kritisch hinterfragt werden (vgl. hierzu auch Kuper/Diemer 2012). So wäre ein gezieltes Trainieren der so genannten *Test-Wiseness* (Vertrautheit mit der Bearbeitung von Tests, eine geschickte Zeiteinteilung, Rate-Strategien usw.) der Schülerinnen und Schüler sicherlich für das Abschneiden in Lernstand 8 förderlich; dies ließe jedoch keine Aussage über die Unterrichtsprozessqualität zu.

Insgesamt ist es aufgrund des gewählten Fallstudiendesigns nicht möglich, an dieser Stelle allgemeingültige gelungene bzw. weniger gelungene Datennutzungsstrategien mit Blick auf die Leistungsentwicklung der Schülerinnen und Schüler abzuleiten. Gleichwohl können diese Ergebnisse dazu dienen, im Rahmen quantitativ vergleichender Erhebungen an Schulen in unterschiedlichen Kontexten geprüft zu werden.

Dies betrifft auch das allgemeine Schulleitungshandeln bzw. die Führungsstile. Innerhalb dieser Studie mit Fallstudiendesign wurde zunächst die Bedeutung des Schulleitungshandelns aus Sicht der Lehrkräfte an Schulen in weniger begünstigter Lage erfragt. Hier zeigte sich, dass die Lehrkräfte dem Handeln der Schulleitung eine hohe Bedeutung beimessen, was den vorliegenden Ergebnissen der internationalen Forschung entspricht. Darauf folgend wurden verschiedene Handlungsmuster erfasst, die unterschiedlichen Führungsstilen zugeordnet werden konnten. Aus Sicht der befragten Lehrkräfte kommt dem visionären und personalen Führungsstil von Schulleitungen an Schulen in schwieriger Lage eine hohe Bedeutung zu. Auch dieses Ergebnis entspricht der internationalen Forschung, die jedoch darüber hinaus die

Partizipation der Lehrkräfte an schulischen Entscheidungsprozessen als wesentlich benennt.

In der vorliegenden Untersuchung werden hinsichtlich der praktizierten Führungsstile Unterschiede zwischen erwartungskonform und erwartungswidrig gut abschneidenden Schulen sichtbar, die sich einerseits in signifikanten Unterschieden in der Umsetzung des partizipativen und organisationsbetonten Führungsstils zugunsten der erwartungswidrig guten Schulen widerspiegeln und andererseits in einem ebenfalls signifikanten Unterschied in der Umsetzung des visionären Führungsstils zugunsten der erwartungskonform abschneidenden Schulen. Der erste Befund entspricht den dargelegten internationalen Befunden zum Schulleitungshandeln an so genannten *Improving Schools in Challenging Circumstances*. Im Rahmen der Schulleitungsinterviews wurde der vorliegende Befund insofern bestärkt, als die befragten Schulleitungen erwartungswidrig guter Schulen insbesondere sowohl die Partizipation als auch die Veränderung organisatorischer, unterrichtsbezogener Rahmenbedingungen zur Optimierung der Schülerleistungen hervorheben. Dieses Resultat entspricht den vorliegenden Forschungsbefunden zur Bedeutung eines partizipativen Führungsstils sowie zu einem starken Fokus der Schulleitung auf die Unterrichtsqualität. Ein entsprechendes die Lehrkräfte unterstützendes Verhalten, das insgesamt einem personalen Führungsstil zugeordnet werden kann, wurde auch im Rahmen der Nutzung von Daten von den befragten Schulleitungen berichtet. Insgesamt zeigt sich in den Schulleitungsinterviews mit Schulleitungen erwartungswidrig guter Schulen gleichwohl ein visionärer Führungsstil, der jedoch nicht nur auf die gegenwärtige Situation bezogen ist, sondern rückblickend auf verschiedene Schritte der Schul- und Unterrichtsentwicklung in einer Umbruchphase rekurriert. Über die erfragten Führungsstile hinaus erachten Schulleitungen erwartungswidrig guter Schulen eine gute Schulleiter-Schülerbeziehung als entscheidend. Die Realisierung einer solchen Beziehung gelingt jedoch insbesondere Schulleitungen an Schulen mit geringen Schülerzahlen.

Insgesamt deuten die Schulleitungsinterviews an, dass je nach individuellen Kontextbedingungen (schulische Wettbewerbssituation, Lernleistungsentwicklung der Schülerinnen und Schüler, Möglichkeiten der Netzwerkbildung, Schulgröße) eine schulspezifische Melange von Führungsstilfacetten hinsichtlich der Schülerleistungen erfolgreich zu sein scheint. Demnach erweist es sich als günstig, wenn eine Schulleitung, die als *Change Agent* wirken will, in der Lage ist, die Ausgangssituation mit Hilfe unterschiedlicher Daten zu analysieren, um dann reflektiert auf unterschiedliche Führungsstrategien zurückgreifen zu können, die sich wiederum im Verlauf des Schulentwicklungsprozesses adaptiv verändern können.

Literatur und Internetquellen

Ackeren, I. van. (2008): Schulentwicklung in benachteiligten Regionen – Eine exemplarische Bestandsaufnahme von Forschungsbefunden und Steuerungsstrategien. In: Lohfeld, W. (Hrsg.): Gute Schulen in schlechter Gesellschaft. Wiesbaden: VS, S. 47-58.

Bauer, K. ([2]2002): Lehrerinteraktion und -kooperation. In: Helsper, W./Böhme, J. (Hrsg.): Handbuch der Schulforschung. Wiesbaden: VS, S. 813-831.

Baumert, J./Stanat, P./Watermann, R. (2006): Herkunftsbedingte Disparitäten im Bildungswesen. Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS.

Berkemeyer, N./Bos, W./Manitius, V./Müthing, K. (Hrsg.) (2008): Unterrichtsentwicklung in Netzwerken. Konzeptionen, Befunde, Perspektiven. Münster u.a.: Waxmann.

Bonsen, M. (2010): Schulleitungshandeln. In: Altrichter, H./Maag Merki, K. (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS, S. 277-295.

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. Hrsg. vom Bundesministerium für Bildung und Forschung, Referat Bildungsforschung. Berlin: BMBF.

Brauckmann, S. (2012): Schulleitungshandeln zwischen deconcentration, devolution und delegation (3D). In: Empirische Pädagogik 26, H. 1, S. 78-102.

Burchardt, S. (2008): Schulen im Sozialraum. Strukturelle Grenzen und institutionelle Chancen im lokalen Kontext. In: Lohfeld, W. (Hrsg.): Gute Schulen in schlechter Gesellschaft. Wiesbaden: VS, S. 59-76.

Cavanagh, R.F. (1997): The Culture and Improvement of Western Australian Senior Secondary Schools. Dissertation, Curtin University of Technology.

Dedering, K. (2011): Hat Feedback eine positive Wirkung? Zur Verarbeitung extern erhobener Leistungsdaten in Schulen. In: Unterrichtswissenschaft 39, H. 1, S. 61-81.

Demski, D./Rosenbusch, C./Ackeren, I. van/Clausen, M./Schmidt, U. (2012): Steuerung von Schule durch evidenzbasierte Einsicht? Konzeption und erste Befunde des Forschungsverbundes EviS. In: Hornberg, S./Parreira do Amaral, M. (Hrsg.): Deregulierung im Bildungswesen. Münster u.a.: Waxmann, S. 131-150.

Diemer, T./Kuper, H. (2012): Formen innerschulischer Steuerung mittels zentraler Lernstandserhebungen. In: Zeitschrift für Pädagogik 57, H. 4, S. 554-571.

Flintham, A. (2006): What's Good about Leading Schools in Challenging Circumstances? Nottingham: National College for Leadership of Schools and Children's Services. URL: http://dera.ioe.ac.uk/6968/1/leading-schools-in-challenging-circumstances.pdf; Zugriffsdatum: 20.01.2013.

Freeman, J.A. (1997): Contextual Contrasts between Improving and Stable Elementary Schools in Louisiana. Paper Presented at the Annual Conference of the Mid-South Educational Research Association, Memphis, TN.

Fussangel, K./Gräsel, C. (2012): Lehrerkooperation aus der Sicht der Bildungsforschung. In: Baum, E./Idel, T.-S./Ullrich, H. (Hrsg.): Kollegialität und Kooperation in der Schule. Theoretische Konzepte und empirische Befunde. Wiesbaden: Springer VS, S. 29-40.

Gorad, S./Taylor, C. (2004): Combining Methods in Educational Research. Maidenhead: Open University Press.

Groß Ophoff, J./Hosenfeld, I./Koch, U. (2007): Formen der Ergebnisrezeption und damit verbundene Schul- und Unterrichtsentwicklung. In: Hosenfeld, I./Groß Ophoff, J. (Hrsg.): Nutzung und Nutzen von Evaluationsstudien in Schule und Unterricht. Landau: Verlag Empirische Pädagogik, S. 411-427.

Hanushek, E.A./Raymond, M.F. (2004): Does School Accountability Lead to Improved Student Performance? In: NBER Working Paper No. 10591. URL: http://www.nber.org/papers/w10591; Zugriffsdatum: 21.07.2013.

Harris, A. (2004): Teacher Leadership and Distributed Leadership: An exploration of the literature. In: Leading and Managing 10, H. 2, S. 1-10.

Harris, A./Chapman, C./Muijs, D./Russ, J./Stoll, L. (2006): Improving Schools in Challenging Contexts: Exploring the Possible. In: School Effectiveness and School Improvement 17, H. 4, S. 409-424.

Hartung-Beck, V. (2009): Schulische Organisationsentwicklung und Professionalisierung. Folgen von Lernstandserhebungen an Gesamtschulen. Wiesbaden: VS.

Henchey, N. (2001): Schools that Make a Difference: Final Report. Twelve Canadian Secondary Schools in Low-Income Settings. Kelowna, BC: Society for the Advancement of Excellence in Education.

Hermann, J. (2013): Erfahrungen aus einer Unterstützungsmaßnahme für „Schulen in schwieriger Lage" in Hamburg: In: Quesel, C./Husfeldt, V./Landwehr, N./Steiner, P. (Hrsg.): Failing Schools. Herausforderungen für die Schulentwicklung. Bern: hep, S. 77-94.

Hopkins, D. (2001): Meeting the Challenge. An Improvement Guide for Schools Facing Challenging Circumstances. London: Department for Education and Skills.

Hopkins, D./Reynolds, D. (2002): The Past, Present and Future of School Improvement. London: Department for Education and Skills.

Horstkemper, M./Tillmann, K.-J. (2008): Schulformvergleiche und Studien zu Einzelschulen. In: Helsper, W./Böhme, J. (Hrsg.): Handbuch der Schulforschung. Wiesbaden: VS, S. 285-320.

Huber, S.G./Ahlgrimm, F./Hader-Popp, S. (2012): Kooperation in und zwischen Schulen sowie mit anderen Bildungseinrichtungen: Aktuelle Diskussionsstränge, Wirkungen und Gelingensbedingungen. In: Huber, S.G./Ahlgrimm, F. (Hrsg.): Aktuelle Forschung zur Kooperation in und zwischen Schulen sowie mit anderen Partnern. Münster u.a.: Waxmann, S. 323-372.

Huber, S.G./Muijs, D. (2010): School Leadership Effectiveness: The Growing Insight in the Importance of School Leadership for the Quality and Development of Schools and Their Pupils. In: Huber, S.G. (Hrsg.): School Leadership – International Perspectives. Dordrecht: Springer Science+Business Media, S. 57-77.

Keys, W./Sharp, C./Greene, K./Grayson, H. (2003): Successful Leadership in Urban Contexts. Full Report. Nottingham: National College for School Leadership. URL: http://dera.ioe.ac.uk/5129/1/media-7ab-b0-successful-leadership-in-urban-contexts.pdf; Zugriffsdatum: 20.01.2013.

Kühle, B./Peek, R. (2007): Lernstandserhebungen in Nordrhein-Westfalen. Evaluationsbefunde zur Rezeption und zum Umgang mit Ergebnisrückmeldungen in Schulen. In: Empirische Pädagogik 21, H. 4, S. 428-447.

Kühn, S.M./Racherbäumer, K. (2011): Von Daten zu Taten. Oder: Wie können Schulen die Ergebnisse internationaler Schulleistungsstudien für sich nutzbar machen? In: schulmanagement 32, H. 4, S. 22-24.

Kullmann, H. (2010): Lehrerkooperation – Ausprägung und Wirkungen am Beispiel des naturwissenschaftlichen Unterrichts an Gymnasien. Münster u.a.: Waxmann.

Ledoux, G./Overmaat, M. (2001): Op zoek naar succes: Een onderzoek naar basisscholen die meer en minder succesvol zijn voor autochtone en allochtone leerlingen uit achterstandsgroepen. Amsterdam: SCO-Kohnstamm Instituut.

Leithwood, K./Harris, A./Strauss, T. (2010): Leading School Turnaround. San Francisco, CA: Jossey-Bass.

Leithwood, K./Steinbach, R. (2002): Successful Leadership for Especially Challenging Schools. In: Journal of Leadership in Education 79, H. 2, S. 73-82.

Maaz, K./Baumert, J. (2011): Risikoschüler in Deutschland. Identifikation und Beschreibung. SchulVerwaltung 22, S. 56-59.

Maier, U. (2008): Rezeption und Nutzung von Vergleichsarbeiten – Ergebnisse einer Lehrerbefragung in Baden-Württemberg. In: Zeitschrift für Pädagogik 54, H. 1, S. 95-117.

Maier, U./Kuper, H. (2012): Vergleichsarbeiten als Instrument der Qualitätsentwicklung an Schulen. Überblick zum Forschungsstand. In: Die Deutsche Schule 104, H. 1, S. 88-99.

Mayring, P. ([11]2010): Qualitative Inhaltsanalyse. Grundlagen und Techniken. Aktualisierte und überarb. Aufl. Weinheim: Beltz.

McDougall, D./Gaskell, J./Flessa, J./Kugler, J./Jang, E.E./Herbert, M./Pollon, D./Russell, P./Fantilli, R. (2006): Improving Student Achievement in Schools Facing Challenging Circumstances: Draft Report. Toronto: Ontario Institute for Studies in Education. URL: http://www.curriculum.org/secretariat/files/Sept28Summary.pdf; Zugriffsdatum: 20.01.2013.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (o.J.): Beschreibung der Standorttypen. URL: http://www.standardsicherung.schulministerium.nrw.de/lernstand8/standorttypenkonzept/standorttypenbeschreibung/; Zugriffsdatum: 20.01.2013.

Muijs, D./Harris, A./Chapman, C./Stoll, L./Russ, J. (2004): Improving Schools in Socioeconomically Disadvantaged Areas. A Review of the Research Evidence. In: School Effectiveness and School Improvement 15, H. 2, S. 149-175.

Nachtigall, C./Jantowski, A. (2007): Die Thüringer Kompetenztests unter besonderer Berücksichtigung der Evaluationsergebnisse zum Rezeptionsverhalten. In: Empirische Pädagogik 21, H. 4, S. 401-410.

Pont, B./Nusche, D./Moorman, H. (2008): Improving School Leadership. Policy and Practice. Paris: OECD.

Quesel, C./Husfeldt, V. (2013): „Failing Schools" im Kontext marktorientierter Bildungspolitik. Eine Analyse zur Reformdynamik in England und den USA. In: Quesel, C./Husfeldt, V./Landwehr, N./Steiner, P. (Hrsg.): Failing Schools. Herausforderungen für die Schulentwicklung. Bern: hep, S. 9-22.

Racherbäumer, K./Funke, C./van Ackeren, I. (2013): Forschungsfeld „Schulen in schwieriger Lage". Erste Befunde eines Projekts in der Metropole Ruhr. In: Schulverwaltung NRW 24, H. 1, S. 24-26.

Reynolds, D./Hopkins, D./Potter, D./Chapman, C. (2001): School Improvement for Schools Facing Challenging Circumstances: A Review of Research and Practice. London: Department for Education and Skills.

Rolff, H.-G. (2010): Datengestützte Schulentwicklung. In: Rolff, H.-G. (Hrsg.): Führung, Steuerung, Management. Seelze: Friedrich Verlag, S. 133-176.

Rolff, H.-G. (2013): Schulentwicklung kompakt. Modelle, Instrumente, Perspektiven. Weinheim/Basel: Beltz.

Scheerens, J./Bosker, R.J. (1997): The Foundations of Educational Effectiveness. Oxford: Pergamon.

Schrader, F.W./Helmke, A. (2003): Evaluation – und was danach? Ergebnisse der Schulleiterbefragung im Rahmen der Rezeptionsstudie WALZER. In: Schweizerische Zeitschrift für Bildungswissenschaften 25, H. 1, S. 79-110.

Schratz, M./Hartmann, M./Schley, V. (2010): Schule wirksam leiten: Analyse innovativer Führung in der Praxis. Münster u.a.: Waxmann.

Schulze, F. (2012): Folgen zentraler Lernstandserhebungen. In: Die Deutsche Schule 104, H. 1, S. 43-57.

Seeley, D.S./Niemeyer, J.S./Greenspan, R. (1990): Principals Speak: Improving Inner-City Elementary Schools. Report on Interviews with 25 New York City Principals. New York, NY: City University of New York.
Solzbacher, C./Minderhop, D. (2007): Bildungsnetzwerke und Regionale Bildungslandschaften. Ziele und Konzepte, Aufgaben und Prozess. Neuwied: LinkLuchterhand.
Steinert, B./Klieme, E./Maag Merki, K./Döbrich, P./Halbheer, U./Kunz, A. (2006): Lehrerkooperation in der Schule. In: Zeitschrift für Pädagogik 52, H. 2, S. 185-203.
Stoll, L./Fink, D. (1996): Changing Our Schools: Linking School Effectiveness and School Improvement. Buckingham u.a.: Open University Press.
Teddlie, C./Stringfield, S./Wimpelberg, R./Kirby, P. (1989): Contextual Differences in Models for Effective Schooling in the USA. Paper Presented at the International Congress for School Effectiveness and Improvement, Rotterdam, The Netherlands.
Tippelt, R./Reupold, A./Strobel, C./Kuwan, H. (Hrsg.) (2009): Lernende Regionen – Netzwerke gestalten. Teilergebnisse zur Evaluation des Programms „Lernende Regionen – Förderung von Netzwerken". Bielefeld: WBV.
West, M. (2012): Improving Student Outcomes in Challenging Contexts by School-to-School Cooperation. In: Huber, S.G./Ahlgrimm, F. (Hrsg.): Aktuelle Forschung zur Kooperation in und zwischen Schulen sowie mit anderen Partnern. Münster u.a.: Waxmann, S. 208-224.
West, M./Ainscow, M./Stanford, J. (2005): Sustaining Improvement in Schools in Challenging Circumstances: A Study of Successful Practice. In: School Leadership and Management 25, H. 1, S. 77-93.
Wissinger, J. (2011): Schulleitung und Schulleitungshandeln. In: Terhart, E./Bennewitz, H./Rothland, M. (Hrsg.): Handbuch der Forschung zum Lehrerberuf. Münster u.a.: Waxmann, S. 98-115.

Kathrin Racherbäumer, Dr. phil., geb. 1978, wissenschaftliche Mitarbeiterin an der Universität Duisburg-Essen.
E-Mail: kathrin.racherbaeumer@uni-due.de

Christina Funke, geb. 1986, wissenschaftliche Mitarbeiterin an der Universität Duisburg-Essen.
E-Mail: christina.funke@uni-due.de

Isabell van Ackeren, Prof. Dr. phil., geb. 1974, Professorin für Bildungssystem- und Schulentwicklungsforschung an der Universität Duisburg-Essen.
E-Mail: isabell.van-ackeren@uni-due.de

Marten Clausen, Prof. Dr. phil., geb. 1968, Professor für Unterrichtsforschung an der Universität Duisburg-Essen.
E-Mail: marten.clausen@uni-due.de

Anschrift: Universität Duisburg-Essen, Fakultät für Bildungswissenschaften, Campus Essen/Weststadttürme, 45117 Essen

DDS – Die Deutsche Schule
Beiheft 12, S. 255-275
© 2013 Waxmann

Laura Fölker

„Da ist einfach 'ne gewisse Leistungs- und auch Schulentwicklungsgrenze erreicht!"

Schulfusionen als zusätzliche Reformimpulse im Zuge der Integration von Haupt- und Realschulbildungsgängen

Zusammenfassung

Das Schulsystem wird einer Neuordnung unterzogen: Demografischer Wandel und ökonomische Zwänge geben aktuell Anlass dazu, kleine Schulstandorte im Zuge von Schulstrukturreformen zusammenzulegen. Durch den in verschiedenster, negativer oder positiver Form beeinflussten schulischen Gestaltungsfreiraum führen dabei Schulfusionen in vielen Fällen zu Identifikationsverlusten, können aber auch gleichzeitig Möglichkeiten für pädagogische Erneuerungen im Sinne von Schulentwicklungsprozessen entstehen lassen. Im Hinblick auf diese Problemstellungen versucht der Artikel, einen Beitrag zu der noch weitestgehend unerforschten Thematik von Schulfusionen zu leisten, indem anhand von dokumentarischen Rekonstruktionen selbige – und die aus ihnen entstehenden Herausforderungen – in ihrer Bewältigung analysiert werden. Fallrekonstruktiv untersucht werden dabei Schulfusionsprozesse zweier Schulen. Anhand dieser soll vorrangig die Komplexität von Schulentwicklungsprozessen für Einzelschulen sichtbar gemacht werden, welche wiederum verpflichtend eingeführte, strukturelle Imperative in ihre Arbeit einflechten müssen. Dabei geht es im Kern um die Analyse der schulprogrammatischen und pädagogischen Prinzipien und Arbeitsformen beider Schulen im Fusionsprozess.

Schlüsselwörter: Schulstrukturreformen, Schulfusionen, rekonstruktive Fallstudien, Schulentwicklungsprozesse

Laura Fölker

"A Certain Limit of Effort and School Development Has Been Reached!"

School Fusions as Additional Reform Impulses in the Course of the Integration of Various Secondary Schools (*Hauptschulen* and *Realschulen*)

Abstract

The school system is being restructured. Demographic changes as well as economic constraints currently lead to the amalgamation of schools over the course of structural school reforms. Because the structural impulse of merging is forced upon the schools, it affects the educational freedom of scope of all involved parties within the merging process. Thus, school merging often leads on to the loss of the school identity on all sides involved – the merger school as well as the merged one. Nevertheless, the combination of schools may also give rise to educational renewals concerning processes of school development. Concerning these challenges the article tries to contribute to the still widely unexplored subject matter of school merging, by analyzing two schools when coping with the challenges just illustrated. Therefore, the article aims at showing the complexity of school development processes within the schools investigated, when being forced to interlace mandatory introduced imperatives within their school work. Essentially, the article thus focuses on the analysis of principles and values to influence the educational practices of the teaching staff of both schools within the merging process.

Keywords: structural educational school reforms, school merging, reconstructive case studies, processes of school development

Vor dem Hintergrund demografischer Entwicklungen wie die anhaltend niedrige Geburtenrate und die dadurch sinkenden Schülerzahlen, aber auch durch die politisch motivierten Kosteneinsparungen vieler Bundesländer werden derzeit bundesweit kleine Schulstandorte im Zuge der Umsetzung von Schulstrukturreformen miteinander fusioniert. Diese Schulfusionen stellen alle an Schule beteiligten Akteure vor neue Herausforderungen und bieten zugleich Chancen für Veränderungen im Schulentwicklungsprozess. Im Zusammenhang mit dem so genannten „Hauptschulsterben" sind es derzeit vor allem niedrigqualifizierende Schulen, deren Akzeptanz aufgrund des immer größer werdenden Interesses an höheren Bildungsabschlüssen innerhalb des Bildungssystems sinkt und die deshalb Fusionsprozessen unterzogen werden (vgl. Meister 2012).

Im Hinblick auf die mit diesen Entwicklungen verbundenen Herausforderungen stellt der folgende Beitrag Befunde aus dem BMBF-geförderten Projekt „Bildungsbezogene Integration unter Bedingungen vielfacher Bildungsrisiken"[1] vor, in welchem zwei

1 Das Projekt „Bildungsbezogene Integration unter Bedingungen vielfacher Bildungsrisiken – Fallstudien zu zwei Schulen mit multikulturellen Schülerschaften in benachteiligten Quartieren" unter der Leitung von Nicolle Pfaff ist Teil der vom Bundesministerium für Bildung und Forschung (BMBF) initiierten Förderlinie „Steuerung im Bildungssystem". Außer der Autorin dieses Beitrages waren Thorsten Hertel als wissenschaftlicher Mitarbeiter, Verena

Schulen in benachteiligten Quartieren nach Faktoren erfolgreichen Lernens und bildungsbezogener Integration untersucht wurden. Beide Untersuchungsschulen sind im Rahmen von aktuell ablaufenden Schulstrukturreformen[2] durch Fusionsprozesse niedrigqualifizierender Schulen entstanden.

Vor diesem Hintergrund besteht das Ziel des folgenden Beitrags darin, die von den Akteuren an beiden Untersuchungsschulen thematisierten Alltagspraxen im Bezug auf die ablaufenden Fusionsprozesse und die dazugehörigen Konfliktlinien zu rekonstruieren, um zentrale Orientierungen im Umgang mit den von den zuständigen Behörden geforderten Fusionen herauszuarbeiten. Die dokumentarischen Rekonstruktionen von Fusionsprozessen in ihrem Vollzug bieten hierbei die Möglichkeit, Bedingungen ihrer Realisierung an Schulen empirisch zu erfassen und somit einen Beitrag zur Generierung von Wissen zur noch weitgehend unerforschten Thematik von Schulzusammenlegungen zu leisten.

1. Theoretische Hintergründe und empirische Befunde

Studien zu Schulzusammenlegungen gehen in Deutschland auf die Umstellung von Schulstrukturen in den neuen Bundesländern zurück, wo seit der Wiedervereinigung fast jede zweite Schule aufgrund der vor allem durch Geburtenrückgang („Geburtenschock") und deutsche Binnenmigration bedingten sinkenden Schülerzahlen geschlossen oder mit einer anderen Schule fusioniert wurde (vgl. Fleischer 2009; Meister 2012, S. 167). In diesem Zusammenhang sind beschreibende „Untersuchungen zur Entwicklung von Schulstandorten im Kontext von demografischen Schrumpfungsprozessen" (vgl. Budde 2007; Weishaupt 2002; beide zitiert nach: Meister 2012, S. 169) zu finden. Darüber hinaus wurden „Tendenzbeschreibungen zu den Auswirkungen der demografischen Entwicklungen auf die Schulstruktur innerhalb der neuen Bundesländer seit den 1990 Jahren" (vgl. Weishaupt/Zedler 1994; zitiert nach: Meister 2012, S. 169) geleistet.

Aktuelle Arbeiten zu Schulfusionen basieren auf einzelnen Erfahrungsberichten sowie Empfehlungen zu Realisierungen von Schulzusammenlegungen (vgl. Buchen/Horster/

Busch als Praktikantin sowie Johanna Wieneke, Nele Kehr und Tirsa Lönker als studentische Hilfskräfte an dem Projekt beteiligt. Ausführliche Projekt- und Ergebnisdarstellungen sind auf der Website URL: www.brennpunkt-schule.de (Zugriffsdatum: 20.07.2013), zu finden.

2 Die in vielen Bundesländern eingeführten Schulstrukturreformen beinhalten vorrangig die Integration von Haupt- und Realschulbildungsgang zu einer integrativen Schulform, in der alle Schulabschlüsse bis zum Abitur nach der 13. Klasse abgelegt werden können. Weitere Ziele der Schulstrukturreformen sind zum einen die durchgängige Umsetzung des individualisierten Lernens, zum anderen die damit verbundene Hoffnung auf mehr Chancengleichheit, indem man Schülerinnen und Schülern aus sozial schlechteren Verhältnissen den Zugang zu höheren Bildungsgängen eröffnet sowie die Integration von Schülerinnen und Schülern mit Behinderung ermöglicht.

Rolff 2009). So gibt Fleischer (2009) eine detaillierte Beschreibung der Umsetzung einer Schulfusion, indem er seine Erfahrungen vor dem Hintergrund der eigenen Tätigkeit als Schulleiter einer Fusionsschule rückblickend schildert. Die Frage danach, wie Betroffene mit einer solchen Situation umgehen, ist ein weiterer Schwerpunkt innerhalb einzelner Erfahrungsberichte zu Schulfusionen (vgl. Müller 2010).

Auf dem Gebiet wissenschaftlicher Studien hat Tischendorf (2007) zu problematischen Intergruppenbeziehungen nach Schulfusionen gearbeitet, wobei die psychologischen Folgen von Schulfusionen sowie die damit verbundenen identitätsrelevanten kognitiven Prozesse aus einer sozialpsychologischen Perspektive untersucht wurden (vgl. ebd., S. 178). Die Arbeit zeigt, dass alle an Schulfusionen beteiligten Akteure analog zu Unternehmensfusionen, sowohl auf der individuellen als auch auf der intergruppalen Ebene, negative psychologische Reaktionen auf die Fusion entwickelten. In diesem Zusammenhang unterlag die Phase im Anschluss an die Schulfusion einem Lagerdenken der Akteure, indem sowohl die Lehrerinnen und Lehrer als auch die Schülerinnen und Schüler der jeweiligen zusammengelegten Schulen eine zunehmend hohe Identifikation mit der Gruppe entwickelten, der sie vor der Fusion angehörten, was in der Folge mit einer negativen Bewertung der jeweils anderen Schule einherging (vgl. ebd.). Die Arbeit konnte zeigen, dass die negativen Bewertungen der jeweils anderen Schule gepaart mit einer hohen Identifikation mit der Arbeitsweise und den Konzepten der eigenen „Präfusions-Schule" sowohl den Erfolg einer Schulfusion bezüglich einer raschen Integration mit harmonischen Intergruppenbeziehungen gefährden als auch eine Verschlechterung des Schulklimas und implizit der individuellen Erfolgskriterien, wie z.B. des schulischen Wohlbefindens und der Schulleistungen, herbeiführen (vgl. ebd.).

Zusätzlich zu diesen inhaltlichen Schwerpunkten wird innerhalb der Arbeit eine begriffliche Abgrenzung verschiedener Formen von Schulfusionen geleistet, indem in Anlehnung an die ökonomische Fusionsforschung fünf Fusionsmodelle vorgestellt und auf Schulzusammenlegungen übertragen werden.

Dabei wird zunächst das Assimilationsprinzip beschrieben (vgl. Marks/Mirvis 2001; Mottola u.a. 1997; beide zitiert nach: Tischendorf 2007), bei dem die Kultur und Identität einer der zu fusionierenden Organisationen überwiegt oder als den Fusionsprozess lenkend wahrgenommen wird. Dieses Prinzip erinnert an das aus der Wirtschaftswissenschaft bekannte Modell einseitiger Machtverhältnisse, welches oftmals in Form von Akquisitionen oder Aufkäufen realisiert wird (vgl. ebd., S. 6). Bezogen auf Schulfusionen bedeutet dies, dass eine der zusammenzulegenden Schulen gezwungen ist, ihre Existenz, wie auch ihr Schulgebäude, die Schulleitung und den Schulnamen, völlig aufzugeben (vgl. ebd., S. 10).

Innerhalb des Integrationsprinzips nach Schweiger und Weber (1989) werden Teile beider Organisationen innerhalb der Fusion berücksichtigt. Giessner und Kollegen

(2006; zitiert nach: Tischendorf 2007, S. 7) differenzieren innerhalb dieses Prinzips zwei untergeordnete Modelle der Zusammenlegung. Während dabei das Gleichheitsprinzip die Integration der Kulturen aller Fusionspartner zu gleichen Teilen anstrebt, beschreibt das Proportionalitätsprinzip unausgewogene Machtverhältnisse beim Erhalt der Kulturen innerhalb des Fusionsprozesses (vgl. ebd.). Überträgt man das Integrationsprinzip auf Schulfusionen, werden zunächst einmal beide Fusionsschulen in der neugegründeten Schule repräsentiert. Dabei könnte eine der beiden Fusionsschulen das Schulgebäude und die andere beispielsweise den Schulleiter in die neue Schule einbringen. Im Anschluss an die Fusion würden dann Arbeitsgemeinschaften, wie beispielsweise eine Steuergruppe, innerhalb der neuen, fusionierten Schule mit Akteuren aus den ehemaligen Fusionsschulen besetzt, und Schulprofile sowie Normen und Werte der zusammengelegten Schulen würden sich zu gleichen bzw. ungleichen Teilen im Schulprofil der neugegründeten Schule wiederfinden (vgl. ebd., S. 10).

Das Transformationsprinzip stellt ein weiteres Fusionsmodell dar. Es bedingt den kompletten Neuaufbau einer Kultur und Identität innerhalb der neugegründeten Organisation (vgl. ebd., S. 7). Bezogen auf Schulfusionen geben beide Fusionsschulen hier ihre Existenzen auf und kommen an einem neuen Ort, in einem neuen Schulgebäude, unter einem neuen Schulnamen und einer neuen Schulleitung zusammen. Dabei können auch Lehrerstellen neu besetzt werden.

Schließlich bezieht sich das Aufrechterhaltungsprinzip als letztes Fusionsmodell trotz der formalen Integration zweier Organisationen auf Kontinuitäten der Gruppenzugehörigkeiten und der über einen längeren Zeitraum etablierten Identitäten (vgl. Marks/Mirvis 2001; zitiert nach: Tischendorf 2007, S. 7). Mit Blick auf Schulfusionen wird dabei eine Koexistenz angestrebt, innerhalb derer alle an der Fusion beteiligten Schulen das jeweilige Fundament an Arbeitsformen und Handlungspraktiken aufrechterhalten.

Eine aktuelle empirische Studie untersucht schließlich kollegiale Kooperationspraxen von Lehrerinnen und Lehrern innerhalb der Entwicklung von Ganztags-Sekundarschulen in Sachsen-Anhalt (vgl. Meister 2012). Dabei kam zusätzlich zur Einführung der Ganztagsschule die Fusionsanforderung für einige Schulen erschwerend hinzu. Vor diesem Hintergrund wurden Formen kollegialer Kooperation sowie deren Bedeutung im Rahmen von Schulfusionen untersucht. Die Ergebnisse der Studie weisen darauf hin, dass „Schulfusionen besondere Organisations-, Koordinations- und Aushandlungsprozesse erfordern", welche zusätzliche Energien aller am Fusionsprozess beteiligten Akteure erforderlich machen und auf kollegiale Kooperationen im Rahmen der alltäglichen Schulentwicklungsarbeit angewiesen sind (ebd., S. 175).

Wie einleitend erwähnt und anhand der knappen Skizze vorliegender Studien bestätigt, mangelt es innerhalb der Schulentwicklungsforschung an empirischen Be-

funden zur Umsetzung von Schulfusionen. Dies gilt insbesondere für die Analyse von Entwicklungsprozessen, die sich vor dem Hintergrund von Restrukturierungsvorgaben entfalten. Analysen zu Einzelschulen, die in besonderer Weise von solchen Reformimpulsen betroffen sind, können dabei systematisches Wissen einerseits über die Verläufe solcher Strukturinnovationen liefern und andererseits die dabei entstehenden und weitgehend unbekannten Dynamiken von Schulentwicklungsprozessen in den Blick nehmen.

2. Materialgrundlage und methodischer Zugang

Im Anschluss an die internationale erziehungswissenschaftliche Forschung zu Schulen unter herausfordernden Bedingungen untersucht das Projekt „Bildungsbezogene Integration unter Bedingungen multipler Bildungsrisiken – Fallstudien zu zwei Schulen mit multikulturellen Schülerschaften in benachteiligten Quartieren" Prozesse der Schulentwicklung in segregierten Stadtteilen unter Bedingungen schulstruktureller Reformen. Die Untersuchung basiert auf rekonstruktiven Fallstudien und zielt im Kern darauf, Faktoren erfolgreichen Lernens und bildungsbezogener Integration an Schulen in benachteiligten Quartieren sichtbar zu machen.

Als Annäherung an das Feld wurde dabei ein offener ethnographischer Zugang gewählt. In diesem Zusammenhang wurden neben Experteninterviews mit den verschiedensten Akteuren der Schulen und Schulverwaltungen auch Gruppendiskussionen geführt sowie Audioaufnahmen pädagogischer Alltagsinteraktionen angefertigt.

Die Auswertung des so gewonnenen Datenmaterials erfolgt mit der dokumentarischen Methode der Interpretation (vgl. Bohnsack 2003; Bohnsack/Nentwig-Gesemann/Nohl 2001; Bohnsack/Pfaff 2010; Nohl 2007). Diese zielt auf die Rekonstruktion von habituellen Orientierungen von Individuen und Kollektiven in spezifischen Milieuzusammenhängen und vollzieht dabei drei Interpretationsschritte (die formulierende Interpretation, die reflektierende Interpretation und die Typenbildung), in denen die diskursive Verhandlung und die Art und Weise der Präsentation von Inhalten aufgearbeitet werden (vgl. z.B. Bohnsack/Nentwig-Gesemann/Nohl 2001), um zentrale Orientierungsfiguren und gemeinsame Handlungspraxen von Akteursgruppen zu rekonstruieren.

Vor diesem Hintergrund werden im vorliegenden Beitrag empirische Rekonstruktionen zu beiden Untersuchungsschulen vorgestellt, die Orientierungen der pädagogischen Professionellen im Bezug auf die Realisierung der von Seiten der regionalen Bildungspolitik und lokalen Schulverwaltung initiierten Fusionen herausarbeiten. Die aus dem Diskurs rekonstruierten, die Handlungspraxis leitenden

Orientierungen der Akteure im Fusionsprozess werden dabei durch die Interpretation von Diskursabschnitten von hoher metaphorischer Dichte herausgearbeitet und im vorliegenden Beitrag anhand einschlägiger Zitate exemplarisch vorgestellt. Vorrangiger Gegenstand der dabei durchgeführten Kontrastierung sind die aus den Orientierungen resultierenden, über lange Jahre etablierten Handlungspraktiken der pädagogischen Professionellen der innerhalb beider Untersuchungsschulen zusammengelegten Fusionsschulen. Darauf aufbauend sollen im Anschluss die Ergebnisse der Rekonstruktionen zunächst an den Forschungsstand rückgebunden werden, um sie abschließend mit Blick auf mögliche Handlungsempfehlungen für die Bildungspolitik zu diskutieren. Mit dem Fokus auf Schulfusionen im Kontext von Schulstrukturreformen konzentrieren sich die folgenden Rekonstruktionen damit auf einen von insgesamt fünf Untersuchungsschwerpunkten des Projekts.[3]

3. Schulische Praxis der Realisierung von Schulfusionen – Empirische Rekonstruktionen zu einem lückenhaften Forschungsfeld

Die erste Untersuchungsschule entstand zum Schuljahr 2009/2010 aus der Fusion einer Hauptschule und einer integrierten Haupt- und Realschule, die an einem der beiden ehemaligen Schulstandorte zusammengelegt wurden und sich im Hinblick auf zentrale strukturelle und kulturelle Bedingungen deutlich unterschieden. Mit der Zusammenlegung beider Schulen entschied man sich für eine kooperative Schulleitung, allerdings operierte die Schulleitung der *ersten Fusionsschule* nach wie vor als offizielle Leitungsinstanz der neuen integrierten Schule. Zum letzten Erhebungszeitpunkt im Projekt (März 2012) wurde aufgrund von Konflikten die Kooperation auf Schulleitungsebene aufgelöst. Die Schulleitung der *ersten Fusionsschule* blieb daraufhin im Amt, die Schulleitung der *zweiten Fusionsschule* verließ die Schule.

Die zweite Untersuchungsschule entstand ebenfalls aus der Zusammenlegung verschiedener niedrigqualifizierender Schulen zum Schuljahr 2009/2010. Hierbei blieben beide Schulstandorte der neugegründeten integrierten Schule als Schulgebäude erhalten. Die Schulleitung der *ersten Fusionsschule* blieb auch hier weiter als neue Schulleitung der integrierten Schule bestehen, während die Schulleitung der *zweiten Fusionsschule* diese bereits vor der Fusion verließ. Auch zahlreiche Lehrkräfte aus dem Kollegium der *zweiten Fusionsschule* ließen sich an andere Schulen versetzen,

3 Die zentralen Analyseschwerpunkte im Projekt beziehen sich auf das Lehren und Lernen unter herausfordernden Bedingungen (1), das Schulmanagement (2) und die professionelle Kooperation (3). Aufgrund der besonderen Situation der Untersuchungsschulen im Kontext von Schulstrukturreformen und Schulfusionen werden auch diese Aspekte als zu den genannten drei Untersuchungsschwerpunkten querliegende Themen untersucht.

sodass nur wenige Lehrerinnen und Lehrer in die neugegründete integrierte Schule aufgenommen wurden.

Beide Schulfusionen wurden im Kontext übergreifender Schulstrukturreformen realisiert, die vor allem auf die Integration von Haupt- und Realschulbildungsgängen zielten. Die Bedingungen, unter welchen die Zusammenlegungen stattgefunden haben, konnten im Bezug auf die erste Untersuchungsschule im Verlauf der jeweils vier Erhebungsphasen im Projekt mit Hilfe von Interviews und Gruppendiskussionen festgehalten werden. An der zweiten Untersuchungsschule kam der Fusionsprozess retrospektiv in den Blick.

Aus methodischer Sicht ergibt sich durch diese Differenz im Material für beide Schulen ein unterschiedlicher Zugang zu Schulfusionsprozessen. So kristallisieren sich im Fusionsprozess der ersten Untersuchungsschule die zahlreichen Herausforderungen, welchen die Schule in der unmittelbaren Umsetzung der strukturellen Imperative ausgesetzt war, heraus: Die dem Fusionsprozess zugrundeliegenden Dynamiken können damit kleinschrittig und detailliert empirisch nachgezeichnet werden, und die damit im Zusammenhang stehenden Konfliktlinien in Bezug auf die Rekontextualisierung der Reformvorgaben und deren Einbettung in den konjunktiven Erfahrungsraum der Akteure werden unmittelbar erfasst.

Im Unterschied dazu werden an der zweiten Untersuchungsschule Erfahrungen mit der Fusion von den befragten Beteiligten im Rückblick und damit auch im Modus einer Reflexion der Ereignisse beschrieben, wodurch nur diejenigen Herausforderungen und Konfliktlinien sichtbar werden, die den Akteuren als solche präsent sind.

Ungeachtet dieser Differenzen im Datenmaterial der Untersuchung zeigen die Rekonstruktionen, dass die Akteure an beiden Untersuchungsschulen zunächst einmal vergleichbar, nämlich mit Skepsis und Unbehagen, auf die Fusionsvorgaben reagierten. Dabei sind die im Folgenden nachzuzeichnenden erfolgreichen und weniger erfolgreichen Fusionsprozesse maßgeblich unterschiedlichen strukturellen und schulgeschichtlichen Bedingungen zuzuschreiben.

3.1 Angleichung an Bestehendes und nachträgliche Legitimierung

Im Bezug auf die Bedingungen, die Freiwilligkeit und das Mitspracherecht der Beteiligten, unter denen die Schulfusionen stattfanden, reflektiert der stellvertretende Schulleiter der *ersten Fusionsschule* an der ersten Untersuchungsschule die Situation wie folgt:

> *Ja neue Schule (.) es kommt aber für uns eben schon noch dazu (.) die Fusion. (.) Also ich meine ich hab im letzten Schuljahr (.) mehrere Fortbildungen ge-*

> *macht also das war (.) dreimal drei zweitägige Module und da sah man dann doch sehr deutlich dass es Unterschiede gibt ahm sehr ahm sehr ahm sehr massiv zwischen den Schulen die diese Umstrukturierung aus sich heraus machen Schule alleine (.) oder Leute die fusionieren müssen* (Fallstudie I, Interview stellv. Schulleiter *erste Fusionsschule*, Erhebung 1, 9/2010, Z. 17-23).

Dieses Zitat zeigt deutlich, dass an der ersten Untersuchungsschule zwischen der Schulstrukturreform und der Fusion unterschieden wird. Bedingt durch die Schulreform ergab sich die Gründung einer *„neue[n] Schule"*. Dabei ist die *„neue Schule"* nicht gleichzusetzen mit dem Fusionsprozess, welcher vorrangig zu bewältigen ist. So bestand die erste Aufgabe der neugegründeten, integrierten Schule zunächst darin die Fusion der beiden zusammengeführten Schulen zu verwirklichen. Diese Anforderung kam zusätzlich zum Aufbau der *„neuen Schule"* dazu. Folglich stellt die Fusion zunächst die zentrale Bedingung der Realisierung der Schulreform dar.

Die vom stellvertretenden Schulleiter angesprochenen *„massiven Unterschiede"* zwischen jenen Schulen, die eine *„Umstrukturierung aus sich heraus machen"*, und den *„Leuten"*, die gezwungen sind zu fusionieren, verweisen auf einen positiven Gegenhorizont der geringeren Herausforderungen im Bereich der Schulentwicklung. Mit Verweis auf diejenigen Schulen, welche die Schulreform auf der Grundlage einer gemeinsamen Basis der Zusammenarbeit umsetzen können, wird die Schulfusion zur Belastung und Voraussetzung für inhaltliche und organisatorische Schulentwicklungsprozesse im Sinne der Strukturreform. Implizit werden hier auch die Entscheidungsträger kritisiert, ungleiche Bedingungen zwischen den Schulen herzustellen.

Diese Art und Weise, Schulfusionen von übergeordneter Ebene verordnet zu bekommen, gerät auch an der zweiten Untersuchungsschule im Projekt als zusätzliche strukturelle Veränderung zur Integration der Haupt- und Realschulbildungsgänge als problematisch in den Blick:

> *zur Fusion selber da haben wir gar nicht so viel Mitspracherecht es is hier der Standort da der Standort peng zusammen so also äh die Lehrer sind ja nicht so, dass die hier sagen können das wollen wir nicht oder so (.) wir wir mussten uns arrangieren ne und äh das nehmen was gerade zur Verfügung stand so is das ne* (Fallstudie II, Interview Lehrerin, Z. 174-185).

> *dass wir damit natürlich einerseits ähm sozusagen etwas gekriegt haben im zweiten Standort den wir überhaupt nicht wollten, dass wir als zweites dann eine (.) Schülerschaft die ja noch da ist also diese schwierigste Schülerschaft übernehmen durften äh die dann sozusagen sich ja irgendwie auch in das neue Konstrukt *integrierte Schulform* eingliedern sollte das war auch ein bisschen*

> *schwierig und wir haben auch ne bestimmte Lehrerschaft gekriegt die auch zum Teil seit Jahren da war* (Fallstudie II, Gruppendiskussion Schulleiter und stellv. Schulleiterin *erste Fusionsschule*, 5/2011, Z. 629-639).

Beide Zitate verdeutlichen zunächst die Irritation und Skepsis der Akteure gegenüber der Zusammenlegung beider Schulen. Die von außen generierten Fusionsbedingungen erscheinen im Kollegium fragwürdig, und die ungewollten Folgen der Zusammenlegung werden durch die Aufzählung der Übernahme eines zweiten Standortes, einer *„schwierigsten"* Schülerschaft und einer bestimmten Lehrerschaft präzisiert. Hier wird explizit das Zustandekommen der Fusion kritisiert, indem offen gelegt wird, dass die Lehrerinnen und Lehrer kein Mitspracherecht hatten, als es darum ging, welche Schulen zusammengelegt werden. So musste man *„das nehmen was gerade zur Verfügung stand"*, was wiederum, ähnlich wie bei der ersten Untersuchungsschule, aufzeigt, dass die bildungspolitische Entscheidung der Zusammenlegung genau dieser beiden Schulen aus Sicht der Lehrerin nicht die optimale Lösung war.

In der Aussage des Schulleiters scheint es, ähnlich wie in der ersten Fallstudie, eine Abgrenzung zwischen der Schulstrukturreform und der Fusion zu geben, indem erläutert wird, dass man zusätzlich zu den ungewollten Fusionsfolgen ebenfalls die Eingliederung der neuen Schülerschaft in *„das Konstrukt *integrierte Schulform*"* umsetzen musste, was im Nachhinein als schwierig beurteilt wird. Interessanterweise wird die Aufgabe der Eingliederung in das Konzept der *integrierten Schulform* nur der neuen Schülerschaft überantwortet, nicht aber der bestehenden Schülerschaft oder gar den pädagogischen Akteuren an der Schule. Dies kann mit Blick in die Schulgeschichte der *ersten Fusionsschule* der zweiten Fallstudie erklärt werden, wo durch einen bereits vor Jahrzehnten durchgeführten Schulversuch die Integration von Haupt- und Realschulbildungsgang an der Schule seit langem Realität ist. Dadurch werden die aktuellen Reformimpulse nach Ansicht der Schulleitung nicht als Innovation wahrgenommen, sondern entsprechen der bereits etablierten Praxis. Durch die Implementierung integrierter Bildungsgänge wurden weitreichende Schulentwicklungsprozesse innerhalb der Schule in Gang gesetzt, welche die pädagogische Arbeit der professionellen Akteure an der Schule bis heute nachhaltig prägen und vor diesem Hintergrund zu einem spezifischen Umgang mit der aktuellen Reform führen.

Auch die *erste Fusionsschule* der ersten Fallstudie nahm schon Mitte der 1980er-Jahre an einem Schulversuch teil, in welchem, so die Ansicht zahlreicher Akteure, die aktuellen Reformimpulse der Strukturreform bereits umgesetzt wurden. Vor diesem Hintergrund scheint die Notwendigkeit schulorganisatorischer Umstrukturierungen an der *ersten Fusionsschule* der ersten Fallstudie ebenfalls fragwürdig:

> *Ja. Dann war für uns die Struktur (.) also wenn ich sage für uns ahm (2) für die Kollegen der ehemaligen *Fusionsschule 1* unser Schulkonzept (.) beinhaltete schon sehr sehr vieles oder fast alles (.) was die *integrierte Schulform* ahm auch beinhaltet (.) und von daher war es für uns relativ leicht (.) ahm zu sagen ahm diese neuen Sachen anzunehmen (.) was uns teilweise von andern sogar vorgeworfen wird dass gesagt wird wir wollen ihnen unsern Stil über [...] ahm sie müssten sich unserem Stil (anpassen) (.) wobei das eigentlich so ist (.) diese ahm die *Fusionsschule 1* hat in den 80er und 90er Jahren schon eine integrative Schulform entwickelt und ahm man könnte fast sagen dass ahm wir eigent- ne ist einfach so einen großen Teil (.) der ahm Dinge die die integrierte Schulform vorhat schon lange praktizieren* (Fallstudie I, Interview stellv. Schulleiter *erste Fusionsschule*, 9/2010, Z. 38-50).

Hier hat es den Anschein, dass die *erste Fusionsschule* das passende Schulkonzept, welches auch im Sinne der Schulstrukturreform ist, schon, bevor diese eingeführt wurde, etabliert hatte. Diesbezüglich schreibt der stellvertretende Schulleiter seiner Schule die Vorwegnahme der Realisierung des integrativen Schulkonzepts im Sinne der Schulreform zu und lässt selbiger damit eine Vorreiterrolle im Vergleich zur *zweiten Fusionsschule* zukommen. Was die integrierte Schulform darstellt, ist demnach kein Gegenstand von Aus- oder Verhandlungen, sondern trägt objektive Eigenschaften und entsprechende Kriterien, die bereits von der *ersten Fusionsschule* erfüllt werden. Aufbauend auf diese Argumentationslinie erfolgt die explizite Forderung an die *zweite Fusionsschule*, sich an das schon etablierte und geeignete, weil von der Schulreform bestätigte, Konzept der *ersten Fusionsschule* anzupassen.

An der hier zusätzlich deutlich werdenden Zurückweisung der Anpassungsforderung von Seiten der *zweiten Fusionsschule* zeigt sich, dass beide Schulen an ihren bestehenden Praktiken festhalten. In einem noch vor der Zusammenlegung durchgeführten Workshop werden diesbezüglich die Perspektiven und Orientierungen der Akteure der *zweiten Fusionsschule* hinsichtlich der bevorstehenden Schulfusion mit der *ersten Fusionsschule* deutlich:

> *ich würde mir wünschen, dass das so bleibt, dass wir auch ohne es vielleicht auszusprechen, wenn wir da rüber gehen, ne verschworene Gemeinschaft bleiben, man muss es ja sozusagen nicht mit dieser- mit diesem Triumphgebaren äh kombinieren, sondern man muss es ganz leise auch transportieren, dass wir der Motor dieser Entwicklung sein- sind und dass wir viele Strukturen haben, die gut sind. Damit will ich aber nicht sagen ähm dass alles sozusagen eins zu eins übersetzbar ist, sondern […] wir werden sag ich mal uns wan- auch wandeln müssen und äh diese Fähigkeit sehe ich aber hier. Und das ist denke ich mal das Gute, das wir haben. So an menschlicher Qualität [...]. Abgesehen mal von den ganzen Strukturen die wir äh etabliert haben* (Fallstudie I, Workshop

> Stärken und Schwächen, stellvertretender Schulleiter *zweite Fusionsschule*, 1/2010, Z. 537-547, gekürzt).

Der stellvertretende Schulleiter der *zweiten Fusionsschule* spricht hier für den Erhalt der Kollegialität der pädagogischen Professionellen an der Schule. Mit der Zurückweisung jeglichen *„Triumphgebaren[s]"* und dem Apell, die Gemeinschaftlichkeit nicht zu thematisieren, sondern zu leben, beschreibt er die Art und Weise des Zugangs auf die neuen Kolleginnen und Kollegen der *ersten Fusionsschule*. Eine Lehrerin betont, dass man nicht als *„Kolonialherren"* oder *„Okkupanten"* (Fallstudie I, Workshop Stärken und Schwächen, Lehrerin, 1/2010, Z. 467-468) auftreten bzw. den Kolleginnen und Kollegen der *ersten Fusionsschule* vorschreiben könne, was sie zu tun haben: *„und ich denke es is auch schwierig wenn wir dahin gehen und sagen: Hey, wir machen, wir ham das bis jetzt immer richtig toll gemacht, wir können das und es wäre schön, wenn ihr euch sozusagen an uns richtet"* (Fallstudie I, Workshop Stärken und Schwächen, Lehrerin, 1/2010, Z. 368-370). In diesen Aussagen, wie auch in der Aussage des stellvertretenden Schulleiters, man sei eine *„verschworene Gemeinschaft"*, dokumentiert sich die Selbstwahrnehmung als kompetente Professionelle und die Absprache ebenjener Kompetenz gegenüber den Lehrkräften der *ersten Fusionsschule*. In den Metaphern des *„Motor[s] der Entwicklung"*, der *„Strukturen […] die gut sind"*, der *„Fähigkeit [sich] zu wandel[n]"* und der *„menschliche[n] Qualität"* stecken vielfältige Hinweise auf die Überlegenheit vor allem der organisationalen Praktiken, aber auch der Kompetenz der Akteure der *zweiten Fusionsschule* gegenüber denen der *ersten Fusionsschule*.

Diese Rekonstruktionen weisen darauf hin, dass eine Angleichung an die bestehenden Strukturen und etablierten Praktiken der jeweils anderen Schule von den Akteuren beider Fusionsschulen bereits im Vorfeld erschwert oder sogar ausgeschlossen werden. Die Praktiken der *zweiten Fusionsschule* werden mit der Begründung abgelehnt, man sei mit den Arbeitsformen der *ersten Fusionsschule* näher an den Erwartungen der Schulreform. Andersherum gilt die innerhalb der *zweiten Fusionsschule* herrschende pädagogische Realität den Lehrkräften als Maßstab für gute schulische Praxis. An dieser messen sie die Praktiken der *ersten Fusionsschule* und werten selbige damit gleichzeitig ab. Bedingt durch die Auseinandersetzungen um die Schulfusion lässt sich an dieser Stelle die Orientierung am Erhalt der eigenen, über lange Jahre etablierten Praktiken und Strukturen festhalten, welche gleichzeitig die inhaltliche Umsetzung der Schulstrukturreform überlagert.

Mit ähnlichen Vorsätzen wie innerhalb der ersten Fallstudie, nämlich die Fusion nicht als *„feindliche Übernahme"* aussehen zu lassen, jedoch mit gegensätzlichem Ausgang, ist man der Zusammenlegung auch innerhalb der zweiten Fallstudie begegnet:

> *Und dann standen wir vor dem Prozess zu sagen hier ist eine Schule die hat man ausgewählt und dann hat man dort ein Anhängsel was dazu kommt und wie machen wir das jetzt eigentlich und was machen wir eigentlich miteinander und dann gab es noch die Situation dass der Schulleiter der drüben an der anderen Schule war der auch ein sehr anerkannter Mensch im Stadtteil ist der ein sehr anerkannter Schulleiter […] und äh der der musste nun gehen also da gab es sozusagen eine Abwägung welcher der beiden Schulleiter soll denn Schulleiter werden und insofern musste er dann gehen das war die eine Situation da kommt 'n Schul- Schulleiter von außen und es war auch so dass die stellvertretende Schulleiterin die möglicherweise wenn sie drauf bestanden hätte dann dabei zu bleiben dann auch hätte bleiben können äh die wollte aber nich aus privaten Gründen […] insofern haben wir das da drüben wirklich feindlich übernehmen dürfen (.) Und äh das war genau der Punkt also wir hätten es wahrscheinlich sowieso nich anders gemacht aber das hat uns natürlich noch dabei äh unterstützt zu sagen äh dass wir grade in diesem Umgang miteinander oder mit diesen Sachen miteinander sehr sehr kleinschrittig vorgehen müssen und von vornerein eben jeden Verdacht von feindlicher Übernahme ähm möglichst möglichst an die Seite stellen wollen das heißt was wir gemacht haben is zunächst einmal natürlich mit ner Steuergruppe zusammen gesessen seit letztem Schuljahr praktisch ähm also bevor das dann mit der Fusion war haben wir uns zusammen gesetzt eigentlich im November schon haben angefangen ähm bestimmte Dinge zu planen wie wir diesen gesamten Prozess gestalten* (Fallstudie II, Gruppendiskussion Schulleiter und stellv. Schulleiterin *erste Fusionsschule*, 5/2011, Z. 681-717).

Auch hier sollte, ähnlich wie innerhalb der Fusionsschulen der ersten Fallstudie, der Zugang auf das neue Kollegium von Seiten der *ersten Fusionsschule* frei von *„jegliche[m] Verdacht einer feindlichen Übernahme"* realisiert werden. Um dieses Vorhaben verwirklichen zu können gründete man eine Steuergruppe, deren zentrale Aufgabe darin bestand, den Fusionsprozess zu gestalten. So begann die Fusion vor der eigentlichen Fusion, indem schon im Vorfeld Konzepte zur Umsetzung der Fusion abgesprochen wurden. Folglich dokumentiert sich darin eine Orientierung, welche die Bereitschaft zum Austausch über Werte, Ziele und Methoden sowie zum Abgleich des eigenen Handelns mit dem der neuen Kolleginnen und Kollegen verdeutlicht. Dabei wurde von keiner der an der Zusammenlegung beteiligten Fusionsschulen die Anpassung der jeweils anderen an bestehende Konzepte erwartet, sondern man entschied sich, in Zusammenarbeit und unter Einbezug der jeweiligen Erfahrungen die bestehenden Konzepte zu ergänzen und weiterzuentwickeln.

3.2 Die Aushandlungsphase – Konflikt vs. Integration von pädagogischen Leitlinien

> *Na für uns ist ne Identität weggebrochen. Das muss man einfach so sagen. Es wurde zwar offiziell erklärt, die *zweite Fusionsschule* spielt ne große Rolle aber wenn man ehrlich ist, tut sie das nicht mehr. Wir sind entmachtet worden auf (irgendne Art und Weise) […] Ähh also uns ist immer nahe- nachgesagt worden wir sind ne Law-and-Order-Schule, wir sind (.) sehr wirtschaftorientiert, neoliberalistisch, und die anderen Kollegen hier haben wir ein bisschen abfällig als Gutmenschen und Sozialromantiker, als Faulpelze manchmal auch definiert, die ihre Schüler gerne nach der 4. Stunde rausschickten, damit sie nicht mehr wiederkommen, ich sag jetzt einfach mal unsere Vorurteile einfach hinten weg* (Fallstudie I, Interview stellv. Schulleiter *zweite Fusionsschule*, 9/2010, Z. 94-119, gekürzt).

In der Aussage, dass die Identität *„weggebrochen"* sei dokumentiert sich, dass die über einen langen Zeitraum erworbenen Praxen und Arbeitsformen durch die Fusion infrage gestellt werden. Laut den Aussagen des stellvertretenden Schulleiters der *zweiten Fusionsschule* wurde diese *„entmachtet"*. Seiner Macht beraubt worden zu sein, geht mit einer Entwertung des Eigenen einher, womit die Orientierung am Erhalt des Eigenen irritiert wird. Schließlich mündet der Fusionsprozess zunächst in Zuschreibungen negativer Identitätsmerkmale an den jeweils anderen. Eine Verhärtung der Fronten wird durch die Zuspitzung und Festschreibung von pädagogischen Leitlinien in ihrer Begegnung beschrieben, was eine Stagnation im Prozess des Zusammenkommens hervorruft:

> *Völlig stagnierend. Es ist keinem irgendetwas klar. Vor allem konzeptionslos. […] Und es ist aber grundsätzlich das Problem (…) an unserer *integrierten Schule* es gibt keinen Überbau, keinen theoretischen und insofern wuselt jeder vor sich hin und das ist das was mich wahnsinnig macht* (Fallstudie I, Interview Sozialpädagogin, 9/2010, Z. 75-85, gekürzt).

Dieses Zitat verdeutlicht die damals herrschende schulische Realität an der neuen integrierten Schule. Es wird deutlich, dass die Schulfusion die strukturellen Aspekte der schulischen Organisation aktuell unklar erscheinen lässt. Bedingt durch die Unsicherheit innerhalb der Kollegien entsteht eine Irritation von Handlungsroutinen, die sich hier allmählich zur Frustration steigert. Dringender Handlungsbedarf besteht hier einerseits in einer neuen Organisation des Schullalltags, denn die Schulfusion hat durch die Zusammenführung unterschiedlicher Arbeitskonzepte dazu geführt, dass die etablierten pädagogischen und kollegialen Handlungsroutinen infrage gestellt wurden, sodass zurzeit keine klaren Zieldefinitionen, beispielsweise im Kontext der Reformimpulse, verfolgt werden können.

Festzuhalten ist hier, dass durch die Auseinandersetzungen um die Schulfusion ebendiese Zieldefinitionen im Kontext der Schulreform ausbleiben. Vor diesem Hintergrund ist andererseits die Erarbeitung eines gemeinsamen *„theoretischen Überbaus"* dringend nötig. An anderer Stelle im Datenmaterial wird diesbezüglich davon gesprochen, dass es eine *„super Schule werden könnte, wenn man Visionen [hätte]"* (Fallstudie I, Interview Lehrer, 9/2010, Z. 209). Der besagte theoretische Überbau oder die Visionen könnten dabei durch die gemeinsame Kommunikation über die jeweiligen Schulkonzepte erreicht werden:

> *das müsste da müssten wir uns zusammensetzen und nicht streiten sondern da müsste einfach mal jemand sagen wie viel Gedanken stecken hinter diesem Konzept (.) was haben wir verbindlich für uns festgelegt (.) bisher an dieser äh *ersten Fusionsschule* was wollen wir gerne übernehmen und weiterführen (4) […] so können wir zusammen kommen (.) das heißt die Kommuni- es passiert nicht (.) nein das passiert nicht (5) es soll jetzt (.) sich Arbeitsgruppen bilden und zwar sollen die von allen Lehrern besetzt werden nicht nur von denen die (.) in die siebte Klasse sondern von allen Lehrern äh einmal geht's um den Ganztag (.) Rhythmisierung AGs (.) Freizeitgestaltung dann individualisiertes Lernen dann äh duales Lernen und ähm (.) gibt's noch ne Arbeitsgruppe (.) das sind diese (.) vier Hauptpunkte der *integrierten Schule* die (.) Aushängeschilder sozusagen (.) dazu sollen äh AGs gegründet werden was aber im Augenblick nen bisschen fremd is (.) weil ich mich dann frage wo bleibt *Fusionsschule 2* und wo bleibt *Fusionsschule 1*? also (.) wir fangen ja nicht einfach ganz neu an (.) sondern wir ham ja unsere Erfahrung und äh die Arbeit wir bringen doch unsere Arbeit da mit ein das müsste auch stattfinden (3)* (Fallstudie I, Interview Teamleiter Jahrgang 7, 12/2010, Z. 257-288).

Rückblickend auf die ersten Monate nach der Schulfusion wird angegeben, dass konfliktreiche Auseinandersetzungen, sogar Streit das Aufeinandertreffen der beiden Schulen bestimmt haben. Weiterhin bekundet sich hierin einerseits das Interesse an der jeweils anderen pädagogischen Praxis, jedoch hebt diese Aussage auch hervor, dass man bis zu diesem Zeitpunkt, wo es die neue integrierte Schule schon knapp fünf Monate gab, noch keine Kenntnis von den anderen Konzepten und Arbeitsformen sowie deren Begründungszusammenhängen hatte. Begründet wird diese Unwissenheit durch die mangelnde, sogar kaum vorhandene Kommunikation zwischen den Kollegien – wohl wissend, dass man durch Kommunikation zusammenkommen könnte. Innerhalb dieser Äußerung wird erstmals vom Zusammenkommen, Sich-Verbinden, eben vom Fusionieren gesprochen. Das Ziel dieses Austausches wäre die gemeinsame Entwicklung eines neuen Konzepts, welches sich aus den bewährten Praxen und Arbeitsformen der schon bestehenden Konzepte beider Schulen erarbeiten ließe. Deutlich lässt sich hierin eine Entwicklung der Orientierung am Erhalt des bewährten Eigenen feststellen, die durch die Offenheit für eine Integration funktionierender Konzepte der anderen Schule ergänzt wird. Jedoch bestätigt sich weiter-

hin, analog zu den Anfangsaussagen, dass eine konzeptionelle Arbeit im Sinne der Forderung der Schulstrukturreform zugunsten des Wunsches einer kommunikativen Bearbeitung der Schulfusion in den Hintergrund gerät.

Anders wird die Notwendigkeit an der konzeptionellen Arbeit der Schule im Sinne der Reformimpulse von einer anderen Akteursgruppe gesehen, wie im weiteren Verlauf des Diskursabschnitts deutlich wird. Die konzeptuelle Arbeit ist von dieser Seite nicht auf die Zusammenführung von Bestehendem und Bewährtem gerichtet, sondern strebt die Erarbeitung neuer Konzepte vor dem Hintergrund der Schulreformimpulse an, was jedoch innerhalb des Kollegiums mit Skepsis aufgenommen wird. Für den Sprecher ist es befremdlich, völlig neue Konzepte ausarbeiten zu müssen, wo etablierte Arbeitsformen, bewährte pädagogische Praxen und Erfahrungen an beiden Fusionsschulen existieren. Die Entwicklung der Orientierung vom Erhalt des Eigenen hin zur Integration von Eigenem und Anderem findet in den Fragen: „*wo bleibt die *zweite Fusionsschule**“, „*wo bleibt die *erste Fusionsschule**“, ihren Ausdruck.

Zusammengefasst lässt sich hinsichtlich der ersten Fallstudie festhalten, dass die Fusion als krisenhaftes Ereignis erlebt wurde. Dabei wird die Schulfusion vorrangig mit Blick auf die besondere Herausforderung thematisiert, welche mit dem Zusammentreffen zweier in weiten Teilen sehr verschieden operierender Schulen verbunden war. Die Hauptproblematik bestand schließlich darin, dass die pädagogischen Akteure, unabhängig davon, welchem Kollegium sie vor der Fusion angehört hatten, an den jeweiligen Konzepten und Arbeitsformen festhielten. Wie die bisherigen Rekonstruktionen verdeutlichen konnten, ließ der Streit um das „richtige“ oder „bessere“ Schulkonzept Neues nur schwer entstehen, sodass die Bewältigung der Schulfusion im Vorfeld der Schulstrukturreform einen Entwicklungsprozess bezüglich der inhaltlichen Innovationsimpulse der Reform erschwerte.

Auch innerhalb der zweiten Fallstudie wurde, ähnlich dem ersten Fall, die Praxis des Aufeinanderzugehens thematisiert: „*und es geht auch darum dass die Kultur von Kollegen und von Schülern sozusagen auch rüber getragen werden und wir haben drüben dann gesagt […] wir schaffen Zeiten dafür dass sie sich treffen können und es soll ein neues Konzept entstehen (.) hier entsteht ne neue Schule*“ (Fallstudie II, Gruppendiskussion Schulleiter und stellv. Schulleiterin *erste Fusionsschule* und *integrierte Schule*, 5/2011, Z. 835-839, gekürzt). Deutlich wird hierin, dass die gegenseitige Anerkennung der Kolleginnen und Kollegen mit ihren jeweiligen Kompetenzen, das daraus entstandene Kommunikationsverhalten im gesamten Kollegium sowie die Empathie und das Einfühlungsvermögen aller beteiligten Akteure, trotz aller Irritationen und Skepsis, eine gemeinsame Handlungsmaxime sowie ein neues Konzept für die neue *integrierte Schulform* entstehen ließen:

> *also ich glaube es ist absolut wichtig is wenn Schulen fusionieren dass so etwas läuft was wir gemacht haben also dass es auf Augenhöhe ein Austausch ist und dass man sich informieren kann was wir auch gemacht haben indem wir den halben Tag in der einen Schule verbracht haben den anderen an der anderen Schule also wirklich sehr sehr ausgeglichen versucht diese Dinge wirklich so auszutragen und nicht zu sagen ihr seid weniger und ihr könnt das sowieso nicht, sondern wirklich von vornherein es deutlich zu machen wir wollen voneinander profitieren* (Fallstudie II, Gruppendiskussion Schulleiter und stellv. Schulleiterin *erste Fusionsschule*, 5/2011, Z. 774-781).

Weiterhin wurde die Fusion innerhalb der zweiten Fallstudie *„von vornherein"* als Chance begriffen, voneinander und miteinander zu lernen. Kollegialität entstand dabei auf der Basis des gemeinsamen Arbeitens und nicht durch rationale Überzeugungsarbeit, wie es im Falle beider Fusionsschulen der ersten Untersuchungsschule versucht wurde. Beispiele dafür sind der Besuch der jeweils anderen Schule sowie der Austausch und die Würdigung der jeweils anderen etablierten Praktiken, was wiederum zur Herausbildung kollektiver Deutungen und Werte im neuen Kollegium führte. Durch den angesprochenen *„Austausch auf Augenhöhe"* wird hier die schon zuvor angedeutete Bereitschaft zur kollektiven Auseinandersetzung über Werte, Ziele und Methoden sowie zum Abgleich des eigenen Handelns mit dem der neuen Kolleginnen und Kollegen besonders deutlich.

So kann für die zweite Fallstudie festgehalten werden, dass die Zusammenlegung beider Schulen allen beteiligten Akteuren ein hohes Maß an Sozialkompetenz sowie einen hohen kommunikativen Aufwand abverlangt hat. Der von den Beteiligten wahrgenommene Erfolg der Fusion basiert auf wechselseitigem Vertrauen, Unterstützung, Solidarität und Partnerschaft bei der Gestaltung der Beziehungen sowohl zwischen Vorgesetzten und Mitarbeitern und Mitarbeiterinnen als auch zwischen Kolleginnen und Kollegen. Grundvoraussetzungen dafür bestanden in der Bereitschaft und im Engagement aller Beteiligten, ihre Einstellungen und Verhaltensweisen zu reflektieren und gegebenenfalls auch zu modifizieren.

Unter Rückgriff auf die soeben rekonstruierten Orientierungen ist zusammenfassend festzuhalten, dass Schulfusionen, neben einer zeitgleich eingeführten Schulstrukturreform, zusätzliche Anforderungen an alle beteiligten Akteure stellen sowie nichtintendierte Nebeneffekte, welche den schulischen Alltag stark belasten, auslösen. So sehen sich die Schulen zusätzlich zu den neuen pädagogischen und schulpolitischen Herausforderungen mit deutlich wahrnehmbaren Ängsten vor Identifikationsverlust durch die Fusionsbedingungen, an welchen sie keinerlei Mitspracherecht haben, konfrontiert. Die daraus resultierenden negativen Zuschreibungen sowie das beharrliche Festhalten an den eigenen, nach Meinung der Akteure jeweils besseren pädagogischen Konzepten, mit welchen sich die Fusionsschulen im ersten Untersuchungsfall gegenseitig begegneten, verschärfen einen kon-

flikthaften Fusionsverlauf, gewähren dadurch den von Seiten der Bildungspolitik initiierten strukturellen Novellierungen nur schwer Eingang in die schulische Realität und lähmen damit einen Schulentwicklungsprozess im inhaltlichen Sinne der Strukturreform.

Umgekehrt ist jedoch festzuhalten, dass das Ausbleiben einer negativen Bewertung der jeweils anderen Fusionsschule und die Bereitschaft der Akteure, ihre Einstellungen und Konzepte zu durchdenken und gegebenenfalls zu ändern, wie es an der zweiten Untersuchungsschule der Fall war, einen positiv verlaufenden Fusionsprozess markieren. Hierbei ist allerdings auf die strukturellen Unterschiede innerhalb der Fusionsbedingungen an der zweiten Untersuchungsschule hinzuweisen. So wurde mit der *zweiten Fusionsschule* der zweiten Fallstudie ein sehr kleiner Teil des Lehrerkollegiums und der Schülerschaft fusioniert und nicht, wie an der ersten Untersuchungsschule, das gesamte Kollegium inklusive der Schülerschaft. Weiterhin operierte die Schulleitung auch nicht als zusammengelegtes Management im Vergleich zur ersten Fallstudie, wo man sich entschied, als kooperative Schulleitung zusammenzuarbeiten, woraus zusätzliches Konfliktpotenzial entstand. Schließlich galt der Erhalt des Schulgebäudes der *zweiten Fusionsschule* an der zweiten Untersuchungsschule als identitätserhaltender Aspekt innerhalb des Fusionsprozesses.

4. Diskussion

Die beiden in diesem Beitrag betrachteten, sehr unterschiedlichen Verläufe von Schulfusionen weisen auf grundsätzliche Bedingungen der Realisierung von Schulzusammenlegungen hin und lassen damit erste Befunde zu einem noch sehr lückenhaften Forschungsfeld im Bereich der Steuerung im Bildungssystem zu. Mit Bezug auf die eingangs vorgestellte und an wirtschaftswissenschaftliche Modelle zu Fusionsprozessen angelehnte Systematik der Fusionsprinzipien nach Tischendorf (2007) lassen sich die in diesem Beitrag rekonstruierten Verläufe von Schulfusionen im Falle der ersten Fallstudie annähernd dem Assimilationsmodell und im Fall unserer zweiten Fallstudie einem Integrationsmodell zuordnen. Für assimilativ verlaufende Schulfusionen, wie in jenem präsentierten Fall nachgezeichnet, wo eine der fusionierenden Institutionen die neu entstehende Schule massiv dominiert, deuten sich in den dargestellten Befunden folgende strukturelle Bedingungen an: Auf der Ebene der Organisation der Schulfusion sind diese einerseits gekennzeichnet durch die Aufgabe zentraler Aspekte der Identität einer der Fusionsschulen, wie Schulleitung, Schulgebäude und Merkmale der Organisation des Schulalltags, und andererseits durch den konflikthaften Verlauf bzw. das Fehlen von Aushandlungsprozessen über die Realisierung im Vorfeld der Fusion. Auf der Ebene der im empirischen Teil des Beitrags rekonstruierten professionellen Orientierungen und Handlungspraktiken dokumentieren sich Überlegenheitsvorstellungen in den Kollegien beider Fusions-

schulen bei einem Erfahrungsvorsprung im Hinblick auf aktuelle bildungspolitische Erfordernisse durch eines der fusionierenden Kollegien.

Für die in der zweiten Fallstudie präsentierte, integrativ verlaufende Schulfusion, die ebenfalls durch die Dominanz einer Fusionsschule geprägt ist, wird auf der Ebene organisatorischer Bedingungen des Verlaufs die Aufgabe von Aspekten der Schulidentität einer Schule, hier der Schulleitung, durch die sorgfältige Realisierung von konsensuell verlaufenden Aushandlungsprozessen im Vorfeld der Fusion aufgefangen. Auf der Ebene der professionellen Orientierungen ist damit die grundsätzliche Anerkennung pädagogischer Praktiken beider Fusionsschulen verbunden, die gleichsam existierende Vorstellungen der Überlegenheit und des Erfahrungsvorsprungs aufhebt.

Diese ersten Hinweise auf die Bedingungen eines erfolgreichen bzw. weniger erfolgreichen Verlaufs von Schulfusionen sind insbesondere dann von steuerungspolitischer und schulpädagogischer Relevanz, wenn, wie aktuell in der Regel realisiert, Schulzusammenlegungen gleichzeitig mit anderen schulstrukturellen Reformimpulsen eingeführt werden. Die vorgestellten Befunde bestätigen die ohnehin verbreitete Annahme, dass konfliktreich verlaufende Fusionsprozesse die Implementation von Reformimpulsen sowie Schulentwicklungsprozessen der neugegründeten Schulen lähmen, während konsensuell verlaufende Fusionen diese eher zu begünstigen scheinen.

Vor dem Hintergrund der soeben diagnostizierten Belastungen von Schulen durch Schulfusionen, die von der bildungspolitischen Steuerungsebene zusätzlich zu anderen Schulentwicklungsaufgaben vorgegeben werden, findet die Aussage einer Schulverwaltungsbeamtin durchaus ihre Berechtigung: *„Das heißt natürlich schon auch dass wenn eine Stadt oder ein Stadtteil sich entschieden hat neben der Schulstrukturreform Schulen die Aufgabe zu geben auch noch zu fusionieren ich halte das für eine große zusätzliche Aufgabe (.) ich finde man kann auch diskutieren ob das Sinn macht (.) Da ist einfach 'ne gewisse Leistungs- und auch Schulentwicklungsgrenze erreicht* (Gruppendiskussion Schulverwaltung, 10/2011, Z. 281-334, gekürzt). Deutlich wird hierin allerdings auch, dass die Kommunikation zwischen den Ebenen (Makro-, Meso- und Mikroebene) des Bildungssystems hinsichtlich des (für Einzelschulen zumutbaren) Möglichen auszubauen ist. Das Zitat sowie die in dem Beitrag geleisteten Rekonstruktionen lassen den Schluss zu, dass die Entwicklungsgrenzen von Schulen erreicht sind, wenn zusätzlich zu Schulstrukturreformen Schulfusionen veranlasst werden. Ganz im Sinne der Selbststeuerung von Schulen sollte ein erster Schritt im Zuge des demografischen Wandels und der sinkenden Schülerzahlen sein, die Notwendigkeit von Schulfusionen von Seiten der Bildungspolitik (Makroebene) an die Einzelschulen (Mesoebene) zu kommunizieren und Handlungsbedarf zu verdeutlichen.

Um tiefergehende, alle Ebenen des Bildungssystems zusammendenkende Handlungsempfehlungen geben zu können, sind weitere Forschungen notwendig. Aktuell mangelt es an Daten und Befunden zur Quantität sowie zum Verlauf von Schulentwicklungsprozessen während und nach Schulfusionen, zu den Bedingungen ihrer Realisierung auf den Ebenen der Organisation sowie der schulischen Mikroprozesse und schließlich zu ihren Folgen für die Segregation von Schülerschaften sowie im Bereich der Schülerleistungen.

Literatur und Internetquellen

Bohnsack, R. (2003): Rekonstruktive Sozialforschung. Einführung in qualitative Methoden. Weinheim: Beltz.

Bohnsack, R./Nentwig-Gesemann, I./Nohl, A.-M. (Hrsg.) (2001): Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. Opladen: Leske + Budrich.

Bohnsack, R./Pfaff, N. (2010): Die dokumentarische Methode: Interpretation von Gruppendiskussionen und Interviews. In: Maschke, S./Stecher, L. (Hrsg.): Enzyklopädie Erziehungswissenschaft Online (EEO). URL: http://www.erzwissonline.de/fachgebiete/methoden_erziehungswissenschaftlicher_forschung/beitraege/07100073.htm; Zugriffsdatum: 18.07.2013.

Buchen, H./Horster, L./Rolff, H.-G. (Hrsg.) (2009): Schulverbünde und Schulfusion – Notlösung oder Impuls? Stuttgart: Raabe.

Budde, H. (2007): Die Entwicklung regionaler Schulstrukturen in peripheren ländlichen Räumen unter dem Paradigma demographischer Schrumpfung. In: Zeitschrift für Pädagogik 53, S. 314-325.

Fleischer, T. (2009): Fusionierung von Regelschulen in Thüringen – Eine Herausforderung für alle Beteiligten. In: Buchen, H./Horster, L./Rolff, H.-G. (2009): Schulverbünde und Schulfusion – Notlösung oder Impuls? Stuttgart: Raabe, S. 41-52.

Giessner, S.R./Viki, G.T./Otten, S./Terry, D.J./Täuber, S. (2006): The Challenge of Merging: Merger Patterns, Pre-Merger Status and Merger Support. In: Personality and Social Psychology Bulletin 32, S. 339-352.

Marks, M.L./Mirvis, P.H. (2001): Making Mergers and Acquisitions Work: Strategic and Psychological Preparation. In: Academy of Management Executive 15, S. 80-94.

Meister, G. (2012): Kooperation in fusionierten Kollegien. In: Baum, E./Idel, T.-S./Ullrich, H. (Hrsg.): Kollegialität und Kooperation in der Schule. Wiesbaden: VS, S. 167-178.

Mottola, G.R./Bachman, B.A./Gaertner, S.L./Dovidio, J.F. (1997): How Groups Merge: The Effects of Merger Integration Patterns on Anticipated Commitment to the Merged Organization. In: Journal of Applied Social Psychology 27, H. 15, S. 1335-1358.

Müller, H. (2010): Schulleitung im Reformprozess. Schulleiter ohne Stellung. In: b:sl, Beruf: Schulleitung, H. 2, S. 15-17.

Nohl, A.-M. ([2]2007): Komparative Analyse: Forschungspraxis und Methodologie dokumentarischer Interpretation. In: Bohnsack, R./Nentwig-Gesemann, I./Nohl, A.-M. (Hrsg.): Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. Erweiterte Auflage. Wiesbaden: VS, S. 255-276.

Schweiger, D.M./Weber, Y. (1989): Strategies for Managing Human Resources during Mergers and Acquisitions: An Empirical Investigation. In: Human Resource Planning 12, H. 2, S. 69-86.

Tischendorf, K. (2007): Problematische Intergruppen-Beziehungen nach Schulfusionen: Die Rolle relativer Eigengruppen-Prototypikalität in Abhängigkeit von der Bewertung der fusionierten Gruppe. Dissertation, Friedrich-Schiller-Universität Jena. URL: http://www.db-thueringen.de/servlets/DerivateServlet/Derivate-11138/Tischendorf/Dissertation.pdf; Zugriffsdatum: 18.07.2013.

Weishaupt, H. (2002): Demographie und Schulentwicklung in den neuen Ländern. In: Döbert, H./Fuchs, H.-W./Weishaupt, H. (Hrsg.): Transformation in der ostdeutschen Bildungslandschaft. Eine Forschungsbilanz. Opladen: Leske + Budrich, S. 51-62.

Weishaupt, H./Zedler, P. (1994): Aspekte der aktuellen Schulentwicklung in den neuen Ländern. In: Rolff, H.-G./Bauer, K.-O./Klemm, K. (Hrsg.): Jahrbuch der Schulentwicklung, Bd. 8. Daten, Beispiele und Perspektiven. Weinheim/München: Juventa, S. 395-429.

Laura Fölker, geb. 1981, wissenschaftliche Mitarbeiterin am Arbeitsbereich Ethnische Differenzierung und Heterogenität am Institut für Pädagogik der Universität Duisburg-Essen.
E-Mail: laura.foelker@uni-due.de

Anschrift: Universität Duisburg-Essen, Campus Essen/Weststadttürme, Fakultät für Bildungswissenschaften, Institut für Pädagogik, Berliner Platz 6-8, 45127 Essen

BERICHT

DDS – Die Deutsche Schule
Beiheft 12, S. 276-284
© 2013 Waxmann

Felicitas Thiel/Katja Thillmann/Jasmin Tarkian

Die Förderung des *Practice Impact* Empirischer Bildungsforschung als Herausforderung für die Koordination von Forschungsverbünden

Bericht der Koordinierungsstelle des Forschungsverbunds „Steuerung im Bildungssystem“

The Advancement of the Practice Impact of Empirical Educational Research as Challenge for the Coordination of Collaborative Research Projects

Report from the Coordination Board of the Collaborative Research Project "Governance in the Educational System"

Der Ausbau der Empirischen Bildungsforschung verdankt sich nicht allein innerwissenschaftlichen Antriebsdynamiken, sondern ebenso sehr einer gezielten Forschungspolitik, die mit der deutlichen Steigerung der Investitionen in Empirische Bildungsforschung auch konkrete Nutzen- und Verwertungserwartungen verbindet. Dass Nutzenerwartungen und wissenschaftliche Autonomieansprüche nicht selten konfligieren, kann an den Debatten um die Ausrichtung der Bildungsforschung auf *„What Works“* in Großbritannien und den USA abgelesen werden (vgl. Biesta 2007).

In diesem Beitrag werden vor dem Hintergrund der Rekonstruktion unterschiedlicher Motive des Ausbaus Empirischer Bildungsforschung (1) und der Darstellung prinzipieller Konflikte, die im Anspruch einer gleichzeitigen Maximierung von *Science Impact* und *Practice* bzw. *Policy Impact* angelegt sind (2), Möglichkeiten und Grenzen einer Förderung der praktischen Nutzung von Forschungsbefunden durch Einrichtungen der Forschungskoordination am Beispiel der Koordinierungsstelle des

vom Bundesministerium für Bildung und Forschung geförderten Forschungsverbunds „Steuerung im Bildungssystem“ (SteBis) diskutiert (3).

1. Forschungspolitische Motive des Ausbaus Empirischer Bildungsforschung

Der Ausbau der Empirischen Bildungsforschung durch eine gezielte Förderung thematisch fokussierter Verbundforschung ist in der Bundesrepublik Deutschland in den letzten zehn Jahren durch die Forschungspolitik stark vorangetrieben worden. Dabei kamen unterschiedliche Motive zum Tragen. Zum einen machten die internationalen Schulleistungsvergleichsstudien nicht nur deutlich, dass bezüglich der sprachlichen, mathematischen und naturwissenschaftlichen Bildung die Leistungen deutscher Schülerinnen und Schüler deutlich hinter dem zurückblieben, was eine der führenden Industrienationen von sich erwartet, sondern auch, dass – aufgrund der spezifisch deutschen Tradition pädagogischer Wissenschaft (vgl. dazu Thiel 2006) – die wissenschaftliche, insbesondere methodologische Kompetenz zur Modellierung und Messung dieser Leistungen sich auf wenige, meist außeruniversitäre Forschungseinrichtungen konzentrierte. Zum anderen wurde zunächst insbesondere in den angelsächsischen Ländern – in England durch New Labour, in den Vereinigten Staaten von Amerika durch die konservative Bush-Regierung – und nachfolgend auch in anderen Ländern ein neues Modell politischer Entscheidungsfindung etabliert, das sich in erster Linie auf wissenschaftliche Evidenz stützt (vgl. Howlett 2009). Beide Motive trugen dazu bei, dass in der Bundesrepublik Deutschland mit dem Rahmenprogramm Empirische Bildungsforschung derzeit neben dem Nationalen Bildungspanel sieben Forschungsschwerpunkte gefördert werden.

Die übergreifenden Ziele des Rahmenprogramms bestehen in der „originär wissenschaftlichen Aufgabe der Erkenntnisgenerierung im Sinne verallgemeinerbarer und geprüfter Theorien“ sowie darin, „empirisch belastbare wissenschaftliche Informationen zu generieren, die eine rationale Begründung bildungspraktischer und bildungspolitischer Entscheidungen ermöglichen“ (Bundesministerium für Bildung und Forschung 2007, S. 4).

Von Anfang an haben die Promotoren des Rahmenprogramms Empirische Bildungsforschung mit der erklärten Absicht einer strukturellen Förderung Empirischer Bildungsforschung betont, dass die Maximierung des *Science Impact* eine unverzichtbare Voraussetzung evidenzbasierter Bildungspolitik darstellt. Anders formuliert: Allein eine an innerwissenschaftlichen Gütekriterien gemessene, leistungsstarke Bildungsforschung kann die Qualität der Daten sicherstellen, die dem Rationalitätsanspruch einer evidenzbasierten Politik genügen.

Maßnahmen zur Maximierung des *Science Impact* sind einerseits die Einbindung von Wissenschaftlerinnen und Wissenschaftlern in Prozesse der Entwicklung von Forschungsagenden und der Begutachtung von Forschungsanträgen. Andererseits wurde für jedes Verbundforschungsvorhaben eine Koordinierungsstelle eingerichtet, die Aufgaben der nationalen und internationalen wissenschaftlichen Vernetzung und einer koordinierten Nachwuchsförderung übernehmen sollte. Defila, di Giulio und Scheuermann (2006) haben in einer Studie Anforderungen an das Management von Forschungsverbünden identifiziert, die im Wesentlichen auch für die Koordinierungsstellen des Rahmenprogramms Empirische Bildungsforschung gelten: Unterstützung der Einigung auf übergreifende Ziele und Fragestellungen, Vernetzung der Forschungsprojekte und Abstimmung der Forschungsbeiträge im Hinblick auf die Gesamtziele des Verbunds, Synthese oder „Zusammenführung der Einzelprojekte" mit dem Ziel, „neues Wissen zu produzieren" (S. 117), Entwicklung gemeinsamer Produkte, definiert als für Dritte – im oder außerhalb des Wissenschaftssystems – nutzbare Forschungsergebnisse, Komposition von Einzelprojekten zu Arbeitsteams mit Blick auf die wechselseitige Nutzung und koordinierte Entwicklung von Kompetenzen, Gewährleistung des Kommunikationsflusses zwischen den Einzelprojekten und Kommunikation mit der fachlichen und interessierten Öffentlichkeit sowie schließlich Planung, Strukturierung und Organisation der Arbeit im Hinblick auf Ziele und Ressourcen.

Im Rahmenprogramm Empirische Bildungsforschung erfolgte die Einsetzung von Koordinierungsstellen explizit mit dem Auftrag, nicht nur den *Science Impact*, sondern auch den *Practice* bzw. *Policy Impact* der Verbundprojekte zu fördern (vgl. exemplarisch für den Fall der Rehabilitationswissenschaft: Herdt/Bengel/Jäckel 2003). Die deutsche Forschungspolitik hatte mit dem Rahmenprogramm Empirische Bildungsforschung von vorneherein auch die Verwertungs- und Nutzungsinteressen politischer und professioneller Akteure im Blick.

2. *Science Impact* versus *Practice Impact* von (Bildungs-)Forschung

Dass diese doppelte Zielstellung Konfliktpotential birgt, liegt auf der Hand. Sie berührt grundsätzliche Differenzen zwischen Wissenschaft und Politik, die bereits in den späten 1980er-Jahren durch eine Gruppe von Soziologen um Ulrich Beck in verschiedenen Untersuchungen beschrieben worden sind. Das Verhältnis von Wissenschaft und Politik sei, so Beck und sein Kollege Wolfgang Bonß in einer Anspielung auf den Positivismusstreit in der deutschen Soziologie, weder mit dem Begriff der Sozialtechnologie noch mit dem Begriff der Aufklärung hinreichend bestimmt (vgl. Beck/Bonß 1989). Vielmehr seien unterschiedliche Logiken des Umgangs mit Wissen in Rechnung zu stellen. Im Anschluss an Beck und Bonß lassen sich Wissenschaft und Politik sowohl hinsichtlich der Selektion und Bewertung von Wissen als

auch hinsichtlich der Bedingungen der Wissensproduktion bzw. -nutzung unterscheiden: Während Wissenschaft Wissen streng nach Maßstäben wissenschaftlicher Geltung selegiert und bewertet, ist für die politische Rezeption in erster Linie die praktische Verwertbarkeit maßgeblich. Im Hinblick auf die Wissensproduktion ist für die Wissenschaft das Prinzip der Handlungsentlastung – und damit ein nicht vorhersehbarer Zeitbedarf – konstitutiv, während Politik unter hohem Zeitdruck und praktischem Entscheidungszwang agiert. Diese strukturellen Differenzen zwischen theoretischem und praktischem Wissen und die Autonomie der sozialen Systeme Wissenschaft und Politik (vgl. Luhmann 1986) erklären die von Beck und Bonß in zahlreichen Studien rekonstruierten Transferprobleme sowie die eigensinnige Verwendung wissenschaftlichen Wissens durch politische Akteure, die „vom ‚Ergebnispflücken und -picken' bis zur Umkehrung und Unkenntlichmachung soziologischer Ergebnisse im Zuge ihrer Nutzung" (Beck/Bonß 1989, S. 10) reiche. Was von der Wissenschaftssoziologin Karin Knorr-Cetina (1977) noch als symbolische und instrumentelle Nutzung von Forschungsergebnissen kritisiert wurde, wird von der durch den Neoinstitutionalismus geprägten Bildungssoziologin Cynthia Coburn als „*Sensemaking Process*" (Coburn 2004) rekonstruiert. *Sensemaking* erscheint dabei als notwendige Voraussetzung von Transfer und nicht als falsche oder mangelhafte Anwendung. Prozesse des *Sensemaking* sind aus diesem Grund in jedem Transferprojekt in Rechnung zu stellen und zu gestalten.

Aufschlussreich für das Verständnis des Verhältnisses von Wissenschaft und Politik ist die Bilanzierung der Erfahrungen mit der Forschungspolitik der Blair-Regierung durch Geoff Whitty, den damaligen Präsidenten der *British Educational Research Association* (BERA). Whitty nennt in einem Beitrag zu dem von Lesley Saunders herausgegebenen Band mit dem Titel „Educational Research and Policy-Making" mehrere Beispiele für Fehlinterpretation und Missbrauch erziehungswissenschaftlicher Befunde durch Politik und kommt zu dem Schluss: „New Labour has sometimes used research quite selectively and has not appeared particularly concerned about quality of research as long as it serves policy purposes" (Whitty 2007, S. 8) Auf der anderen Seite von Whittys Bilanz stehen allerdings auch positive Beispiele der Verwendung erziehungswissenschaftlicher Forschung, wie die Rezeption von Befunden zu *Assessment* und *Accountability* und unterschiedlichen Befunden zur frühkindlichen Erziehung.

Zwei grundlegende Schwierigkeiten der Übersetzung wissenschaftlicher Befunde in politische Entscheidungen lassen sich vor dem Hintergrund der vorausgehenden Ausführungen unterscheiden:

- Nutzungsbarrieren, die sich fehlenden Kompetenzen oder mangelnden Ressourcen zur Interpretation empirischer Befunde bei den Akteuren in Politik und Praxis verdanken. Diese Barrieren resultieren in falschen Interpretationen oder in der Ignoranz wissenschaftlicher Befunde;

- einseitige oder verzerrte Rezeptionen wissenschaftlicher Befunde, die aus der Logik des politischen „*Sensemaking*" und der Eigenrationalität politischer Entscheidungsprozesse zu erklären sind.

Wie kann unter diesen Voraussetzungen der *Practice Impact* wissenschaftlichen Wissens maximiert werden? Neben der Frage der Definition von Forschungsagenden ist hier vor allem die Frage aufgeworfen, durch welche konkreten Maßnahmen Barrieren der Nutzung wissenschaftlichen Wissens gesenkt und seine Rezeptionschancen erhöht werden können.

- Auf der Seite der Politik wird vor allem *Capacity Building* im Hinblick auf den Umgang mit Daten und eine entsprechende Bereitstellung von Ressourcen für systematische Analysen angemahnt (vgl. Howlett 2009, S. 162). In einem OECD-Bericht zu Forschungs- und Entwicklungsaktivitäten im Bildungsbereich in Großbritannien werden unterschiedliche Initiativen zur Unterstützung von Politik im Umgang mit wissenschaftlicher Evidenz beschrieben. Dazu gehört neben der Bereitstellung von Zeit und Unterstützungssystemen die Entwicklung einer „‚culture' of evidence-based policy" (OECD Review 2002, S. 730). Für die Bundesrepublik Deutschland kommt eine umfangreiche Untersuchung zur Verarbeitung empirischer Evidenz aus Leistungsvergleichsstudien durch Bildungspolitik und Bildungsadministration verschiedener Bundesländer zu dem Ergebnis, dass seit der Veröffentlichung der ersten PISA-Ergebnisse eine Routinisierung und Professionalisierung im administrativen Umgang mit empirischen Befunden zu verzeichnen sei (vgl. Dedering 2009, S. 494), mithin *Capacity Building* in diesem Bereich einer evidenzbasierten Politik offensichtlich bereits erfolgt ist.
- Auf der Seite der Forschung kann durch eine zielgruppenspezifische Aufbereitung und eine konzentrierte Dissemination von Befunden die Nutzung wissenschaftlicher Evidenz durch bildungspolitische und administrative Akteure erheblich erleichtert werden (vgl. Sylva 2004). Eine gleichzeitige Dissemination der Befunde in der Öffentlichkeit könnte die oben beschriebene selektive und instrumentelle Nutzung wissenschaftlicher Evidenz durch Politik und Administration beschränken, weil eine informierte Öffentlichkeit den Rechtfertigungsdruck und die Begründungspflicht politischer Entscheidungen erhöht. Ganz in diesem Sinn sieht Geoff Whitty in einer entsprechenden Stärkung der „public mind on education" eine Möglichkeit zur Korrektur von „superficial but seemingly attractive policies" (2007, S. 13).

3. Steigerung des *Practice Impact* im Rahmen der Forschungskoordination von SteBis

Was können Einrichtungen der Forschungskoordination zur Steigerung des *Practice Impact* beitragen und wo bestehen Grenzen ihrer Wirksamkeit?

Diese Frage wollen wir abschließend am Beispiel des Forschungsverbunds „Steuerung im Bildungssystem" (SteBis) diskutieren. Der Forschungsverbund „Steuerung im Bildungssystem" stellt einen besonderen Fall des Zusammenwirkens von Empirischer Bildungsforschung und Bildungspolitik dar, weil hier das Modell der evidenzbasierten Bildungspolitik mit seinen wichtigsten Instrumenten – Vergleichsarbeiten, zentrale Abschlussprüfungen und Schulinspektionen – selbst Gegenstand der Forschung ist. Wie die mit Hilfe dieser Instrumente erzeugten Daten durch Bildungspolitik, Bildungsadministration und Schulen verarbeitet werden, welche Faktoren Rezeption und Nutzung begünstigen, einschränken oder gar verzerren und welche Nebenfolgen sowie Kosten dieses Modell evidenzbasierten Entscheidens verursacht, sind Fragen, die in den einzelnen Projekten des Forschungsverbunds untersucht werden. Ähnlich wie bei der Rückmeldung von Evaluationsdaten an Schulen ist bei der Rückmeldung von Forschungsbefunden zur Steuerungswirkung evidenzbasierter Bildungspolitik an politische Steuerungsakteure mit besonderen Empfindlichkeiten zu rechnen, insbesondere dann, wenn die Befunde auf ausbleibende oder nichtintendierte Wirkungen hinweisen. In welcher Form die Befunde an Vertreter der Bildungspolitik und -administration adressiert werden, verdient in diesem Forschungsfeld besondere Sensibilität.

Die Koordinierungsstelle des Forschungsverbunds setzt neben Instrumenten zur Erhöhung des *Science Impact,* wie der Organisation von Qualifizierungsangeboten für Nachwuchswissenschaftlerinnen und -wissenschaftler, der Unterstützung der Vernetzung mehrerer Forschungsprojekte sowie der Unterstützung gemeinsamer Symposien und Arbeitsgruppen bei nationalen und internationalen Kongressen, auch folgende Instrumente zur Optimierung des *Practice Impact* ein:

- Vorstellung der Arbeit des Forschungsverbunds auf nationalen und internationalen Fachtagungen von Bildungspraktikern und -praktikerinnen und Organisation einer eigenen Fachtagung für Repräsentanten und Repräsentantinnen der Bildungspolitik und Bildungsadministration;
- Organisation von Podiumsdiskussionen mit Repräsentanten und Repräsentantinnen aus Politik und Verbänden sowie Wissenschaftlerinnen und Wissenschaftlern aus den Projekten;
- Bereitstellung von Projektergebnissen des Forschungsverbunds auf einer eigenen Website (URL: www.stebis.de) sowie mittels eines Newsletters für die fachliche und nicht-fachliche Öffentlichkeit;

- Organisation regionaler und überregionaler Pressegespräche bzw. Hintergrundgespräche mit der Presse zur Dissemination der Ergebnisse in der nicht-fachlichen Öffentlichkeit.

Bei der Handhabung dieser Instrumente wird das oben beschriebene Spannungsverhältnis von Wissenschaft und Politik virulent. Die Entwicklung neuer Kommunikationsformen zur Senkung von Rezeptionsschwellen gehört deshalb ebenso zu den zentralen Aufgaben dieses Typs der Forschungskoordination wie die Anregung der Reflexion über Bedingungen und Grenzen der Applikation konkreter wissenschaftlicher Befunde in praktischen Kontexten. Dass der hohe Zeitbedarf und die prinzipielle Ergebnisunsicherheit wissenschaftlicher Forschung für eine unter Entscheidungsdruck agierende Politik manchmal schwer verständlich sind, muss genauso ein Thema der wechselseitigen Verständigung zwischen Wissenschaft und Politik/Praxis sein wie der bisweilen esoterische Charakter wissenschaftlicher Kommunikation.

Die Grenze der Steigerung des *Practice Impact* durch Forschungskoordination liegt bei der Forderung nach einer direkten Übersetzung wissenschaftlichen Wissens in die Praxis (vgl. Shavelson/Towne 2002). Eine Unterstützung von Transfer im Sinne der Implementation von Innovationen bedarf anderer Instrumente, wie zum Beispiel einer wissenschaftlichen Begleitung der Entwicklung und Implementation von Konzepten im Sinne einer „evidenzbasierten Strategie" des Transfers (vgl. Gräsel 2010). Die eingangs beschriebenen *Sensemaking*-Prozesse und die inzwischen in zahlreichen Studien identifizierten Barrieren der Nutzung empirischer Evidenz durch Praktikerinnen und Praktiker müssen in einer intelligenten Transferstrategie antizipiert werden. Die Koordinierungsstelle eines Forschungsverbunds kann allerdings lediglich Vernetzungsaktivitäten von Akteuren der Bildungspolitik und -administration oder von Schulen auf den Weg bringen bzw. unterstützen, indem sie einzelne Projekte des Forschungsverbunds als Impulsgeber für solche Aktivitäten gewinnt. Dies wurde im Rahmen des Forschungsverbunds SteBis für Schulen in schwierigen Lagen realisiert (vgl. die Beiträge von Racherbäumer u.a. sowie Fölker in diesem Band). Mit der Ausrichtung eines Expertenzirkels, innerhalb dessen die Befunde zweier SteBis-Projekte zur Schulentwicklung unter Bedingungen vielfacher Bildungsrisiken vorgestellt und in Workshops mit schulischen Praktikerinnen und Praktikern auf ihre konkreten Verwertungsmöglichkeiten hin diskutiert wurden, hat die Koordinierungsstelle den organisatorischen Rahmen für die Aufnahme eines Dialogs zwischen Wissenschaft und Praxis geschaffen. Eine kontinuierliche Begleitung solcher Aktivitäten geht allerdings über die Möglichkeiten der Forschungskoordination weit hinaus.

Zu beachten ist, dass die Unterstützung evidenzbasierter Entscheidungen im Bildungssystem durch Einrichtungen der Forschungskoordination insofern grundsätzlich beschränkt ist, als die Aufbereitung und Kommunikation empirischer Befunde nicht über den eigenen Forschungsverbund hinausreichen kann. Eine umfassende,

flächendeckende und kontinuierliche Bereitstellung wissenschaftlicher Evidenz kann nur durch forschungsverbundübergreifende, intermediäre Instanzen geleistet werden, die die Nutzungsbarrieren senken, indem sie die gesamte wissenschaftliche Evidenz zu einem Themenfeld zielgruppenspezifisch aufbereiten und damit die politische, professionelle und öffentliche Nutzung erleichtern.

Beispiele für solche intermediären Instanzen sind das 1993 an der University of London gegründete *Evidence for Policy and Practice Information and Co-ordinating Centre* (EPPI-Centre) oder die im Jahr 2002 durch das *Institute of Education Sciences* (IES) des US-Bildungsministeriums gestartete Initiative *What Works Clearinghouse* (WWC). Während das EPPI-Centre mit seiner *Online Evidence Library* systematische Reviews über eine breite Spanne von Themen, unter anderem aus den Bereichen Bildung, Sozialpolitik und Gesundheitsförderung, für Akteure aus Politik und Praxis zugänglich macht, strebt das WWC danach, *die* zentrale Stelle speziell für Fragen der Bildungswissenschaften zu sein. Auf der Website des WWC können Nutzer nach (wirksamen) Interventionen suchen und sich systematisch aufbereitete und kritisch bewertete Forschungsbefunde zu unterschiedlichen Maßnahmen im Bildungsbereich ausgeben lassen. Mit dieser Initiative soll Bildungspraktikern und -praktikerinnen eine Orientierung über die zahlreichen Befunde Empirischer Bildungsforschung erleichtert und damit eine bessere Nutzung für bildungspolitische oder professionelle Entscheidungen ermöglicht werden.

Diese Aktivitäten gehen allerdings über die Aufgaben einer Koordinierungsstelle schon weit hinaus.

Literatur

Beck, U./Bonß, W. (1989): Verwissenschaftlichung ohne Aufklärung? Zum Strukturwandel von Sozialwissenschaft und Praxis. In: Beck, U./Bonß, W. (Hrsg.): Weder Sozialtechnologie noch Aufklärung? Analysen zur Verwendung sozialwissenschaftlichen Wissens. Frankfurt a.M.: Suhrkamp, S. 7-45.

Biesta, G. (2007): Why "What Works" Won't Work: Evidence-Based Practice and the Democratic Deficit in Educational Research. In: Educational Theory 57, H. 1, S. 1-22.

Bundesministerium für Bildung und Forschung (2007): Rahmenprogramm zur Förderung der empirischen Bildungsforschung. URL: http://www.bmbf.de/pubRD/foerderung_der_empirischen_bildungsforschung.pdf; Zugriffsdatum: 22.07.2013.

Coburn, C.E. (2004): Beyond Decoupling: Rethinking the Relationship between the Institutional Environment and the Classroom. In: Sociology of Education 77, H. 3, S. 211-244.

Dedering, K. (2009): Evidence-Based Education Policy: Lip Service or Common Practice? Empirical Findings from Germany. In: European Educational Research Journal 8, H. 4, S. 484-496.

Defila, R./di Giulio, A./Scheuermann, M. (2006): Forschungsverbundmanagement. Handbuch für die Gestaltung inter- und transdisziplinärer Projekte. Zürich: Vdf Hochschulverlag.

Gräsel, C. (2010): Stichwort: Transfer und Transferforschung im Bildungsbereich. In: Zeitschrift für Erziehungswissenschaft 13, H. 1, S. 7-20.
Herdt, J./Bengel, J./Jäckel, W.H. (2003): Wissensmanagement und Verwertung von Forschungsergebnissen in den Rehabilitationswissenschaften. In: Phys Rehab Kur Med 13, H. 4, S. 199-207.
Howlett, M. (2009): Policy Analytical Capacity and Evidence-Based Policy-Making: Lessons from Canada. In: Canadian Public Administration/Administration Publique du Canada 52, H. 2, S. 153-175.
Knorr-Cetina, K. (1977): Producing and Reproducing Knowledge: Descriptive or Constructive? Toward a Model of Research Production. In: Social Science Information 16, H. 6, S. 669-696.
Luhmann, N. (1986): Ökologische Kommunikation. Kann die moderne Gesellschaft sich auf ökologische Gefährdungen einstellen? Opladen: Westdeutscher Verlag.
OECD Review (2002): Educational Research and Development in England: Examiners' Report. In: European Educational Research Journal 1, H. 4, S. 705-738.
Shavelson, R.J./Towne, L. (Hrsg.) (2002): Scientific Research in Education. Committee on Scientific Principles for Education Research. Washington, DC: National Academy Press.
Sylva, K. (2004): Interventions that Work. In: Fox, N. (Hrsg.): Social and Moral Development: Emerging Evidence in the Toddler Years. Princeton: Johnson & Johnson, S. 269-280.
Thiel, F. (2006): Die Etablierung der akademischen Pädagogik als Reflexionstheorie des Erziehungssystems: Ein anderer Blick auf die „Erfolgsgeschichte" der geisteswissenschaftlichen Pädagogik. In: Zeitschrift für Erziehungswissenschaft 9, H. 1, S. 81-96.
Whitty, G. (2007): Education(al) Research and Educational Policy-Making. Is Conflict Inevitable? In: Saunders, L. (Hrsg.): Educational Research and Policy-Making. Exploring the Border Country between Research and Policy. New York: Routledge, S. 1-18.

Felicitas Thiel, Prof. Dr., geb. 1961, Leiterin des Arbeitsbereichs Schulpädagogik/Schulentwicklungsforschung an der Freien Universität Berlin.
E-Mail: felicitas.thiel@fu-berlin.de

Katja Thillmann, Dr., geb. 1978, bis 2013 wissenschaftliche Mitarbeiterin an der Freien Universität Berlin.

Jasmin Tarkian, geb. 1975, wissenschaftliche Mitarbeiterin an der Freien Universität Berlin.
E-Mail: jasmin.tarkian@fu-berlin.de

Anschrift: Arbeitsbereich Schulpädagogik/Schulentwicklungsforschung, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin