

Wirkungen und Wirkungswege einer Bildungsstandards-Reform: Die Sichtweise von Lehrpersonen und Schulleitungen in österreichischen Primarschulen

Manuela Gamsjäger  · Herbert Altrichter  · Regina Steiner

Angenommen: 24. Mai 2019 / Online publiziert: 7. Juni 2019
© Der/die Autor(en) 2019

Zusammenfassung Die Bildungsstandard-Reform in Österreich ist ein Beispiel für aktuelle Governance-Veränderungen in Schulsystemen. In der Logik evidenzbasierter Steuerung setzt sie normative und funktionale Erwartungen in Richtung einer Verbesserung von Unterrichtsprozessen und -ergebnissen durch individualisierenden, kompetenzorientierten Unterricht und durch gezielte Unterrichts- und Schulentwicklung auf der Basis der Rückmeldung von standardorientierten Kompetenztestergebnissen. Der Beitrag prüft durch die Analyse qualitativer Fallstudien dreier Primarschulen, wie Schulleiterinnen und Lehrerinnen diese Bildungsreform rekontextualisieren und inwieweit die durch die Reformpolitik postulierten Wirkungs- und Prozessenerwartungen der Bildungsstandard-Reform, die in Form eines Wirkungsmodells rekonstruiert wurden, in den Primarschulen aufzufinden sind. Dabei zeigt sich, dass im Kontext der Bildungsstandard-Reform in den untersuchten Schulen tatsächlich Veränderungen des Unterrichts unterschiedlicher Qualität stattfanden, die von Anpassung an die neuen Testformate bis zu einer stärkeren Kompetenzorientierung des Unterrichts reichte. Von den – aufgrund des Wirkungsmodells erwarteten – Prozessen des Wirksamwerdens der Reform konnten zwei rekonstruiert werden: Durch die Bildungsstandard-Reform geschieht ein Aufbau von Erwartungen, der durch

M. Gamsjäger

Institut für Sekundarstufenpädagogik, Pädagogische Hochschule Oberösterreich,
Kaplanhofstraße 40, 4020 Linz, Österreich
E-Mail: manuela.gamsjaeger@ph-ooe.at

H. Altrichter (✉)

Linz School of Education, Abt. f. Bildungsforschung, Johannes Kepler Universität, Altenberger
Straße 69, 4040 Linz, Österreich
E-Mail: herbert.altrichter@jku.at

R. Steiner

Institut Elementar- und Primarstufenpädagogik, Pädagogische Hochschule Oberösterreich,
Kaplanhofstraße 40, 4020 Linz, Österreich
E-Mail: regina.steiner@ph-ooe.at

die Bereitstellung von Unterrichtsmaterialien verstärkt wird. Von Datenfeedback, der Einbeziehung von Stakeholdern der Schule sowie von neuen Notwendigkeiten zur innerschulischen Koordination scheinen dagegen weniger Entwicklungsimpulse auszugehen.

Schlüsselwörter Bildungsstandards · Kompetenzüberprüfungen · Datenfeedback · Rekontextualisierung · Unterrichtsentwicklung

How a Performance-Standards Policy Becomes Effective: The Perception of Teachers and School Leaders in Austrian Primary Schools

Abstract The performance standards reform in Austria is an exemplar of contemporary governance changes in school systems. Following the logic of “evidence-based governance” it sets normative and functional expectations for improving the schools’ processes and outcomes through individualized and competence-oriented learning and through well-focused teaching improvement by using data feedback about student competences. The paper uses qualitative case studies of three Austrian primary schools for studying how teachers and school leaders interpret this policy and whether or not the effects and processes proposed by this reform are reflected by their perceptions and experiences. By analyzing qualitative case studies of three primary schools, the article examines how principals and teachers reconstruct this educational reform and to what extent the expected effect and processes (which were reconstructed in a “conceptual model”) are found in the primary schools under study. The results indicate that there were actually changes in teaching (of different quality, ranging from adaptation to the new test formats to a more competence-oriented teaching) in the three case schools. Two of the change processes expected by the “impact model” have been reconstructed: the educational standard reform is “setting expectations”, which are reinforced by the “provision of teaching materials”. By contrast, “data feedback”, “stakeholder involvement” as well as new needs for “in-school coordination” seem to have less developmental momentum.

Keywords Performance standards · Competence assessment · Data feedback · Rekontextualisation · Teaching development

1 Einleitung

Nach dem ‚PISA-Schock‘ wurde versucht, die ‚Governance‘ der meisten deutschsprachigen Bildungssysteme, d.h. die Prinzipien und Prozesse ihrer Steuerung und Gestaltung, einer Modernisierung zu unterziehen, wofür meist Konzepte einer „outputorientierten“ bzw. „evidenzbasierten Steuerung“ als Vorbild dienten (Altrichter et al. 2005). In diesen sollen Steuerungsentscheidungen ‚rationalisiert‘ und Entwicklungsmaßnahmen fokussiert werden: Dazu müssen mindestens (1) die Ziele in klarer und messbarer Form an die Systemakteure kommuniziert, (2) die Zielerreichung in verschiedenen Bildungseinrichtungen in vergleichender Form gemessen und (3) deren Ergebnisse an die Akteure auf verschiedenen Systemebenen zurück-

gespielt werden, von denen (4) Entwicklungsmaßnahmen zur Behebung von Ist-Soll-Diskrepanzen und zur Verbesserung der Systemleistung erwartet werden (Ehren et al. 2013, S. 4). „Neue Schulinspektionen“ sowie Bildungsstandards und darauf bezogene Lernstandserhebungen sind die in den deutschsprachigen Schulsystemen bekanntesten Realisierungsformen dieser evidenzbasierten Steuerungslogik.

Bildungsreformen werden in der Perspektive der Governanceforschung (Schimank 2007) erst sozial bedeutsam, wenn sie von relevanten Akteurinnen und Akteuren auf verschiedenen Ebenen des Bildungssystems rezipiert und „rekontextualisiert“ (Fend 2006) werden, d. h. in Praktiken und Strukturen übersetzt werden, die der Handlungslogik der jeweiligen Ebene entsprechen. Der vorliegende Beitrag verwendet die österreichische Bildungsstandard-Politik als ein Beispiel für ‚evidenzbasierte Reformen‘ im Schulwesen. Uns interessiert, wie diese Reform an österreichischen Primarschulen ‚rekontextualisiert‘ wird und ob diese bildungspolitische Intervention zu solchen Handlungsweisen von Lehrpersonen und Schulleitungen führt, die von den Verfechter/innen der Reform intendiert wurden (vgl. Kap. 2). Nach einem Überblick über den themenbezogenen Forschungsstand (Kap. 3) untersucht der Beitrag an qualitativen Fallstudien (vgl. Kap. 4), ob sich die normativen Erwartungen über Ergebnisse und Wirkungsprozesse einer solchen evidenzbasierten Reform in den Interviewaussagen von Schulleitungen und Lehrpersonen dreier österreichischer Primarschulen widerspiegeln (Kap. 5).

2 Merkmale der österreichischen Bildungsstandard-Politik

Die 2008 in Österreich eingeführte Bildungsstandard-Politik (BiSt-Politik) ist als konkrete Realisierungsform ‚evidenzbasierter Steuerung‘ zu verstehen. Eine Analyse offizieller Darstellungen des Ministeriums und der staatlichen Unterstützungsgagentur BIFIE (Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens) zeigt folgende zentrale *Elemente*: (1) *Bildungsstandards* spezifizieren transparente und vergleichbare Bildungsziele. Sie halten fest, welche Kompetenzen „Schüler/innen bis zum Ende der 4. Schulstufe in Deutsch und Mathematik [...] nachhaltig erworben haben sollen“ (BIFIE 2012). Dadurch sollen sie schulische Akteur/innen für Qualitätskriterien sensibilisieren, zu Qualitätsentwicklung stimulieren und insgesamt eine stärker kompetenz-, ergebnisorientierte und individualisierte Unterrichtskultur befördern (BIFIE 2016). (2) Die darin festgelegten Schülerkompetenzen werden in *periodischen Standardüberprüfungen* alle zwei Jahre abwechselnd für Mathematik (erstmalig 2013) und Deutsch (erstmalig 2015) flächendeckend erhoben. (3) Deren Ergebnisse werden den Akteuren auf allen Systemebenen jeweils in aggregierter Form (z. B. erhalten Lehrpersonen die aggregierten Ergebnisse der eigenen Klasse, Schulleitungen jene der gesamten Schule usw.) und bezogen auf ‚Erwartungswerte‘ (diese berücksichtigen u. a. die soziale Zusammensetzung der Schülerschaft) rückgemeldet. Im österreichischen Standardkonzept erfolgt die Testung jeweils am Ende von Bildungsstufen (also in der 4. Schulstufe der Primarschule¹) und die Ergebnisse werden erst im darauffolgenden Schuljahr

¹ Wir verzichten aus Platzgründen auf die Erläuterung analoger Regelungen für die Sekundarstufe.

rückgemeldet. Dadurch hat die überwiegende Zahl der getesteten Schüler/innen die jeweilige Schule bereits verlassen. (4) Verschiedene *Begleitmaßnahmen zur Implementierung*, wie Informationsangebote, Fortbildungen, Unterrichtsmaterialien und ein standardbezogenes Diagnoseinstrument (*Informelle Kompetenzmessung, IKM*), sollen die „Neuorientierung unterrichtlicher Praktiken“ (BIFIE 2016) erleichtern. Zur Unterstützung bei der Interpretation des Datenfeedbacks können sog. Rückmeldemoderator/innen angefordert werden. (5) Schließlich erfolgt ein *Einbeziehen der Schulpartner/innen*, da die Testergebnisse im Schulpartnerschaftsgremium mit Eltern- und Lehrerververtretungen verpflichtend besprochen werden müssen (BMUKK 2012).

In einer Analyse der österreichischen BiSt-Politik (Altrichter und Kanape-Wilingshofer 2012, S. 359 ff.) wurden zwei allgemeine *Ziele dieser Reform* herausgearbeitet: nachhaltig verbesserte Schülerkompetenzen sowie mehr Chancengleichheit und Bildungsgerechtigkeit. Dabei wurde auch gezeigt, dass Zielbeschreibungen in den offiziellen Dokumenten oft in Form von Prozesszielen erfolgen: So wird von der BiSt-Reform ein Beitrag zur ‚evidenzbasierten Unterrichts- und Schulentwicklung‘ und zum ‚Aufbau einer individualisierenden, ergebnis- und kompetenzorientierten Unterrichtskultur‘ erwartet. Auf diese Prozessziele wird sich die hier vorgestellte Analyse konzentrieren.

3 Forschungsstand

3.1 Untersuchungen zu Wirkungen und Verarbeitungsprozessen

Begleitend zur Einführung verschiedener Versionen von BiSt-Politiken in vielen deutschsprachigen Schulsystemen² wurden in den letzten Jahren auch eine Reihe von Studien durchgeführt, die die Wirksamkeit dieser Innovation thematisieren. Dabei kann man zwei Hauptzielrichtungen der Analyse unterscheiden: Einerseits wird untersucht, ob durch die Formulierung und Kommunikation von Bildungsstandards auch *die darin formulierten Kompetenzen den Unterricht vermehrt leiten. Oder allgemeiner: Ob kompetenzorientierter Unterricht vermehrt umgesetzt* wird bzw. ob entsprechende *Erwartungen aufgebaut* werden. Die Befunde zu diesen Fragen erwiesen sich – verglichen mit den bildungspolitischen Erwartungen – als eher ernüchternd (u. a. Wacker 2008; Grillitsch 2010; Lenski et al. 2015; Pöhlmann et al. 2014; Richter et al. 2014). Für Österreich zeigte sich bereits in der Pilotphase der BiSt-Reform eine Skepsis der beteiligten Lehrpersonen über die Nützlichkeit von Bildungsstandards für den Unterricht; Lehrpersonen (Freudenthaler und Specht 2005, S. 31 f; Grillitsch 2010, S. 85 f), auch Primarschullehrkräfte (Freudenthaler und Specht 2006, S. 17 f), nutzten Bildungsstandards selten für ihre laufende Unterrichtsplanung. Insgesamt boten Bildungsstandards Primarschullehrkräften selten Orientierung und wurden

² Bei der Analyse empirischer Erfahrungen beschränken wir uns auf Studien aus deutschsprachigen Schulsystemen. Obwohl BiSt-Politiken als ein Musterbeispiel einer „travelling policy“ in vielen Ländern zu finden sind, ist nicht anzunehmen, dass sie unter verschiedenen Systembedingungen (z. B. high vs. low stake-Kontexte) in ähnlicher Weise wirken (Ozga und Jones 2006).

eher additiv zum bisherigen Unterricht implementiert (Frühwacht und Maier 2012; Feldhoff et al. 2012; Zuber 2019).

Eine zweite Analyserichtung, zu der weit mehr Studien in den deutschsprachigen Ländern vorliegen, zielt auf die *Rezeption und Nutzung von Datenfeedback*. Die Mehrzahl der Studien (u. a. Nachtigall und Jantowski 2007; Kühle und Peek 2007; Groß Ophoff et al. 2006; Grabensberger et al. 2008) zeigt, dass Datenfeedback – auch wenn es als informativ und nützlich empfunden wurde – nur von einer kleinen Zahl von Lehrpersonen für Unterrichtsentwicklung genutzt wurde. Wenn überhaupt, wurden oft nur ‚oberflächliche Unterrichtsveränderungen‘ vorgenommen, die ohne großen Aufwand umsetzbar sind und nicht als qualitätsverbessernde Schritte zu kompetenzorientiertem Unterrichten interpretiert werden können (Groß Ophoff 2013). Darüber hinaus scheinen Informiertheit über die Neuerung, professionelle Vorstellungen, Art der Implementation und Haltungen gegenüber Evaluationen die Rezeption und Nutzung von Datenfeedback zu moderieren (Kuper und Hartung 2007; Grillitsch 2010).

Insgesamt gibt es wenige Studien, die sich genauer und längerfristig mit den *Prozessen der Rezeption und ‚Rekontextualisierung‘* dieser bildungspolitischen Reform auf der Ebene der Einzelschule beschäftigen. Asbrand et al. (2012, S. 38) haben den Diskurs in Fachkonferenzen untersucht und dabei festgestellt, dass beispielsweise Lehrpersonen mit wenig Vorerfahrung kompetenzorientiert gemeinte Materialien für „inputorientierte Unterrichtsgestaltung“ einsetzen. Eine Veränderung etablierter Muster der Unterrichtsvorbereitung erfordere mehr als das Angebot von Unterrichtsmaterial, nämlich „differenzierte Fortbildungsszenarien“ und „längerfristige Entwicklungsbegleitung“ (Asbrand et al. 2012, S. 40).

In qualitativen Interviews zeigte sich bei Zuber (2019), dass Grundschullehrpersonen auch dann Handlungen zur Implementation der BiSt-Reform setzen, wenn sie dieser negativ oder ambivalent gegenüberstehen. Die Autorin schließt daraus, dass die ‚wahrgenommene Verhaltenskontrolle‘ (die sich einstellt, wenn die Lehrkräfte meinen, über entsprechende Kompetenzen und brauchbares Unterstützungsmaterial zu verfügen) bedeutsamer ist als die positive Bewertung der Reform. Ein weiterer – aufgrund der Theorie von Ajzen und Fishbein (1980) – als handlungsförderlich postulierter Faktor, nämlich subjektive Normen³, taucht praktisch nicht in der Wahrnehmung der Lehrpersonen auf: Schulleitungen, die die Norm ‚Umsetzung der BiSt-Reform‘ verkörpern könnten, üben keine diesbezügliche Orientierungsfunktion für die Lehrkräfte aus (Zuber 2019, S. 98).

3.2 Wirkungswege der BiSt-Reform

Insgesamt wissen wir aber zu wenig über die – postulierten und realen – Prozesse zwischen der BiSt-Reform und den (oft unbefriedigenden) Wirkungen. Husfeldt (2011, S. 259) hat vor einigen Jahren den damals unbefriedigenden Stand der Schulinspektionsforschung kritisiert und „brauchbare Wirkungsmodelle“ gefordert, die es

³ Eine „subjektive Norm“ entsteht aus der Abwägung von Akteuren, „wie für sie bedeutsame Personen die Ausführung des Verhaltens bewerten würden“ (Zuber 2019, S. 109).

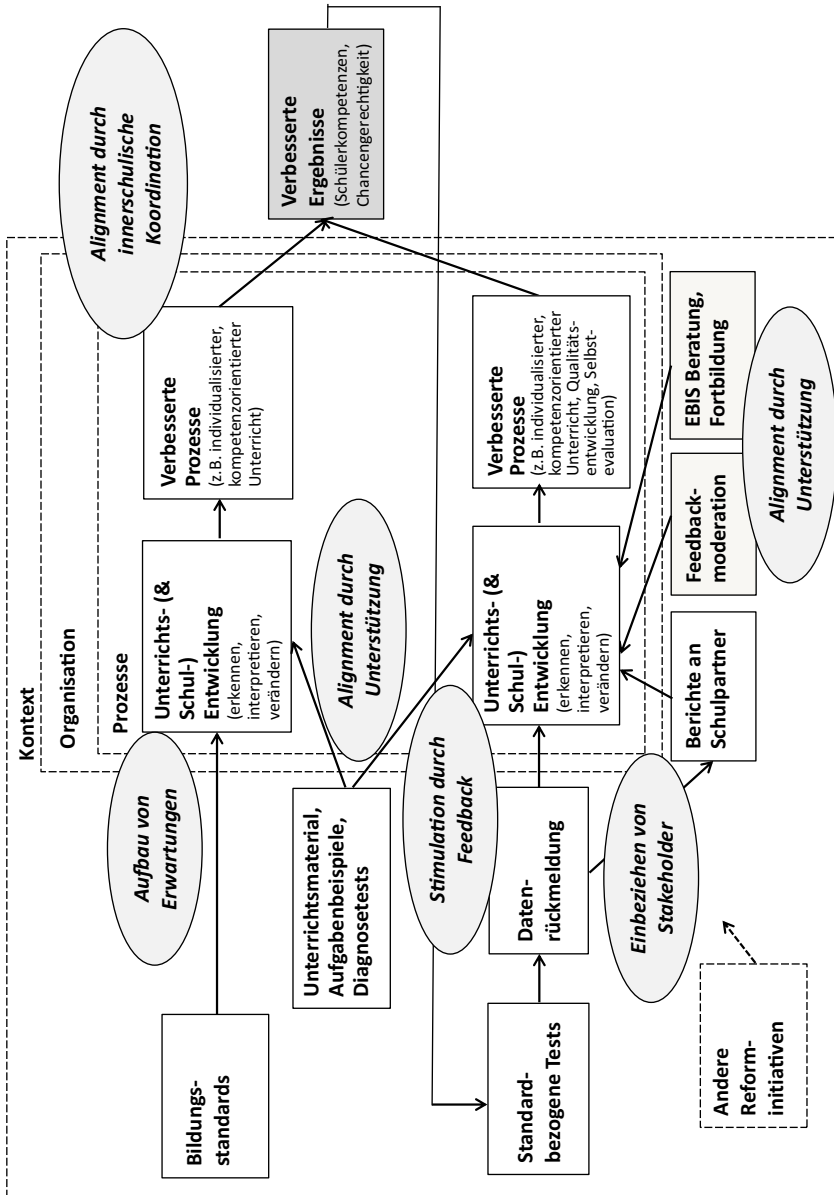


Abb. 1 Konzeptuelles Wirkungsmodell der österreichischen BiSt-Politik (deutsche Übersetzung nach Altrichter und Gamsjäger 2017)

erlauben, „die schulinternen Verarbeitungsprozesse“ zwischen Intervention und erhofften Wirkungen genauer zu untersuchen.

In diesem Sinne soll in diesem Beitrag versucht werden, ein kürzlich veröffentlichtes „Wirkungsmodell“ der österreichischen BiSt-Reform einer ersten Überprüfung zu unterziehen. Einem Vorschlag von Leeuw (2003) und Ehren et al. (2015) entsprechend haben Altrichter und Gamsjäger (2017) durch eine Analyse offizieller Dokumente die normativen (d. h. von der Reformpolitik unterstellten) Wirkungsw-

ge dieser Reform rekonstruiert. Dabei wurden fünf *Prozesse* (vgl. Ellipsen in Abb. 1) auf der Ebene der Einzelschule identifiziert, die als entscheidend für die Erreichung der Ziele der BiSt-Politik anzusehen sind:

1. *Aufbau von Erwartungen*: Bildungsstandards sollen als messbare Kompetenzen ein eindeutiges und transparenteres Referenzsystem für professionelles Handeln und Lernen bieten. Dadurch sollen Erwartungen bei Lehrpersonen in Hinblick auf Ziele, aber auch auf einen stärker kompetenzorientierten, differenzierten und individualisierten Unterricht aufgebaut werden. Dies soll sich in der Unterrichtsplanung und -durchführung der Lehrkräfte und in weiterer Folge in besseren Schülerleistungen und in einer Angleichung der Leistungsanforderungen zwischen Schulen und zwischen Lehrpersonen niederschlagen.
2. *Stimulation durch Datenfeedback*: Die Ergebnisse der periodischen Standardüberprüfungen werden als Datenfeedback den schulischen Akteur/innen mitgeteilt. Insbesondere Lehrpersonen und Schulleitungen sollen diese Ergebnisse lesen, mit den Zielen abgleichen und interpretieren (BIFIE 2012). Dadurch sollen Reflexions- und Entwicklungsprozesse angestoßen werden, die – wo notwendig – zu veränderter Unterrichtsplanung und -gestaltung (und in weiterer Folge zur Erreichung der Reformziele) führen. Für den Fall, dass die Ergebnisse unter oder über den Erwartungswerten liegen, sind keine spezifischen Konsequenzen vorgesehen. Im Rahmen des seit dem Schuljahr 2013/14 eingeführten Qualitätsmanagementsystems Schulqualität Allgemeinbildung (SQA; vgl. Altrichter 2017) kann die Schulaufsicht allerdings die Aufnahme entsprechender Entwicklungsvorhaben in die verpflichtenden Entwicklungspläne der Schulen begehren.
3. *Alignment durch Unterstützung*: ‚Alignment‘ meint die „möglichst kohärente Abstimmung“ (Oelkers und Reusser 2008, S. 432), in diesem Fall der Arbeit der Lehrpersonen und Schulen mit den Zielen der Reforminitiative. Verschiedene Begleitmaßnahmen zur Implementierung zeigen im Falle der BiSt-Politik konkret, wie ihre normativen Erwartungen in Aktionen und Strukturen auf Schul- und Unterrichtsebene übersetzt werden können: was Lehrpersonen tun müssen, wenn sie nach den Intentionen dieser Reform ‚kompetenzorientiert unterrichten‘ wollen; wie die Leistungen ihrer Schüler/innen ‚kompetenzorientiert‘ zu interpretieren sind usw. Weitere Supportangebote (z. B. Rückmeldemoderation, Entwicklungsberatung) sollen die Rezeption und Verarbeitung des Datenfeedbacks unterstützen.
4. *Einbeziehen von Stakeholdern*: Durch das verpflichtende Gespräch über Schulergebnisse in den Schulpartnerschaftsgremien (BMUKK 2012) und deren Thematisierung in den periodischen Zielvereinbarungen mit der regionalen Schulaufsicht sollen Schulleitung und Lehrpersonen zu konsequenter und nachhaltiger Entwicklung im Gefolge des Datenfeedbacks motiviert werden.
5. *Alignment durch innerschulische Koordination*: Darüber hinaus wird erwartet, dass die Umsetzung der BiSt-Politik innerschulische Koordination zwischen Lehrpersonen und mit der Schulleitung intensiviert (z. B. weil Absprachen bei Unterrichtsplanung und Interpretation des Datenfeedbacks notwendig werden). Dadurch soll die Bindung von Pädagoginnen und Pädagogen an die Reform gestärkt werden.

Folgt man den Überlegungen von Ehren et al. (2015), dann sind diese fünf Prozesse des Wirksamwerdens der BiSt-Politik theoretisch nicht unplausibel (vgl. auch Altrichter und Gamsjäger 2017): Die Prozesse (1) und (3) können mittels neo-institutionalistischer Theorien (Scott 2001) erklärt werden: Für Entscheidungen in Organisationen sind nicht allein Effizienzkriterien leitend; vielmehr soll auch Legitimität durch die Erfüllung normativer Erwartungen der Umwelt erreicht werden. Indem sie die ‚Erwartungen‘ der BiSt-Politik durch entsprechende Entwicklungsprozesse noch *vor* den Standardüberprüfungen erfüllen, versuchen Schulen zu verhindern, dass sie ‚normativ auffällig‘ werden. Auch die Nutzung der staatlich zur Verfügung gestellten Unterrichts- und Diagnosematerialien sowie Unterstützungs- und Fortbildungsmöglichkeiten – also Prozess (3) – signalisiert der Umwelt, dass eine Schule ernsthaft versucht, an einer Reformpolitik mitzuwirken.

Prozess (2) ‚Stimulierung durch Feedback‘ findet seine theoretische Begründung in Feedbacktheorien (Visscher und Coe 2002): Akteure passen ihre Handlungen entsprechend ihrer Interpretation der verfügbaren Feedback-Informationen an (Hattie und Timperley 2007). Für Prozesse (4) und (5) können Theorien der sozialen Koordination (Schimank 2002; de Boer et al. 2007) eine theoretische Grundlage bieten: Der Einbezug einer ‚dritten Partei‘ – der ‚Stakeholder‘ in Prozess (4) und der Peers in Prozess (5) – verstärkt die normativen Erwartungen des Reformvorschlags (der das ‚Gemeinwohl‘ gegenüber ‚Individual- und Partialinteressen‘ in Stellung bringt) und soll die Notwendigkeit der sozialen Koordination zwischen Akteuren und Ebenen bewusst halten.

3.3 Fragestellungen

Der vorliegende Artikel soll zum besseren Verständnis der Prozesse des Wirksamwerdens und der Wirkungen von ‚evidenzbasierten Reformen‘ beitragen. Er untersucht, ob die normativen Erwartungen der österreichischen BiSt-Reform in Hinblick auf Wirkprozesse sowie auf Wirkungen der Schul- und Unterrichtsentwicklung sich in den Wahrnehmungen und Erfahrungen von Schulleitungen und Lehrpersonen widerspiegeln bzw. welche Abweichungen von den postulierten Erwartungen auftreten. Dies geschieht entlang zweier Fragestellungen:

1. Werden durch die BiSt-Reform Prozesse der Schul- und Unterrichtsentwicklung an Primarschulen stimuliert und, wenn ja, welche?

Da wir in unserer qualitativen Studie keine Ergebnismaße für die Reformziele ‚verbesserte Schülerkompetenzen‘ und ‚Chancengleichheit‘ zur Verfügung haben, rekurren wir (wie auch die offiziellen Zielbeschreibungen) auf die Prozessziele ‚Schul- und Unterrichtsentwicklung‘, indem wir in den Interviewdaten Aussagen über diese Prozesse suchen. Diese werden in Hinblick auf ihre Art (oberflächliche Anpassung/Entwicklung zu kompetenzorientiertem Unterrichten; vgl. Groß Ophoff 2013) bzw. ihre Zielrichtung (Kompetenzverbesserung/Chancengleichheit; vgl. Altrichter und Kanape-Willingshofer 2012) analysiert.

2. Wie erklärungskräftig sind die fünf postulierten Wirkungsprozesse der BiSt-Reform auf Einzelschulebene für die in den Fallschulen vorgefundenen Entwicklungsprozesse?

Um die reale Wirkungsweise der BiSt-Reform zu erfassen, werden die in den Interviewdaten berichteten Prozesse mit den bildungspolitisch postulierten Wirkungsvorstellungen abgeglichen.

4 Methode

An drei Primarschulen in einem österreichischen Schulbezirk wurden qualitative Fallstudien (Yin 2013) in einem Längsschnittdesign durchgeführt. Zu drei Erhebungszeitpunkten wurden mit jeweils adaptierten Leitfäden Interviews mit Lehrpersonen, Schulleiterinnen, Eltern und Schüler/innen über Erfahrungen und Meinungen zur Reformpolitik und ihrer Implementation durchgeführt: vor der ersten Standardüberprüfung (t_1 = April 2013), nach den ersten Ergebnisrückmeldungen an die Schulen (t_2 = Mai 2014) und vor der zweiten Testung (t_3 = März 2015). Die Leitfäden⁴ enthielten direkte Fragen nach Aktivitäten der Unterrichts- und Schulentwicklung (Beispielfragen: Hat sich Ihr Unterricht seit Einführung der BiSt verändert? Wenn ja, inwiefern? [Lehrpersonenleitfaden, 1. Welle]; Sind in den Konferenzen Beschlüsse für derartige Weiterentwicklungen gefasst worden? [Lehrpersonenleitfaden, 2. Welle]). Für Forschungsfrage 2 wurden Aussagen über individuelle und kollektive Aktivitäten der Lehrpersonen und Schulleitungen angesichts der verschiedenen Elemente der BiSt-Reform (Beispielfrage: Wenn Sie die Ergebnisse der BiSt-Testung überlegen, ergeben sich daraus für Sie neue Ideen für den Unterricht? Welche ersten Ideen sind Ihnen dazu aufgefallen? [Lehrpersonenleitfaden, 2. Welle]) interpretativ den fünf Prozessen zugeordnet.

Die Auswahl der Schulen erfolgte auf Basis einer Befragung von Expert/innen mit Feldexpertise (Schulaufsicht) im Sinne einer ‚maximum variation‘ (Lodico et al. 2010, S. 167) zu den Merkmalen geographische Lage, Entwicklungsbereitschaft und Schülerpopulation (soziokulturelle und Leistungsmerkmale). Zu jedem Erhebungszeitpunkt wurden jeweils die Schulleitung und die Klassenlehrerinnen überprüfter Klassen (4. Schulstufe) interviewt sowie ergänzende Gespräche mit weiteren Lehrerinnen geführt (diese waren ausschließlich weiblich). Schüler/innen wurden zu Erhebungszeitpunkt t_1 und t_3 , Eltern zu t_1 interviewt. Die Auswertungen in Kap. 5 stützen sich nur auf die Schulleitungs- und Lehrpersonendaten ($n_{t_1-t_3} = 27$ Interviews). Die Fallschulen lassen sich auf Basis der Einstiegsfragen in den Interviews von t_1 (Wie charakterisieren Sie Ihre Schule? Was sind besondere Stärken Ihrer Schule?) kurz folgendermaßen beschreiben:

⁴ Verwendet wurden zum Zeitpunkt t_1 vier Leitfäden jeweils adaptiert für Schüler/innen, Lehrpersonen, Eltern und Schulleiter/innen, zu t_2 zwei Leitfäden für Lehrpersonen und Schulleiter/innen sowie zu t_3 drei Leitfäden für Schüler/innen, Lehrpersonen und Schulleiter/innen. Die Interviewleitfäden sind über die Erstautorin (manuela.gamsjaeger@ph-ooe.at) zugänglich.

1. Primarschule A ($n_{t1} = 4_{1SL/1E/1S/1L}$; $n_{t2} = 3_{1SL/2L}$; $n_{t3} = 4_{1SL/2L/1S}$ ⁵): Schule im ländlichen Umfeld einer Kleinstadt, 174 Schüler/innen, neun Klassen, einige Kinder aus sozial schwierigen Milieu. Selbstbeschreibung: Wohlfühlen ist das wichtigste Merkmal.
2. Primarschule B ($n_{t1} = 5_{1SL/1E/1S/2L}$; $n_{t2} = 3_{1SL/2L}$; $n_{t3} = 5_{1SL/2L/2S}$): große Schule in kleinstädtischem Milieu, 245 Schüler/innen, 13 Klassen, hoher Migrationsanteil und viele Schüler/innen mit sonderpädagogischem Förderbedarf, Musikschwerpunkt. Selbstbeschreibung: individuelle, liebevolle Betreuung, Heterogenität wird als Ressource betrachtet.
3. Primarschule C ($n_{t1} = 7_{1SL/1E/2S/2L/1Studentin}$; $n_{t2} = 3_{1SL/2L}$; $n_{t3} = 3_{1SL/1L/1S}$): zweiklassig geführte Kleinstschule in ländlicher Region, 38 Schüler/innen, die Schulleiterin ist zugleich die Klassenlehrerin der gemeinsam geführten 3. und 4. Schulstufe und wird von einer Lehrerin unterstützt, die Stunden in Deutsch und Mathematik übernimmt. Selbstbeschreibung: begehrte Schule, auch für Kinder aus anderen Gemeinden.

Die Interviews wurden vollständig, zeilenweise und in Standardorthographie transkribiert (nach den Regeln bei Altrichter et al. 2018, S. 129f.) und ergaben einen Datenkorpus von etwa 280 Seiten. Für die Zwecke dieses Artikels wurden die qualitativen Daten nicht, wie sonst häufig, explorierend, typen- oder theoriegenerierend ausgewertet. Vielmehr wurde das normative Modell der BiSt-Nutzung als Satz von ex-ante-Hypothesen aufgefasst, deren Erwartungen über Sinndeutungen und Handlungen der professionellen Akteure an dem Datenmaterial einer ersten Prüfung unterzogen werden sollten (Hopf 1996; Meinefeld 2012). Dazu wurden die Interviews mittels strukturierender Inhaltsanalyse (Kuckartz 2012) mit einem im Forscherteam erstellten deduktiven Kategorienschema, das induktiv ergänzt wurde, nach Sinneinheiten analysiert und für jeden Erhebungszeitpunkt in Fallberichte über die einzelnen Schulen überführt. Die Kodierungen und Interpretationen wurden in regelmäßigen Forscherworkshops mittels konsensueller Validierung zwischen den beteiligten Forscher/innen überprüft; für jede Schule wurde eine zusammenfassende Fallstudie erstellt. Diese drei Fallstudien wurden abschließend durch eine Cross Case Analyse in Hinblick auf überlappende und widersprechende Muster analysiert (vgl. Altrichter et al. 2018, S. 248f.). Für die in diesem Beitrag berichteten Ergebnisse waren bei der Analyse jene sieben Kategorien leitend, die sich aus den beiden Forschungsfragen ergeben (Unterrichts- bzw. Schulentwicklung; fünf Wirkungsprozesse).

Durch die methodische Anlage der Studie versuchen wir, Wirkungen und Wirkungswege der Reform über die Berichte von unterschiedlichen Betroffenengruppen über Rekontextualisierungs- und Entwicklungsprozesse abzubilden. Auch sind keine generalisierenden Aussagen angestrebt, doch bietet die Überprüfung, ob die durch die BiSt-Politik postulierten Entwicklungen und Prozesse überhaupt in Bei-

⁵ Schulleitungen werden mit SL abgekürzt, Eltern mit E, Lehrpersonen mit L und Schüler/innen mit S. Die Zitierung erfolgt in Kap. 5 durch folgende Kürzel: „C2/L1/125“ bedeutet: Die Interviewaussage stammt aus Schule C, Erhebungszeitpunkt t2, Lehrperson 1 und findet sich im Transkript des Interviews auf Zeile 125. SL steht für Schulleitung, L für Lehrperson.

spielsschulen aufzufinden sind, Ansatzpunkte für Hypothesen und weiterführende theoriegeleitete Forschung.

5 Rekontextualisierung der Bildungsstandard-Politik in Primarschulen

Im Folgenden werden die Ergebnisse zu den beiden forschungsleitenden Fragen dargestellt. Abschn. 5.1 erörtert im Sinne von Forschungsfrage 1, ob im Kontext der BiSt-Reform Aktivitäten der Schul- und Unterrichtsentwicklung angeregt wurden. In den Abschn. 5.2–5.6 wird – im Sinne von Forschungsfrage 2 – jeweils einer der durch das ‚normative Wirkungsmodell‘ der BiSt-Reform postulierten Prozesse anhand der aus den drei Fallstudien vorliegenden Daten diskutiert.

5.1 Schul- und Unterrichtsentwicklung durch Bildungsstandard-Reform

Die BiSt-Reform ist eine Intervention auf Systemebene, die an den Schulstandorten Entwicklungsprozesse anstoßen soll. Dabei kann man Aktivitäten der Unterrichtsentwicklung von solchen der Schulentwicklung unterscheiden. Für die Zwecke der hier vorgelegten Analyse arbeiten wir mit einer einfachen Definition: Unter dem ersten Konzept verstehen wir alle didaktischen und erzieherischen Aktivitäten, die sich auf einzelne Klassen beziehen, während mit Schulentwicklung alle klassenübergreifenden oder die Koordination der Lehrerverarbeit betreffende Entwicklungsmaßnahmen bezeichnet werden.

In den Fallschulen finden sich nur wenige Hinweise auf *Schulentwicklung*. Die Schulleiterin der Primarschule B nimmt die bevorstehende Standardüberprüfung zum Anlass, um von den Klassenlehrerinnen der 4. Schulstufe eine gemeinsame Planung des Unterrichtsjahrs zu fordern. Sie nutzt Ergebnisse aus Datenfeedback und diagnostischen Tests für Entwicklungsgespräche mit Lehrpersonen, bei denen folgende Informationen berücksichtigt werden: „... [der] Vergleich erstens IKM und dann die tatsächliche Standardtestung, die Konsequenzen, die die jeweiligen Teams daraus ziehen, wie sie arbeiten damit.“ (B2/SL/279–283) Die betroffenen Klassenlehrerinnen verstärken ihre Koordination und regen dadurch auch andere Lehrkräfte zur Kommunikation über den eigenen Unterricht an. In den Schulen A und C werden keine neuen klassenübergreifenden Aktivitäten gesetzt, aber bereits bestehende Koordinationsmechanismen genutzt.

Demgegenüber finden sich mehr Befunde zur *Unterrichtsentwicklung*. In geringem Ausmaß kommt es zu umfassenderen unterrichtsbezogenen Änderungen, wie Arbeiten in leistungsgemischten Kleingruppen (A2/SL/491–497; A3/SL/529–532). In allen Schulen (am intensivsten in Schule B) wird über punktuelle, individuell gesetzte Veränderungen berichtet, die (1) mit dem Ziel eines kompetenzorientierten Unterrichts verbunden werden können, wie bspw. Förderung von selbstständigem und logischem Denken, Vorgabe unterschiedlicher Aufgabenschemata, verstärkter Fokus auf sinnerfassendes Lesen etc. Diese Veränderungen werden auch von Lehrpersonen in den nicht getesteten Klassen bzw. den Schulleiterinnen bemerkt und scheinen auch nach den Standardüberprüfungen fortgesetzt zu werden. „[...] und was vielleicht schon auch dazukommt, die Förderung des logischen Denkens. In Ma-

thematik, so kombinatorische Sachen, auf die auch Gewicht gelegt wird, werden sicherlich etwas vermehrt in den Unterricht eingebunden.“ (A1/LT/5–8).

In Schule A und B werden aber auch adaptive und ‚oberflächlichere‘ (im Sinne von Groß Ophoff 2013) Veränderungen von Unterricht vorgenommen, die der (2) Vorbereitung auf die Standardüberprüfung in der 4. Schulstufe dienen, der Strategie ‚teaching to the test‘ (Au 2007) entsprechen und nach den Standardüberprüfungen eingestellt werden. Bspw. werden Arbeitsblätter zwecks besserer Prüfungsvorbereitung umgestaltet, Testformate geübt, Klassenarbeiten an die Inhalte der Standardüberprüfung angepasst, mehr Unterrichtszeit in den Testgegenstand Mathematik investiert, mehr Frontalunterricht für die schnellere Vermittlung von Inhalten eingesetzt oder Lehrstoff entsprechend der Prüfungsanforderungen zeitlich umgruppiert. In manchen Fällen führen diese adaptiven Veränderungsprozesse zu einer Auseinandersetzung mit kompetenzorientiertem Unterricht.

[Wir] haben dann schon auch immer wieder diese Beispiele [aus dem Internet oder aus Begleitmaterialien] eingebaut, die das logische Denken oder den Hausverstand eigentlich fördern. (B2/L1/225–229)

Die Bildungsstandard-Reform hat in den drei untersuchten Schulen offenbar zu unterschiedlichen Impulsen für Unterrichtsentwicklung geführt, die von häufigeren, relativ ‚oberflächlichen‘ Anpassungsprozessen an neue Anforderungen bis hin zu einzelnen Ansätzen einer koordinierten Umsetzung kompetenzorientierten Unterrichts reichen. Allerdings werden Entwicklungsbemühungen ausschließlich in Hinblick auf das Reformziel ‚Kompetenzverbesserung‘ thematisiert. Ein etwaiges positives Potential der Reform für das Ziel ‚Chancengleichheit‘ kommt in den Beschreibungen der Entwicklungsmaßnahmen nicht vor (explizit wurde danach nicht gefragt); dagegen gibt es an einigen wenigen Stellen skeptische Kommentare, wenn Lehrpersonen (v. a. in Schule A) meinen, das zur Verfügung gestellte kompetenzorientierte Material wäre überfordernd für leistungsschwächere Schüler/innen und erschwere diesen Erfolgserlebnisse (vgl. Abschn. 5.4).

5.2 Aufbau von Erwartungen durch Bildungsstandards

In der BiSt-Reform wird die Absicht formuliert, durch individualisierten Unterricht kompetenzorientiertes Lernen zu ermöglichen. Diese Zielsetzung wird für das Primarschulwesen von fast allen befragten Lehrerinnen und Schulleiterinnen als längst notwendige Veränderung befürwortet und in Übereinstimmung mit dem eigenen Professionsverständnis gesehen. Viele Lehrerinnen verbinden mit dem Begriff Bildungsstandards aber beinahe ausschließlich die Standardüberprüfung. Damit überblenden testbezogene Assoziationen vor allem in der ersten Erhebungswelle der Studie die mit der BiSt-Politik (auch) verbundene Intention, durch Formulierung eines transparenteren Referenzsystems individualisierten Unterricht bzw. kompetenzorientiertes Lernen zu fördern. Standards werden dann mit ‚standardisierter, gleichmachender‘ Behandlung von Lernenden assoziiert und damit als unvereinbar mit der Forderung nach individualisierendem Unterricht angesehen.

Weil ja wieder da jeder das Gleiche abgefragt wird, und wenn ich da sage, ich habe ein Kind, das hat in die Richtung totale [spezielle] Fähigkeiten, ich fördere sehr in diese Richtung. [...] und plötzlich wird das Kind aber genau da [gemeint: das Gleiche wie alle anderen] abgefragt. Dann passt das für mich nicht zusammen. (C3/L1/60)

In allen Schulen wurde das ungewohnte Testformat mit den Schüler/innen geübt. Ein Großteil der Lehrpersonen der Schulen A und B verteilte Unterrichtszeit zugunsten des überprüften Gegenstands um; in Schule A wird verstärkt Frontalunterricht eingesetzt, um Inhalte ‚schneller vermitteln‘ zu können. Während in Schule A kompetenzorientierter Unterricht häufig als zu schwierig und überfordernd für leistungsschwächere Schüler/innen angesehen wird (A2/L2/120–123), befürworten die Lehrerinnen und die Schulleiterinnen der Schulen B und C die Kompetenzorientierung der Bildungsstandards als Unterstützung des eigenen offenen und individualisierten Unterrichtskonzepts, auch wenn die Schulleiterin der Schule C dem Testaspekt der BiSt-Politik sehr kritisch gegenübersteht. In Schule C wurde die positive Teilnahme an einem Mathematikwettbewerb als Erfolgsbestätigung für den eigenen kompetenzorientierten Unterricht und das hohe Leistungsvermögen der Lernenden betrachtet und kein Anlass zu einer besonderen Vorbereitung auf die Standardtests gesehen. Vor der zweiten Standardüberprüfung in Deutsch wurden an den Fallschulen im Vergleich zur ersten Testung geringere Veränderungen berichtet. Ob dies einer beginnenden ‚Routiniertheit‘ oder dem Fach zuzusprechen ist, kann in unserer Studie nicht beantwortet werden.

Wenn Unterrichtsveränderungen in Richtung eines stärker kompetenzorientierten Unterrichts stattfinden, werden sie meist nicht nur auf die BiSt-Reform, sondern auch auf eigene Weiterbildung, veränderte Rahmenbedingungen, andere Reformmaßnahmen (z. B. Qualitätsmanagementstrategie SQA) und Unterrichtsmaterial zurück geführt. *„Durch die Lehrbücher, auch schon mal, da ist es [gemeint: kompetenzorientierte Lernangebote] jetzt intensiver drinnen, durch die Erfahrung, dass das eigentlich wirklich viel bringt, durch meine eigene Erfahrung und eben auch durch die Motopädagogik.“ (B2/L2/430–432).*

5.3 Stimulation durch Datenfeedback

Zu Beginn der BiSt-Implementierung nimmt die Standardüberprüfung in allen Schulen am meisten Aufmerksamkeit in Anspruch. Die Testkonzeption (v. a. deren Merkmale: am Ende der 4. Schulstufe/kurz vor dem Schulwechsel, lange Testdauer, Testformate) wird – zu Beginn der Studie teilweise heftig – als für die Primarstufe ungeeignet kritisiert und als Einschränkung des eigenen Gestaltungsspielraums wahrgenommen. Gegen Ende der Studie nimmt diese Kritik ab. Nur die Schulleiterin der Schule C steht der „Testkultur“ in der Primarstufe auch gegen Ende der Studie skeptisch gegenüber und sieht diese im Widerspruch zu dem Anspruch, Kinder in der Primarschule nicht „unter Druck“ (C2/SL/114) zu setzen. *„Wobei ich sagen muss, wir haben bei den Mathematik-Bildungsstandards haben wir wirklich sehr gut abgeschnitten. Das ist auch nicht das Problem. Oder dass man sich scheut. Aber diese Testungsart, das ist der Volksschule so weit entfernt und fremd und weit*

weg noch ... diese lange Zeit [gemeint: lange Testdauer], da [...] sind die Kinder doch sehr überfordert [...].“ (C3/SL/16–21).

Auch wird die Standardüberprüfung zu Beginn der Studie als externe Evaluierung interpretiert, die der *Überprüfung der Lehrarbeit* diene. Ängste vor einem möglichen Schulranking oder der Zurechnung von Ergebnissen auf einzelne Lehrpersonen durch hierarchisch höhere Ebenen werden geäußert. Von allen Interviewpartnerinnen bewerten nur die Schulleiterinnen der Schulen A und B die Standardüberprüfungen bereits zu Beginn der Studie explizit positiv, weil sie durch die Rückmeldung positive Auswirkungen auf die Unterrichtspraxis der Lehrpersonen erwarten. „Wir wollen ja nicht, dass alle die gleichen Kompetenzen erreichen, sondern wir schauen ja nur, wer wie weit kommt, eigentlich“ (B3/SL/589–590).

Da die Ergebnisse in allen Schulen innerhalb oder über den Erwartungswerten der Standardtests liegen und keine negativen Konsequenzen für Schulen beobachtet werden, nimmt die ablehnende Haltung gegen Ende der Studie deutlich ab, und manchen Lehrerinnen ist es möglich, Datenfeedback als Rückmeldung und Reflexionsmöglichkeit der eigenen Arbeit anzuerkennen.

Ich fühle mich jetzt auch nicht persönlich angegriffen oder ich sehe es jetzt auch nicht als ‚Lehrer-Austestung‘ und ich kriege eine Note oder so – sondern man muss selber einmal reflektieren. (B3/L1/249–252)

Etwa drei bis vier Monate nach der Datenrückmeldung ist nur einer Lehrerin dieser Studie ein konkretes Detailergebnis in Erinnerung, das sie im darauffolgenden Schuljahr stärker berücksichtigen will (C2/L1/198–203). Die meisten befragten Lehrerinnen erinnern gegen Ende der Studie nur, dass „gute Ergebnisse“ rückgemeldet wurden; sie scheinen das Feedback nur in Hinblick auf die Gesamtaussage („innerhalb oder oberhalb der Erwartungswerte“), nicht aber in den Detailhinweisen auszuwerten. Gute Gesamtergebnisse werden als Erfolg der bisherigen Arbeitsweise und als Bestätigung der eigenen Einschätzung des Leistungsvermögens der Klasse interpretiert und machen Veränderungen im Unterrichtshandeln unnötig.

Die Schulleiterinnen sehen im Datenfeedback insgesamt mehr Vorteile als die befragten Lehrerinnen: Innerhalb der Schule könne von Klassen mit guten Ergebnissen gelernt werden, allerdings aus Perspektive der Leiterinnen der Schulen A und C nur freiwillig, da sie als Führungspersonen „keine Konsequenzen“ (C3/SL/15) ziehen und Lehrpersonen nicht zu Veränderungen bzw. Fortbildungen zwingen könnten.

5.4 Alignment durch Unterstützung

Die durch die BiSt-Implementation bereit gestellten Begleitmaßnahmen werden von den Befragten durchwegs positiv gesehen und setzen wahrnehmbare Impulse für die Umsetzung der BiSt-Politik. Die Leiterinnen der Schulen A und B sehen v. a. in der *Informellen Kompetenzmessung (IKM)* ein brauchbares Entwicklungsinstrument, dessen Ergebnisse sie auch für Einblicke in die Unterrichtspraxis und für Entwicklungsgespräche mit Lehrpersonen nutzen. Die IKM wird auch von den Lehrerinnen dieser Schulen geschätzt, da sie eine zeitnahe Rückmeldung über eigenen Unterricht und Schülerleistungen erlaubt, die Lehrpersonen Weiterentwicklung ermöglicht: „Als Lehrer [muss ich] soweit sein, dass ich das ausmerze, was bei den

IKM-Testungen aufgefallen ist. Da muss auch jeder Lehrer an sich selbst arbeiten, das finde ich auch gut.“ (A3/L1/470–473).

Auch die zur BiSt-Implementation bereitgestellten *Unterrichtsmaterialien* werden oft als Konkretisierung der Bildungsstandards und als Unterstützung für den eigenen kompetenzorientierten Unterricht empfunden (ähnlich Frühwacht 2012, S. 168 ff., für bayerische Lehrkräfte). Je nach Interpretation der Bildungsstandards und dem eigenen professionellen Bezugssystem (Asbrand et al. 2012) werden diese aber offenbar unterschiedlich wahrgenommen und eingesetzt: Die Schulleiterin der Schule C, für die die Kompetenzorientierung der Bildungsstandards im Vordergrund steht, sieht in den Materialien großes Potential für leistungsschwächere Schüler/innen und differenzierenden Unterricht. Andere Lehrerinnen, welche mit Bildungsstandards eher die Standardüberprüfung assoziieren, verwenden sie nur als Zusatzaufgaben für leistungsstärkere Schüler/innen und kritisieren, dass schwächere Schüler/innen durch die Aufgabenstellungen überfordert seien. Diese bekämen dadurch ihre Schwächen „vorgeführt“ (A2/L2/238–243). Während einige Lehrkräfte die Übungsmaterialien auch weiterhin einsetzen möchten, war für andere nur die bevorstehende Standardüberprüfung der Grund für ihre Verwendung.

... so jetzt wirklich diese Arbeitsblätter jetzt noch zusätzlich durcharbeiten und zusätzliche Gruppenarbeiten machen, würde ich wahrscheinlich jetzt nicht so intensiv machen, als wie wenn eine Testung ansteht. (B2/L2/94–100)

Dennoch, die Unterstützungsmaßnahmen, insbesondere Unterrichtsmaterialien und IKM, drücken als „gebrauchsfertige Rekontextualisierungen“ (Altrichter und Gamsjäger 2017) die operativen Vorstellungen der BiSt-Politik am deutlichsten aus und scheinen aufgrund der Akzeptanz durch einen Großteil der Interviewten die Erreichung der Reformziele nachhaltiger zu unterstützen als die anderen vier Wirkprozesse.

5.5 Einbeziehen von Stakeholdern

Gespräche mit der Schulaufsicht zur Umsetzung der Bildungsstandards und den Ergebnissen des Datenfeedbacks werden von keiner interviewten Person thematisiert. Ihre Wahrnehmung wird möglicherweise in den ersten Jahren der Umsetzung durch Gespräche im Zuge des neuen Qualitätsmanagementsystems SQA überdeckt. Der Informationspflicht im *Schulpartnerschaftsgremium* kommen alle Schulleiterinnen nach: Dabei wird der Ergebnisbericht durchgehend sehr allgemein gehalten und Informationen über etwaige Leistungsunterschiede zwischen Klassen oder Schüler/innen vermieden.

Ich habe [gemeint: im Schulforum unter Anwesenheit von Eltern] aber auch gleich gesagt, es sind beide Klassen ... sind so gut wie gleich. Die hat da mal drei Punkte mehr ... wirklich. Und gar nichts. Und es ist da eigentlich gar keine Diskussion aufgekommen, wer da schlechter oder besser war. Gott sei Dank. (A2/SL/280–283)

Die Ergebnisse würden von den anwesenden Eltern zur Kenntnis genommen, ohne dass es jedoch zu weiteren Reaktionen oder Rückmeldungen an die Schulen

kam. Dies wird von den Interviewten mit der späten Datenrückmeldung im nächsten Schuljahr begründet, in dem die getesteten Schüler/innen nicht mehr an der Schule sind.

5.6 Alignment durch innerschulische Koordination

In den Schulen A und C werden die bestehenden engen Kooperationsbeziehungen zwischen den Lehrerinnen für einen Austausch über Unterricht genutzt, nicht jedoch explizit für weitere Entwicklungsaktivitäten in Hinblick auf Standards oder deren Überprüfung. In Schule B verstärken die bevorstehenden Standardtests die Koordination und Abstimmung zwischen den Klassenlehrerinnen, die von der Schulleiterin gefordert und als Entwicklungsziel formuliert, aber von manchen Lehrpersonen als ungewünschter „Gleichschritt“ (B3/L1/89–92) kritisiert wird.

Wir [Klassenlehrer/innen der 4. Schulstufe] helfen zusammen, ja. Und wir haben schon gesagt, ja, wie machen wir das jetzt und tja, dann natürlich auch in anderen Schulen erkundigen, wie macht ihr das? (B1/LT/5–6)

Hier scheinen Unterschiede in der Führungskonzeption der Schulleiterinnen deutlich zu werden: Nur die Schulleiterin der Schule B überträgt die Verantwortung für die Elterninformation den Klassenlehrerinnen, stellt konkrete Koordinationsanforderungen an die Lehrkräfte der 4. Schulstufe und will die Ergebnisse der diagnostischen Tests für Entwicklungsgespräche nutzen. Zusätzliche Strukturen innerschulischer Koordination für die Umsetzung der BiSt-Politik werden aber kaum aufgebaut.

6 Zusammenfassung und Diskussion

Der Beitrag ging der Frage nach, ob die österreichische BiSt-Reform zu Schul- und Unterrichtsentwicklung anregt und ob etwaige Veränderungen durch fünf erwartete Wirkungsprozesse erklärt werden können. Die Ergebnisse lassen sich für die drei hier untersuchten Schulen folgendermaßen zusammenfassen:

Während kaum *Schulentwicklung* im Sinne klassenübergreifender Aktivitäten oder Veränderung der Arbeitsstruktur von Lehrpersonen beobachtbar ist, kommt es im Kontext der BiSt-Reform zu *Unterrichtsentwicklung* – sowohl im Sinne von längerfristiger Entwicklung kompetenzorientierten Unterrichts (z.B. logisches Denken in Mathematik über die Schulstufen hinweg) als auch von kurzfristiger Anpassung an die Inhalte, Testformate und Aufgabentypen der Standardtests. Die Art der Umsetzung scheint maßgeblich vom Professionsverständnis der Lehrpersonen (v. a. Lernverständnis, Bedeutung von Kompetenzorientierung und Bildungsstandards; Asbrand et al. 2012) und von Kontextfaktoren (z.B. Führungsstil der Schulleitung, Verständnis der Reform als Druck oder Unterstützung) beeinflusst.

Die Idee, dass Bildungsstandards nicht nur zur Kompetenzentwicklung, sondern auch zur Chancengleichheit im Bildungswesen beitragen könnten, taucht in den Daten nicht auf. Wenn differentielle Wirkungen der Reform überhaupt angesprochen werden, dann wird im Gegenteil v. a. in Schule A die Befürchtung geäußert, dass die anspruchsvolleren Aufgaben schwache Schüler/innen überfordern könnten.

Von fünf erwartbaren Prozessen des Wirksamwerdens der BiSt-Politik zeigen sich zwei erklärungskräftig: Durch die curricularen Kompetenzformulierungen (Bildungsstandards) und den Mechanismus der Standardüberprüfung werden *Erwartungen* über Ziele und Formen des Unterrichtens *aufgebaut*, die – vor allem bei jenen Lehrpersonen, deren Klassen eine Standardtestung bevorsteht – zu einer Überprüfung und Veränderung ihrer Unterrichtspraktiken führen. Besonders wirksam scheinen dafür die *bereitgestellten Unterrichtsmaterialien und diagnostischen Tests* zu sein (ähnlich: Zuber 2019). Sie werden überwiegend positiv konnotiert und machen Ziele und Erwartungen an kompetenzorientiertes Unterrichten für die beteiligten Lehrerinnen und Schulleiterinnen konkret und operabel (vgl. Wacker 2008). Dabei sind sowohl Entwicklungen in Richtung eines kompetenzorientierten Unterrichts zu beobachten, die wahrscheinlich die Erfahrungsmöglichkeiten für Schüler/innen erweitern, als auch adaptive Anwendungen von Materialien, die als *teaching to the test* (Au 2007) zu interpretieren sind.

Von *Standardüberprüfung und Datenfeedback*, eigentlich Kernstücke ‚evidenzbasierter Steuerung‘, scheint in den drei untersuchten Schulen wenig Stimulation für Unterrichtsentwicklung auszugehen. Auch wenn die skeptische Einschätzung durch Lehrpersonen im Verlaufe der Implementierung abnimmt, ist auffallend, dass die Botschaften des Datenfeedbacks nur in einer sehr allgemeinen Weise zur Kenntnis genommen, nicht jedoch im Detail analysiert werden. Von einer ‚Rationalisierung‘ der Unterrichtsentwicklung durch eine bessere Informationslage kann in den untersuchten Fällen nicht die Rede sein. Die Datenrückmeldung wird offenbar nach der Frage „Sind wir innerhalb oder über dem Erwartungswert?“ prozessiert; ist die Antwort positiv, erscheint eine weitere Analyse von Detailinformationen und auch Entwicklungsmaßnahmen unnötig.

Ebenfalls wenige Entwicklungsimpulse ergeben sich durch *Innerschulische Koordination* und *Einbeziehen von Stakeholdern*. Die Anforderungen der BiSt-Politik können offenbar mit bestehenden Koordinationsinstrumenten bewältigt werden; allerdings nimmt in einem Fall (Schule B) die Schulleiterin die Reform zum Anlass, mehr Koordination zwischen den Lehrkräften der 4. Schulstufe einzufordern. Dass die *Einbeziehung von Stakeholdern* für die Umsetzung der BiSt-Reform unerheblich ist, mag mit der mangelnden Praxis der Elternmitwirkung in den Schulen, aber auch mit der Tatsache, dass die Eltern der getesteten Schüler/innen meist nicht mehr im Schulpartnerschaftsgremium mitwirken, zusammenhängen.

Was sagen die Ergebnisse nun über die bildungspolitischen Vorstellungen zur Implementation von ‚neuer Steuerung‘ durch Bildungsstandards aus? Für Befürworter/innen ‚neuer Steuerung‘ wird es zunächst enttäuschend sein, dass die Umsetzung der Reformideen in unseren Fällen eher durch den (traditionellen) Mechanismus der Erwartungsbildung als durch den auf der Idee der Evidenzbasierung fußenden Feedback-Mechanismus getrieben wird. Die Standardüberprüfung scheint damit weniger über das Datenfeedback als vielmehr über die ‚drohende‘ Testung und die Vorbereitung darauf zu wirken. Bedeuten diese Ergebnisse – sollten sie sich in weiteren Studien replizieren lassen –, dass die BiSt-Überprüfungen sinnlos sind? Auch in dem von uns beobachteten Szenario bauen die Kompetenztestungen erst jenen ‚normativen Druck‘ auf, der allein durch die Verlautbarung curricularer Ziele nicht gegeben wäre (Vollstädt et al. 1999) und der dazu führt, dass den Standards und den Unter-

stützungsmaterialien jene Aufmerksamkeit zukommt, die zu den Veränderungen des Unterrichts vor den Testungen führt.

Auch innovationsstrategisch kann man zwei unterschiedliche Schlüsse ziehen: In den Aufbau von Erwartungen und in Unterstützungsprozesse und -materialien (weiter) investieren, weil diese den aktuell sichersten Weg zur Umsetzung der Reformintentionen (allerdings in unterschiedlicher Qualität) darstellen. Wenn man aber die Analyse von Datenrückmeldungen (wie bildungspolitisch postuliert) zu einem regelmäßigen Element rationaler Schulentwicklung machen will, braucht es dafür – auf den verschiedenen Ebenen, auf denen sich die Wirkungsweise von Schule konstituiert – sicherlich mehr als das bloße Verfügbarmachen von Leistungsergebnissen. Hier brauchen wohl viele Lehrpersonen und Schulen entsprechende zeitliche Rahmenbedingungen und inhaltliche Unterstützung, um die differenzierte Interpretation von BiSt-Ergebnissen und die Konstruktion von didaktischen Konsequenzen zu einem regelmäßigen Merkmal der schulischen Arbeit zu machen.

Aus der *methodischen Anlage* unserer Studie deuten unseres Erachtens die partiell unterschiedlichen Ergebnisse in den drei Wellen sowie die unterschiedlichen Aneignungs- und Umsetzungsformen von Reformelementen darauf hin, dass eine längsschnittliche Betrachtung von Reformen und ein genauer qualitativer Blick auf die Prozesse der Aneignung und handlungspraktischen Übersetzung von Reformelementen Sinn machen. Rekontextualisierungen bilden sich vor dem Hintergrund anderer Bezugssysteme (Professionsverständnis, Definition von Begriffen wie Kompetenz oder Bildungsstandards, Verständnis der Schulkultur), die in unterschiedlichen Interpretationen ähnlicher Handlungen sichtbar werden. Auch meinen wir, dass der Bezug der Studie auf die Rekonstruktion eines Wirkungsmodells der Reform sinnvoll ist. Durch ein solches konzeptuelles Modell können verschiedene Elemente einer Reform in ihrem Zusammenwirken betrachtet und mit allgemeineren Theorien in Beziehung gesetzt werden.

Als *Limitationen* der Studie und *weitere Forschungsaufgaben* sind zu nennen: Insbesondere wäre es wünschenswert, mehr Fälle in die Analyse miteinzubeziehen. Dies gilt besonders für Schulen mit *unter* den Erwartungen liegenden Ergebnissen. Da alle drei Schulen Ergebnisse im oder über dem Erwartungsbereich aufwiesen, können keine Aussagen über Rekontextualisierung in Schulen unter dem Erwartungsbereich getroffen werden. Dies limitiert die Analyse sowohl in ihren innovationsstrategischen als auch theoretischen Konsequenzen, die in künftiger Forschung geklärt werden könnten: Analog zur Interpretation von Ehren et al. (2015) in Hinblick auf Schulinspektionen wird man weder bei ‚guten‘ noch bei ‚schlechten‘ Ergebnissen Entwicklungsaktivität *nach* dem Datenfeedback erwarten, weil die ersteren sich schon vorsorglich und rechtzeitig den Erwartungen angepasst haben, während den letzteren unter dem durch ungünstige Ergebnisse gestiegenen Handlungsdruck produktive Entwicklung schwerfällt. Die vorherrschende Sichtweise evidenzbasierter Steuerung, die wir zu Beginn skizziert haben, geht dagegen davon aus, dass gerade Diskrepanzerlebnisse zwischen Zielen und letztlich erzielten Ergebnissen zu Entwicklungsaktivität motivieren.

Funding Open access funding provided by Johannes Kepler University Linz.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Literatur

- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Engelwood Cliffs: Prentice-Hall.
- Altrichter, H. (2017). Von der Schulinspektion zum systemischen Qualitätsmanagement: Eine veränderte Strategie der Qualitätssicherung im österreichischen Schulsystem. *Zeitschrift für Bildungsverwaltung*, 33(2), 27–41.
- Altrichter, H., & Gamsjäger, M. (2017). A conceptual model for research in performance standard policies. *Nordic Journal of Studies in Education Policy*, 3(1), 6–20.
- Altrichter, H., & Kanape-Willingshofer, A. (2012). Bildungsstandards und externe Überprüfung von Schülerkompetenzen. In B. Herzog-Punzenberger (Hrsg.), *Nationaler Bildungsbericht Österreich 2012* (Bd. 2, S. 355–394). Graz: Leykam.
- Altrichter, H., Brüsemeister, T., & Heinrich, M. (2005). Merkmale und Fragen einer Governance-Reform am Beispiel des österreichischen Schulwesens. *Österreichische Zeitschrift für Soziologie*, 30(4), 6–28.
- Altrichter, H., Posch, P., & Spann, H. (2018). *Lehrerinnen und Lehrer erforschen ihren Unterricht*. Bad Heilbrunn: UTB/Klinkhardt.
- Asbrand, B., Heller, N., & Zeitler, S. (2012). Die Arbeit mit Bildungsstandards in Fachkonferenzen. *Die Deutsche Schule*, 104(1), 31–43.
- Au, W. (2007). High stakes testing and curricula control. *Educational Researcher*, 36(5), 258–267.
- BIFIE (2012). Bildungsstandards. <https://www.bifie.at/bildungsstandards>. Zugegriffen: 10. Mai 2012.
- BIFIE (2016). Implementierung. <https://www.bifie.at/node/50>. Zugegriffen: 6. Mai 2016.
- BMUKK (2012). *Brief der Bundesministerin vom September 2012 an österreichische Sekundarschulen*. Wien: Bundesministerium für Unterricht, Kunst und Kultur.
- de Boer, H., Enders, J., & Schimank, U. (2007). On the way towards new public management? The governance of university systems in England, The Netherlands, Austria, and Germany. In D. Jansen (Hrsg.), *New forms of governance in research organisations* (S. 137–152). Dordrecht: Springer.
- Ehren, M. C. M., Altrichter, H., McNamara, G., & O'Hara, J. (2013). Impact of school inspections on teaching and learning – describing assumptions on causal mechanisms in six European countries. *Educational Assessment, Evaluation and Accountability*, 25(1), 3–43.
- Ehren, M. C. M., Gustafsson, J.-E., Altrichter, H., Skedsmo, G., Kemethofer, D., & Huber, S. G. (2015). Comparing effects and side effects of different school inspection systems across Europe. *Comparative Education*, 51, 375–400.
- Feldhoff, T., Durrer, L., & Huber, G. (2012). Steuerung eines Schulsystems. *Die Deutsche Schule*, 104, 71–87.
- Fend, H. (2006). *Neue Theorie der Schule*. Wiesbaden: Springer VS.
- Freudenthaler, H. H., & Specht, W. (2005). *Bildungsstandards aus Sicht der Anwender*. Graz: ZSE.
- Freudenthaler, H. H., & Specht, W. (2006). *Bildungsstandards: Der Implementationsprozess aus der Sicht der Praxis*. Graz: ZSE.
- Frühwacht, A. (2012). *Bildungsstandards in der Grundschule*. Bad Heilbrunn: Klinkhardt.
- Frühwacht, A., & Maier, U. (2012). Konzeptionelle Vorstellungen über Bildungsstandards bei deutschen und finnischen Grundschullehrkräften. In F. Hellmich, S. Förster & F. Hoya (Hrsg.), *Bedingungen des Lehrens und Lernens in der Grundschule* (S. 185–188). Wiesbaden: VS.
- Grabensberger, E., Freudenthaler, H. H., & Specht, W. (2008). *Bildungsstandards: Testungen und Ergebnisrückmeldungen auf der achten Schulstufe aus der Sicht der Praxis*. Graz: BIFIE.
- Grillitsch, M. (2010). *Bildungsstandards auf dem Weg in die Praxis*. Graz: Leykam.
- Groß Ophoff, J. (2013). *Lernstandserhebungen: Reflexion und Nutzung*. Münster: Waxmann.
- Groß Ophoff, J., Koch, U., Helmke, A., & Hosenfeld, I. (2006). Vergleichsarbeiten für die Grundschule – und was diese daraus machen (können). *Journal für Schulentwicklung*, 10, 7–12.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112.

- Hopf, C. (1996). Hypothesenprüfung und qualitative Sozialforschung. In R. Strobel & A. Böttger (Hrsg.), *Wahre Geschichten? Zur Theorie und Praxis qualitativer Interviews* (S. 11–21). Baden-Baden: Nomos.
- Husfeldt, V. (2011). Wirkungen und Wirksamkeit der externen Schulevaluation. *Zeitschrift für Erziehungswissenschaft*, *14*, 259–282.
- Kuckartz, U. (2012). *Qualitative Inhaltsanalyse*. Weinheim: Beltz.
- Kühle, B., & Peek, R. (2007). Lernstandserhebungen in Nordrhein-Westfalen. *Empirische Pädagogik*, *21*, 428–447.
- Kuper, H., & Hartung, V. (2007). Überzeugungen zur Verwendung des Wissens aus Lernstandserhebungen. *Zeitschrift für Erziehungswissenschaft*, *10*, 214–229.
- Leeuw, F. L. (2003). Reconstructing program theories. *American Journal of Evaluation*, *24*(1), 5–20.
- Lenski, A., Richter, D., & Pant, H. A. (2015). Kompetenzorientierung im Unterricht aus der Perspektive von Lehrkräften und Schülerinnen und Schülern. *Zeitschrift für Pädagogik*, *61*(5), 712–737.
- Lodico, M. G., Spaulding, D. T., & Voegtler, K. H. (2010). *Methods in educational research* (2. Aufl.). San Francisco: Jossey-Bass.
- Meinefeld, W. (2012). Hypothesen und Vorwissen in der qualitativen Sozialforschung. In U. Flick, E. v. Kardorff & I. Steinke (Hrsg.), *Qualitative Forschung* (9. Aufl., S. 265–275). Reinbek: Rowohlt.
- Nachtigall, C., & Jantowski, A. (2007). Die Thüringer Kompetenztests unter besonderer Berücksichtigung der Evaluationsergebnisse zum Rezeptionsverhalten. In I. Hosenfeld & J. Groß Ophoff (Hrsg.), *Nutzung und Nutzen von Evaluationsstudien in Schule und Unterricht* (S. 401–410). Landau: Empirische Pädagogik.
- Oelker, J., & Reusser, K. (2008). *Qualität entwickeln – Standards sichern – mit Differenzen umgehen*. Reihe Bildungsforschung, Bd. 27. Bonn, Berlin: BMBF.
- Ozga, J., & Jones, R. (2006). Travelling and embedded policy. *Journal of Education Policy*, *21*, 1–17.
- Pöhlmann, C., Pant, H., Frenzel, J., Roppelt, A., & Köller, O. (2014). Auswirkungen einer Intervention auf die Auseinandersetzung und Arbeit mit Bildungsstandards. *Zeitschrift für Erziehungswissenschaft*, *17*, 113–133.
- Richter, D., Böhme, K., Becker, M., Pant, H. A., & Stanat, P. (2014). Überzeugungen von Lehrkräften zu den Funktionen von Vergleichsarbeiten. *Zeitschrift für Pädagogik*, *60*, 225–244.
- Schimank, U. (2002). *Handeln und Strukturen. Einführung in die akteurtheoretische Soziologie*. Weinheim: Juventa.
- Schimank, U. (2007). Die Governance-Perspektive: Analytisches Potenzial und anstehende konzeptionelle Fragen. In H. Altrichter, T. Brüsemeyer & J. Wissinger (Hrsg.), *Educational Governance* (S. 231–260). Wiesbaden: VS.
- Scott, W. R. (2001). *Institutions and organizations*. London: SAGE.
- Visscher, A. J., & Coe, R. (2002). *School improvement through performance feedback*. London: Routledge.
- Vollstädt, W., Tillmann, K.-J., & Rauin, U. (1999). *Lehrpläne im Schulalltag*. Opladen: Leske + Budrich.
- Wacker, A. (2008). *Bildungsstandards als Steuerungselemente der Bildungsplanung*. Bad Heilbrunn: Klinkhardt.
- Yin, R. K. (2013). *Case study research: design and method*. Los Angeles: SAGE.
- Zuber, J. (2019). Einstellungsbildung als Gelingensbedingung für die Umsetzung einer Bildungsstandardpolitik? In J. Zuber, H. Altrichter & M. Heinrich (Hrsg.), *Bildungsstandards zwischen Politik und schulischem Alltag*. (S. 105–127). Wiesbaden: Springer VS.