

Diemer, Tobias; Kuper, Harm

Formen innerschulischer Steuerung mittels zentraler Lernstandserhebungen

Zeitschrift für Pädagogik 57 (2011) 4, S. 554-571



Quellenangabe/ Reference:

Diemer, Tobias; Kuper, Harm: Formen innerschulischer Steuerung mittels zentraler Lernstandserhebungen - In: Zeitschrift für Pädagogik 57 (2011) 4, S. 554-571 - URN: urn:nbn:de:0111-opus-87466 - DOI: 10.25656/01:8746

<https://nbn-resolving.org/urn:nbn:de:0111-opus-87466>

<https://doi.org/10.25656/01:8746>

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

ZEITSCHRIFT FÜR PÄDAGOGIK

Heft 4

Juli/August 2011

■ *Thementeil*

Bildungspolitik zwischen Expertenwissen und Öffentlichkeit

■ *Allgemeiner Teil*

Die Einrichtung von Gemeinschaftsschulen in Schleswig-Holstein – der Kontext des Systemwandels in lokaler Entscheidung

Formen innerschulischer Steuerung mittels zentraler Lernstandserhebung

Unbekannt, unverstanden, unterschätzt – Gregor Girards facettenreiche Schulpädagogik

Inhaltsverzeichnis

Thementeil: Bildungspolitik zwischen Expertenwissen und Öffentlichkeit

Marcelo Caruso

Bildungspolitik zwischen Expertenwissen und Öffentlichkeit.

Einführung in den Thementeil 453

Michele S. Moses

Education-Policy-by-Ballot-Initiative: Considerations

Related to Democracy and Justice 458

Lucien Criblez

Vox populi – Zur Herausforderung der Bildungspolitik

durch die halbdirekte Demokratie 471

Florian Waldow

Juristen oder Testspezialisten? Zur Rolle von Experten bei der Herstellung

von Notengerechtigkeit in Deutschland und Schweden 487

Bernd Zymek/Sabine Wendt/Moritz Hegemann/Frank Ragutt

Regional Governance und kommunale Schulentwicklungspolitik im Prozess

des Rück- und Umbaus regionaler Schulangebotsstrukturen 497

Markus Maurer

Der *Free Education* Mythos und die Legitimationsprobleme privater Akteure

im Hochschulsektor Sri Lankas 513

Deutscher Bildungsserver

Linktipps zum Thema „Bildungspolitik zwischen Expertenwissen

und Öffentlichkeit“ 526

Allgemeiner Teil

Jürgen Wiechmann

Die Einrichtung von Gemeinschaftsschulen in Schleswig-Holstein –
der Kontext des Systemwandels in lokaler Entscheidung 534

Tobias Diemer/Harm Kuper

Formen innerschulischer Steuerung mittels zentraler Lernstandserhebungen 559

Beat Bertschy

Unbekannt, unverstanden, unterschätzt –
Gregor Girards facettenreiche Schulpädagogik 572

Besprechungen

Christine Schmid/Rainer Watermann

Wulf Hopf: Freiheit – Leistung – Ungleichheit. Bildung und soziale Herkunft
in Deutschland 589

Michaela Schmid

Hildegard Macha/Monika Witzke (Hrsg.):
„Familie – Kindheit – Jugend – Gender“.
Handbuch der Erziehungswissenschaft, Band III 592

Klaus Zierer

Karl-Heinz Arnold/Sigrid Blömeke/Rudolf Messner/
Jörg Schlömerkemper (Hrsg.): Allgemeine Didaktik und Lehr-Lernforschung –
Kontroversen und Entwicklungsperspektiven einer Wissenschaft
vom Unterricht 595

Dokumentation

Erziehungswissenschaftliche Habilitationen und Promotionen 2010 598

Pädagogische Neuerscheinungen 633

Impressum U 3

Table of Contents

Topic: Educational Policy between Expert Knowledge and Public Discussion

Marcelo Caruso

Educational Policy between Expert Knowledge and Public Discussion. An introduction	453
---	-----

Michele S. Moses

Education-Policy-by-Ballot-Initiative: Considerations Related to Democracy and Justice	458
---	-----

Lucien Criblez

Vox populi – On the challenges of semi-direct democracy with regard to educational policy	471
--	-----

Florian Waldow

Lawyers or Test Specialists? On the role of experts in establishing equity of grades in Germany and in Sweden	484
--	-----

Bernd Zymek/Sabine Wendt/Moritz Hegemann/Frank Ragutt

Regional Governance and Local School Development Policy in the Process of De- or Reconstructing Regional Structures of Educational Institutions	497
--	-----

Markus Maurer

The Myth of Free Education and the Problems of Legitimization Private Operators are confronted with in the Field of Higher Secondary Education in Sri Lanka	513
---	-----

Deutscher Bildungsserver

Tips on links relating to the topic of “Educational Policy between Expert Knowledge and Public Discussion”	526
---	-----

Contributions

Jürgen Wiechmann

The Implementation of the Comprehensive School (Gemeinschaftsschule) in Schleswig-Holstein – Systemic change and local decision making	534
---	-----

Tobias Diemer/Harm Kuper

Forms of School-Internal Control by means of Central Surveys on Student Achievement 572

Beat Bertschy

Unknown, Misunderstood, Underestimated – Gregor Girard’s multifaceted theory of education 572

Book Reviews 589

New Books 633

Impressum U 3

Tobias Diemer/Harm Kuper

Formen innerschulischer Steuerung mittels zentraler Lernstandserhebungen

Zusammenfassung: Zentrale Lernstandserhebungen sind Teil eines neuen, mit dem Schlagwort der „neuen Steuerung“ versehenen Steuerungsparadigmas. Anders als andere dem Paradigma zuordnende Instrumente wie internationale Schulleistungsvergleichsstudien oder nationale Bildungsberichterstattung tragen zentrale Lernstandserhebungen das neue Steuerungsparadigma direkt an Schulen und Lehrerinnen und Lehrer heran. Der Artikel geht auf der Grundlage einer qualitativen Interviewstudie der Frage nach, in welcher Weise im Zusammenhang der Rückmeldungen aus Lernstandserhebungen Steuerungsprozesse in Schulen stattfinden. Im Ergebnis zeigen sich dabei erhebliche Variationen von Steuerungslogiken, die sich nicht auf die paradigmatische Erwartung outputorientierter Steuerung reduzieren.

1. Zentrale Lernstandserhebungen und neue Steuerung – Einleitung

Neben der regelmäßigen Teilnahme an den internationalen Schulleistungsuntersuchungen PISA, TIMSS und PIRLS/IGLU, der zentralen, stichprobenbasierten Überprüfung des Erreichens der KMK-Bildungsstandards im Ländervergleich sowie der gemeinsamen Bildungsberichterstattung von Bund und Ländern gehören zentrale Lernstandserhebungen zu den vier wesentlichen Instrumenten der von der Kultusministerkonferenz der Länder im Jahr 2006 beschlossenen Gesamtstrategie zum Bildungsmonitoring (vgl. KMK, 2006). Wie jedes dieser Instrumente können Lernstandserhebungen als ein Mittel zur Verwirklichung dreier oft genannter Strukturmerkmale des mit dem Schlagwort der „neuen Steuerung“ versehenen zeitgenössischen Steuerungsparadigmas verstanden werden: der Standard-, der Evidenz- und der Outputorientierung. Hinter diesen Strukturmerkmalen stehen verschiedene Prinzipien und Vorstellungen, die die Strategie des zeitgenössischen Paradigmas neuer Steuerung kennzeichnen (vgl. Böttcher, 2002, 2004, 2006; Bellmann, 2006). Dazu zählen erstens die Dezentralisierung von Entscheidungen und die Stärkung der Autonomie von Schulen im Sinne der zunehmenden Verlagerung von Steuerungsprozessen in den Verantwortungsbereich von Einzelschulen. Zweitens gehört dazu die im Rahmen der empirischen Wende von Bildungspolitik und Bildungswissenschaften vertretene Forderung, Entscheidungen auf der Grundlage wissenschaftlichen empirischen Wissens zu treffen. Und drittens sind Bestrebungen der Sicherung und Steigerung von Effektivität und Effizienz im Bildungssystem zu nennen.

Da es sich bei zentralen Lernstandserhebungen im Unterschied zu stichprobenbasierten Schulleistungsvergleichsstudien um Vollerhebungen ganzer Jahrgangsstufen jeweils eines Bundeslandes handelt, liegt die Relevanz der empirischen Informationen dabei nicht nur auf den im Mehrebenensystem des Bildungswesens oberhalb der Einzelschule angesiedelten Systemebenen, sondern vor allem auf den Ebenen der Einzelschule und der Schulklasse (vgl. Bonsen & von der Gathen, 2004, S. 226-227). Durch zentrale

Lernstandserhebungen wird somit – im Unterschied zu den übrigen Instrumenten zum Bildungsmonitoring – das neue Steuerungsparadigma direkt an Schulen und Lehrerinnen und Lehrer herangetragen. Damit geht die Erwartung einher, dass in Schulen Formen standard-, evidenz- und outputorientierten Steuerungshandelns realisiert werden. Einem Positionspapier des Netzwerks Empiriegestützte Schulentwicklung (EMSE) zufolge sollen zentrale Lernstandserhebungen so u.a. die Funktion erfüllen, eine „stärkere kriteriale Orientierung [...] an den Standards“, „weitergehende Lernbedarfsdiagnosen“, die „Weiterentwicklung des Unterrichts“ sowie weitergehende Prozesse der Schulentwicklung wie etwa die „Revitalisierung der Fachgruppen und Fachkonferenzarbeit“ anzustoßen (EMSE, 2006, S. 1).

Ob und inwiefern derlei tatsächlich geschieht ist eine empirische Frage, die die praktische Nutzung der Informationen durch die Rezipienten betrifft. Zunächst stellen Lernstandserhebungen ein Testinstrument zur Erhebung von Schülerleistungen und zur Rückmeldung der Ergebnisse an die Schulen dar (vgl. Kuper & Schneewind, 2006). Zum Steuerungsinstrument werden sie erst, wenn die durch sie erhobenen Ergebnisse in der Verantwortung von Professionellen und Organisationen im Schulsystem zu Steuerungszwecken genutzt werden. In den bislang vorliegenden empirischen Studien zur innerschulischen Nutzung der Informationen aus Lernstandserhebungen wurden diesbezüglich bereits Ausprägungen, Zusammenhänge und interne Konfigurationen verschiedener Dimensionen wie bspw. Akzeptanz, Rezeptionsintensität, Verständlichkeit, Nützlichkeit, Nutzungsformen, Bereitschaft zur Nutzung der Ergebnisrückmeldungen für die Unterrichtsentwicklung, realisierte Konsequenzen usw. untersucht (vgl. Bellenberg, 2008; Helmke & Jäger, 2001; Schrader & Helmke, 2004; Groß Ophoff, Koch & Hosenfeld, 2006; Koch, Groß Ophoff & Hosenfeld, 2006; Groß Ophoff et al., 2007; Moser, 2003; Tresch, 2007; Schneewind, 2007; Bonsen, Büchter & Peek, 2006; Kühle & Peek, 2007; Nachtigall & Jantowski, 2007; Maier & Rauin, 2006; Maier, 2008 a, b). Sehr deutlich ist außerdem die Bedingtheit der Verwendung von Lernstandserhebungen durch die schulinternen Voraussetzungen herausgestellt worden (vgl. Kuper & Hartung-Beck, 2007; Hartung-Beck, 2009). Die im Folgenden vorgestellte qualitative empirische Studie fokussiert darüber hinaus die bislang offene Frage, inwiefern im Zusammenhang der Verarbeitung von Lernstandserhebungen in Schulen Formen innerschulischer Steuerung in Erscheinung treten und inwiefern sich in diesen das Paradigma neuer Steuerung spiegelt.

2. Begriff und Modell innerschulischer Steuerung

Die elementare Funktion von Steuerung besteht darin, Differenzen zwischen Eingangs- und Ausgangs- sowie zwischen tatsächlichen und angestrebten Systemzuständen zu vergrößern oder zu vermindern (vgl. Luhmann, 1994, S. 328). Übertragen auf den besonderen Fall von Schulen bedeutet Steuerung so, mittels der Gestaltung von Lehr- und Lernprozessen die Kompetenzentwicklung von Schülerinnen und Schülern zu fördern und dabei die Differenz zu Erwartungswerten zu vermindern bzw. die Differenz zum Aus-

gangszustand zu vergrößern. Diese allgemeine Bestimmung des Begriffs der Steuerung fügt sich in ein Modell, bei dem Systemzustände als Eingangsbedingungen (inputs) über dadurch bedingte Prozesse in daraus resultierende Ergebnisse (outputs, outcomes) transformiert werden (vgl. Dubs, 2004, S. 26ff.). Steuerung lässt sich dann allgemein beschreiben als Veränderung oder Einstellung von Zustands- und Prozessvariablen eines Systems mit dem Ziel der zweckgerichteten Beeinflussung der resultierenden Wirkungen dieses Systems. Eine wesentliche Bedingung der Steuerung besteht im Rahmen dieses Input-Prozess-Output-Modells darin, dass Eingriffe in das System auf Annahmen darüber basieren, wie diese sich auf Zustände (inputs), Prozesse oder Wirkungen (outputs) des Systems auswirken. Nach einer starken Lesart dieser Bedingung – die etwa im Rahmen kybernetischer Steuerungsmodelle vertreten wird – ist hierbei grundsätzlich eine verlässliche Kausalrelation zwischen Output einerseits sowie Input und Prozessen andererseits die Voraussetzung für steuernde Eingriffe. An das Prozess-Produkt-Paradigma angelehnte Steuerungskonzeptionen (vgl. Helmke, 2004) arbeiten hingegen lediglich mit schwachen Kausalannahmen, basieren aber gleichwohl auf der Idee, Relationen zwischen Bildungserfolg und den Prozeduren im System so weit zu spezifizieren, dass daraus im Sinne von Steuerung wirksame Rückschlüsse gezogen werden können.

Auf dieser Grundlage lassen sich wesentliche Charakteristika der Verwendung zentraler Lernstandserhebungen als Steuerungsinstrument, insbesondere im Hinblick auf Kernelemente des Begriffs neuer Steuerung, näher bestimmen. Allen voran steht dabei der Begriff der outputorientierten Steuerung. Ein System outputorientiert zu steuern heißt demnach, auf der Grundlage von Informationen über den Output eines Systems sowie auf der Grundlage von Annahmen über den Zusammenhang von Input-, Prozess- und Outputvariablen Eingriffe in Input- oder Prozessvariablen vorzunehmen. Das folgende Schaubild bietet eine bildliche Darstellung einiger Aspekte dieser Grundstruktur outputorientierter Steuerung:

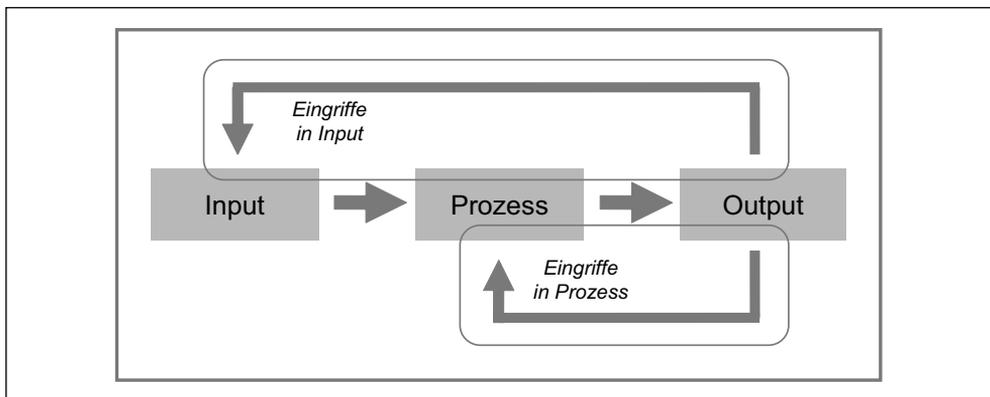


Abb. 1: Grundstruktur outputorientierter Steuerung

Des Weiteren lässt sich anhand des Modells feststellen, dass sich der Typus der outputorientierten Steuerung von input- und prozessorientierter Steuerung grundsätzlich nicht darin unterscheidet, was gesteuert wird. Der wesentliche Unterschied liegt vielmehr

darin, *wie* Input-, Prozess- und Outputvariablen zueinander in Beziehung gesetzt werden bzw. welche Interpunktion des zeitlichen Ablaufs der Ereignisse die Steuerung dominiert. Luhmann (2000, S. 261ff.) unterscheidet in diesem Zusammenhang zwischen Konditional- und Zweckprogrammierung, die dichotome strukturlogische Varianten der Relation kennzeichnen, in die Input und Prozess auf der einen und Output auf der anderen Seite gebracht werden können. Im Fall konditionalprogrammierender Steuerung werden im Bereich der Input- und Prozessvariablen Bedingungen gesetzt und die dann folgenden Konsequenzen im Licht der Bedingungen beurteilt. Im Fall outputorientierter Steuerung werden demgegenüber primär zukünftig zu realisierende Zwecke gesetzt oder identifiziert, um in deren Licht Input- und Prozessvariablen als Mittel zur Erreichung der Zwecke zu prüfen und zu bestimmen.

Charakteristisch für zweckprogrammierende Steuerung ist, dass prospektiv Zwecke gesetzt und die Wahl der Mittel in Abhängigkeit von der Zwecksetzung variabel ist. Besonders komplex wird die zeitliche Struktur dieses Steuerungstyps, wenn aus Erfahrungen über die Auswirkungen vergangener Inputs und Prozesse auf das Erreichen von Zielen auf die Gestaltung zukünftiger Inputs und Prozesse Rückschlüsse gezogen werden. Hier zeigt sich die aus Retrospektion beziehbare Lernfähigkeit zweckprogrammiert gesteuerter Systeme ebenso wie das Risiko dieses Steuerungstyps, aus falschen Annahmen über kausale Ursache-Wirkungs-Relationen heraus falsche Steuerungsentscheidungen zu treffen. Konditionalprogrammierte Systeme sind insofern einfacher strukturiert, als sie im Regelfalle keiner Retrospektion bedürfen. Inputs sind Auslöser von Prozessen, sie können daher in einem Anlass-Folge-Schema betrachtet werden, das nicht auf die komplizierten Voraussetzungen von Kausalitätsannahmen angewiesen ist. Die Lernfähigkeit konditionalprogrammierter Steuerung besteht darin, den Differenzierungsgrad in der Identifikation von Inputs zu erhöhen, die Aktivitäten im System auslösen; das Risiko dieses Steuerungstyps liegt darin, relevante Inputs zu vernachlässigen (vgl. Luhmann, 2000, S. 263ff.).

Ogleich Lernstandserhebungen als Instrument outputorientierter und damit im Sinne zweckprogrammierender Steuerung diskutiert werden, erscheint grundsätzlich auch eine konditionalprogrammierende Nutzung denkbar und möglich. Die durch Lernstandserhebungen erzeugten empirischen Informationen beziehen sich *prima facie* nicht zwangsläufig auf die Outputdimension des jeweiligen Systems, sondern können prinzipiell auch als Zustandsbeschreibungen der Input- oder der Prozessdimension verstanden werden. Je nachdem, wovon die Nutzung der Informationen ausgeht, kann so entweder konditional- oder zweckprogrammierendes Steuerungshandeln realisiert werden. Es ist deshalb eine offene Frage, ob und inwiefern Lernstandserhebungen tatsächlich als Instrument outputorientierter, zweckprogrammierender Steuerung realisiert werden.

Außerdem ist bemerkenswert, dass Lernstandserhebungen keine empirischen Informationen über die Zusammenhänge zwischen Input- und Prozess- sowie Prozess- und Outputdimension liefern (vgl. Stanat, 2008, S. 19). Im Hinblick auf den weiteren Begriff der evidenzbasierten Steuerung bedeutet dies, dass Lernstandserhebungen empirische Informationen und normative Beurteilungsmaßstäbe, aber noch kein rationales Steuerungswissen liefern. Für die Verwirklichung rationalen Steuerungshandelns ist

es vielmehr zusätzlich erforderlich, dass empirisches Wissen über den Zusammenhang von Outputs einerseits und Prozess- und Inputvariablen andererseits dezentral im Erfahrungskontext der Einzelschule als Grundlage für die Formulierung input- und prozessbezogener Entscheidungen und Maßnahmen gebildet wird. Steuerung auf der Grundlage von Instrumenten, die ausschließlich empirische Informationen über einen Teilbereich eines Systems liefern, ist somit grundsätzlich mit Unsicherheit behaftet (S. 17). Dies gilt sowohl für die retrospektive, durch Zweckprogrammierung geleitete Analyse von Gründen und Ursachen, also die Rückbeziehung von Outputinformationen auf Input- und Prozessmerkmale, als auch für die konditionalprogrammierende, prospektive Formulierung geeigneter erscheinender Maßnahmen und Entscheidungen auf der Grundlage von Input- und Prozessinformationen.

Ein Grund hierfür liegt zum einen im faktisch stets mehr oder weniger begrenzt verfügbaren und erzeugbaren empirischen Wissen. Darüber hinaus sind zum anderen aber auch zwei grundsätzliche strukturelle Charakteristika von Schulen als pädagogische Organisationen zu beachten, die mit den Schlagworten der „losen Kopplung“ (engl.: loose coupling; Weick, 1976; Orton & Weick, 1990) und des „Technologiedefizits“ (Luhmann & Schorr, 1979) belegt sind. Der Begriff der losen Kopplung bezeichnet allgemein die Verbindung zweier Systeme oder Systemelemente, bei denen die unabhängigen Variablen auf der einen Seite in mehr oder weniger schwachen und offenen Wirkungsbeziehungen zu den abhängigen Variablen auf der anderen Seite stehen. Im Rahmen des umrissenen allgemeinen Steuerungsmodells bezieht sich der Begriff der losen Kopplung zunächst auf das Verhältnis von Input- und Prozessdimension. Das darauf folgende Verhältnis von Prozess- und Outputdimension wird des Weiteren mit dem Begriff des Technologiedefizits durch einen spezielleren Begriff der losen Kopplung charakterisiert. Kern dieses Begriffs ist die Feststellung, dass sich die im Interaktionssystem des Lehrens und Lernens stattfindenden Prozesse durch doppelte Kontingenz im Sinne der Referenz auf psychische und soziale Systeme (Personen und Kommunikation) auszeichnen, die sich jeweils auf kontingente, d.h. nicht vollkommen vorhersehbare Weise verhalten. Wegen der daraus resultierenden Unmöglichkeit einer auf Kausalplänen mit strikt kontrollierbaren Ursache-Wirkungs-Relationen basierenden Unterrichtstechnologie erweist sich auch die Beziehung zwischen Prozess und Output als grundsätzlich mehr oder weniger lose gekoppelt.

Insgesamt stellt sich Steuerung auf der Grundlage von Lernstandserhebungen damit als Aufgabe dar, die Professionalität im Sinne der konstruktiven und situationspezifischen Deutung und Anwendung des aus Lernstandserhebungen stammenden wissenschaftlichen Wissens und des kompetenten Umgangs mit den darin involvierten Unsicherheiten erfordert (vgl. Tenorth, 2006). Im Besonderen bezieht sich dies im Teilbereich inputbezogenen Steuerungshandelns auf die Profession des Leitens und des Managements der Organisation sowie im Teilbereich prozessbezogenen Steuerungshandelns auf die Profession des Lehrens und Erziehens im Rahmen des Interaktionssystems Unterricht.

Die im Vorangegangenen ausgeführten Überlegungen bilden insgesamt eine heuristische Matrix im Sinne einer strukturierten Darstellung des Möglichkeitsraums der

innerschulischen Verwendung zentraler Lernstandserhebungen. Diese Matrix dient im Folgenden als Grundlage für empirische Untersuchungen zu der Frage, inwiefern Lernstandserhebungen von ihren Adressaten in Schulen tatsächlich zu Steuerungszwecken genutzt werden. Im Einzelnen thematisiert die Studie folgende Untersuchungsfragen:

- *Inwiefern werden Informationen aus Lernstandserhebungen als Outputinformationen wahrgenommen?*
- *Inwiefern wird im Zusammenhang der innerschulischen Verarbeitung der Ergebnisse aus Lernstandserhebungen input- und prozessbezogene Ursachenforschung betrieben?*
- *Inwiefern wird des Weiteren input- und prozessbezogenes Steuerungshandeln realisiert?*

3. Gegenstand und Methode der Studie

Die empirische Grundlage der Studie bildet eine qualitative Interviewerhebung, die über einen Zeitraum von zwei Jahren (April 2007 bis März 2009) im Rahmen eines von der Deutschen Forschungsgemeinschaft (DFG) geförderten Forschungsprojekts mit dem Titel „Die Auswirkungen zentraler Lernstandserhebungen auf die professionelle Arbeit von Lehrern und die Organisation von Schulen“ durchgeführt wurde. An der Erhebung nahmen vier Sekundarschulen aus zwei Bundesländern (ein Gymnasium und eine Gesamtschule in Berlin sowie ein Gymnasium und eine Regelschule in Thüringen) teil, die nach einem theoretical sampling (Glaser & Strauss, 1967) ausgewählt wurden. Um ein größtmögliches Spektrum von Nutzungsphänomenen erfassen zu können, sollten die Schulen aktive Nutzungsprozesse erwarten lassen und im Hinblick auf ihr Leistungsniveau bei den jeweiligen Lernstandserhebungen deutlich kontrastieren. Ersteres wurde über Selbsteinschätzungen der Schulen kontrolliert, letzteres über die Unterscheidung von leistungsbezogen positiv und negativ selektiven Schultypen. Bei den jeweils in den Interviews thematisierten zentralen Lernstandserhebungen handelte es sich in den beiden Thüringer Schulen um den in Jahrgangsstufe 6 in den Fächern Deutsch, Englisch und Mathematik sowie in Jahrgangsstufe 8 im Fach Mathematik durchgeführten Kompetenztest und in den beiden Berliner Schulen um den in Jahrgangsstufe 10 in den Fächern Deutsch, Englisch und Mathematik durchgeführten Mittleren Schulabschluss (MSA).

Ein wesentlicher Unterschied dieser beider Lernstandserhebungen besteht darin, dass der Thüringer Kompetenztest auf der Ebene der Schüler einen sog. low stakes test (vgl. Mitchell, 2006) darstellt, der sich ausschließlich als Diagnoseinstrument ohne bzw. mit nur schwachem Notenbezug versteht, wohingegen der Berliner MSA gleichzeitig ein Prüfungsverfahren mit Zertifikatsstatus und damit auf der Ebene der Schüler einen high stakes test darstellt. Darüber hinaus unterscheiden sich beide Tests deutlich im Hinblick auf Differenziertheit und Komplexität der Ergebnisrückmeldung. Während die MSA-Rückmeldung aus der Mitteilung von Bestehensquoten sowie von Lösungs-

anteilen bezogen auf die getesteten Fächer und von durchschnittlichen Jahrgangs- und Prüfungsnoten besteht, umfasst die Ergebnismeldung des Thüringer Kompetenztests ein Spektrum unterschiedlicher Ergebnisdarstellungen: vom Gesamtergebnis der Schule bzw. jeweiligen Klasse je Fach, den Gesamtergebnissen in den Teilaufgaben der Tests, über die Verteilung der erreichten Punktzahlen je Klasse und einen Vergleich der Testergebnisse mit der Halbjahresnote, bis hin zur Darstellung der Ergebnisse differenziert nach Kompetenzen und Anforderungsbereichen. Hinzu kommt, dass bei den Kompetenztestrückmeldungen im Sinne des Begriffs des fairen Vergleichs (vgl. Nactigall & Kröhne, 2006) adjustierte Landesmittelwerte in die Ergebnisdarstellung integriert sind, wohingegen im Berliner Fall zusätzliche Listen mit entsprechenden auf Bezirks- und Landesebene aggregierten und nach Schulformen differenzierten nicht-adjustierten Mittelwerten zum Vergleichen zur Verfügung standen.

Neben solchen deutlichen Unterschieden weisen die beiden Testverfahren des Weiteren einige wesentliche Gemeinsamkeiten auf, die der Grund sind, warum beide Verfahren trotz der erheblichen Unterschiede als vergleichbare Instanzen des Konzepts zentraler Lernstandserhebungen zusammengefasst und untersucht werden können. Erstens handelt es sich in beiden Fällen zwar nicht auf der Ebene der Schüler, aber der der Organisation um sog. low-stakes Tests, bei denen mit über- oder unterdurchschnittlichen Ergebnissen keine externen Sanktionen oder Gratifikationen verbunden werden. Zweitens orientiert sich die Entwicklung und Formulierung der jeweiligen Testaufgaben in beiden Fällen inhaltlich und fachdidaktisch an den kompetenzorientierten Anforderungen der länderübergreifenden KMK-Bildungsstandards. Und drittens beinhalten beide Verfahren über die Rückmeldung landesweiter Referenzwerte einen gleichartigen vergleichenden Sozialnormbezug.

Im Rahmen der Studie wurden in den vier Schulen Interviews mit den Schulleitern, den für die schulinterne Durchführung der Erhebungen verantwortlichen Koordinatoren sowie mit Fachgruppenleitern und einzelnen Fachlehrern der Fächer Deutsch, Mathematik und Englisch geführt. Die Befragung erfolgte in Form problemzentrierter Interviews (Merton & Kendall, 1979; Hopf, 1995) auf der Basis eines strukturierten Leitfadens. Die Erhebung erstreckte sich auf drei über einen Zeitraum von eineinhalb Jahren verteilte Erhebungsphasen. Die thematischen Schwerpunkte änderten sich dabei von Erhebungszeitpunkt zu Erhebungszeitpunkt entlang der Unterscheidung von vier Phasen der Verarbeitung von Informationen aus Lernstandserhebungen, wie sie im Rahmen eines von Helmke und Schrader (2001) vorgeschlagenen und in der Folge in verschiedenen quantitativen Studien (z.B. WALZER: vgl. Schrader & Helmke, 2004; VERA: vgl. Groß Ophoff et al., 2006) bereits eingesetzten heuristischen Rahmenmodells unterschieden wurden: die Rezeption im Sinne der Informationsentnahme und -verarbeitung (1. Erhebungswelle), die Reflexion im Sinne der Analyse von Gründen und Ursachen (2. Erhebungswelle), die Aktion im Sinne der Ableitung und Durchführung von Maßnahmen sowie die Evaluation im Sinne der Überprüfung der Richtigkeit der Reflexions- und der Wirksamkeit der Aktionsphase (3. Erhebungswelle). Für die Auswertung wurden die aufgezeichneten Interviews normalsprachlich transkribiert. Die Materialbasis umfasst insgesamt 70 Einzel-, Doppel- und Dreierinterviews bzw. 86 Interviewein-

heiten (1 Doppelinterview entspricht dabei 2 Intervieweinheiten, 1 Dreierinterview entspricht 3 Intervieweinheiten).

Die Auswertung des Materials folgte dem Verfahren der qualitativen Inhaltsanalyse nach Mayring (1996). In einem ersten von zwei Schritten wurde das jeweils in einer Erhebungswelle erhobene Material strukturierend auf der Basis theoriegeleitet festgelegter und mit Definitionen und Kodierregeln versehener hierarchisch geordneter Strukturierungsdimensionen kodiert. Nach der Kodierung der Interviews der ersten Welle wurde das Kodiersystem revidiert und ein Konzept systematischer, kompositorischer Mehrfachkodierungen eingeführt. In einem zweiten, auf den Abschluss der Kodierarbeiten der dritten Welle folgenden Schritt wurden die Kategorien zweiter Ordnung auf induktive Weise weiter strukturiert. Darauf schließlich folgte die Erstellung einer deskriptiven Grundausswertung des Materials (Diemer, Rucht, Schulze & Kuper, 2009), die eine zusammenfassende Darstellung der in den Interviews zu findenden Informationen über den Bearbeitungsprozess differenziert nach den Kategorien des dreistufigen Kategoriensystems bietet. Die beschriebenen Kodierungs- und Auswertungsschritte wurden in arbeitsteiliger Teamarbeit von drei Personen durchgeführt; Validität und Reliabilität der Kodierung und Auswertung wurden dabei fortlaufend durch wechselseitiges Lektorieren und Überarbeiten kommunikativ kontrolliert.

Die so entstandene Grundausswertung stellt zum einen eine systematische inhaltliche Dokumentation der erhobenen Informationen des Projekts dar. Zum anderen erschließt sie das Material für weiter- und tieferreichende Analysen und Interpretationen wie die im Folgenden unternommenen.

4. Ergebnisse aus den Fallstudien

4.1 *Inwiefern werden Informationen aus Lernstandserhebungen als Outputinformationen wahrgenommen?*

Mit den Rückmeldungen der Ergebnisse aus Lernstandserhebungen erhalten Lehrer und Schulen verschiedene Informationen. Je nach Funktionsrolle der Rezipienten zeigen sich allgemein unterschiedliche Tendenzen der Wahrnehmung dieser Informationen. So dominiert auf der Ebene der einzelnen Fachlehrer und der Fachschaftsleiter deutlich die Bezugnahme auf die Ergebnisse der jeweils einzelnen Klassen in den einzelnen Aufgaben und Aufgabenbereichen des Tests sowie auf die Ergebnisse einzelner individueller Schüler. Auf der Ebene der Koordinatoren, die in der Regel selbst auch an den Tests beteiligte Fachlehrer sind, und der Fachschaftsleitungen als Leiter der Fachkonferenzen, richtet sich die Aufmerksamkeit darüber hinaus auf den schulinternen Vergleich zwischen Parallelklassen und den Vergleich zwischen den Fächern. Und auf der Ebene der Schulleitung und der Koordinatoren in ihrer Rolle als Beauftragte der Schulleitung ist ein besonderes Interesse an den Gesamtergebnissen der Fächer und dem Gesamtergebnis der Schule im Vergleich zu den Landesmittelwerten sowie zu Schulen vergleichbarer

Schultypen im näheren Umfeld der Schule, also im Bezirk, der Stadt, der Gemeinde und in benachbarten Bezirken, Städten und Gemeinden zu verzeichnen.

Die Aufgabenergebnisse werden von den Fachlehrern und Fachschaftsleitern insbesondere daraufhin wahrgenommen, bei welchen Aufgaben bzw. in welchen Aufgabebereichen besondere Stärken und Schwächen zu verzeichnen sind (Belegstellen (BS) in 42 Intervieweinheiten (IE)). Wenn Landesmittelwerte auf Aufgabenebene als Vergleichswerte berichtet werden, werden diese bisweilen als besondere Referenz zur Einschätzung und Bewertung der einzelnen Aufgabenergebnisse herangezogen (BS in 13 IE).

Das find ich auch recht interessant, dass man mal guckt, wie liegt man im LANDES-durchschnitt. ... So dass man sagen kann: Also hier liegen wir unterdurchschnittlich, hier liegen wir überdurchschnittlich, hier überdurchschnittlich, hier wieder unterdurchschnittlich. So dass man sieht: Wo hält es sich die Waage. Und wo (...) liegt man eben krass drunter. (GS1_1_FSLD1)

Des Weiteren fällt auf, dass bei der Betrachtung der Aufgabenergebnisse die Identifikation negativ entweder vom Landesmittelwert oder den eigenen Erwartungen abweichender, defizitärer Bereiche besonderes Interesse auf sich zieht (BS in 25 IE).

Es waren Aufgaben dabei, (...) wo man gemerkt hat, die haben sie alle relativ gut bewältigt. Und interessant sind natürlich die, wo es eben genau entgegengesetzt war (...). Also da gab's so bestimmte Aufgabentypen, wo sie richtig, die gesamte Klasse, runter ging... Also an bestimmten Aufgaben, da haben wir das dann gesehen, wo wir gesagt haben: Aha, das sind wirklich Schwerpunkte, wo wir noch mal Wert drauf legen müssen. (GY1_2_FLM1)

Wie die zitierte Passage beispielhaft zeigt, schließt sich an die Rezeption der Ergebnisse auf Aufgabenebene oft in direkter Weise, d.h. ohne zwischengeschaltete Analyse von Gründen und Ursachen, eine Art von Maßnahme an: die verstärkte Behandlung von Aufgabeninhalten und -formaten, bei denen die Schüler im Test schlecht abgeschnitten haben (BS in 45 IE). Dies ist insofern bemerkenswert, als die Informationen so unmittelbar in die Gestaltung der Prozesse des Lehrens integriert werden. Die Information scheint in diesem Fall weniger als ein Ergebnis vorangegangener Prozesse, also als Outputinformation, denn vielmehr als Zustandsbeschreibung des aktuellen Prozesses, also als Prozessinformation, rezipiert zu werden.

Anders verhält es sich, wenn die Ergebnisse im Rahmen von schulinternen Vergleichen zwischen parallelen Klassen rezipiert werden (BS in 28 IE). Ein solcher Vergleich wird in der Regel vorgenommen, um Ursachenforschung zu betreiben. Ziel ist es, feststellbare Unterschiede zwischen Klassen zu erklären, indem Bedingungen identifiziert werden, die die Ergebnisse positiv oder negativ beeinflusst haben könnten (BS in 13 IE).

In der Schule ist das insofern wichtig für mich, diese Vergleichsarbeiten, dass man sagen kann: Woran lag es denn, dass bei dir das so ist und bei mir ist das Ergebnis so. Muss ja Gründe dafür geben. (GS1_1_FLD3)

Das heißt, die Ergebnisse werden in diesem Zusammenhang deutlich als Produkt von Input- und Prozessvariablen und auf diese Weise als Outputinformationen wahrgenommen.

Eine andere Art der Rezeption der Ergebnisse als Outputinformationen lässt sich bei der Rezeption des Schulergebnisses identifizieren. Dieses wird zum Teil vergleichend in Beziehung zum globalen Landesmittelwert gesetzt (BS in 24 IE). Zum Teil werden die Ergebnisse der Schule, so diese Informationen geliefert werden, mit den Mittelwerten von Schulen des gleichen Typs aus dem selben Gebiet (Stadt, Bezirk, Region) (BS in 11 IE) oder auch über verschiedene personen- oder funktionsrollenbezogene kommunikative Kanäle mit den Ergebnissen einzelner Schulen in der näheren Umgebung verglichen (BS in 9 IE). Die Aufmerksamkeit liegt bei diesen Vergleichen vornehmlich auf der Ergebnisinformation selbst, eine Bezugnahme auf Input- und Prozessvariablen findet in diesem Zusammenhang kaum statt, wie ein Schulleiter gleichermaßen deutlich wie kritisch feststellt:

Ich muss auch sagen, ich spüre hier in LAND so nach und nach einen Trend, wo ausschließlich auf der Grundlage bestimmter Zahlen diskutiert wird und dann auch wirklich mit Zahlen hantiert wird, (...). Das ist auch so ein Gefühl, das sich entwickelt hat, dass ich sage: Hier wird abstrakt nicht der Prozess, sondern das Ergebnis und das Ergebnis auch nur anhand von Zahlen gewertet. (GS1_1_2_SL2)

Zusammengefasst zeigt sich: Erstens finden sich im Material der Studie deutlich erkennbare Hinweise auf eine in bestimmten Kontexten dominante Wahrnehmung der Informationen als Prozessinformationen. Dies ist ein empirischer Beleg für die konzeptionelle These, dass Informationen aus Lernstandserhebungen nicht zwangsläufig als Outputinformationen wahrgenommen werden. Werden die Informationen tatsächlich als Outputinformationen wahrgenommen, können zwei Formen unterschieden werden. Die eine dieser Formen zeichnet sich aus durch die Einbettung in weitergehende Reflexionsprozesse, die ein notwendiges Strukturmerkmal outputorientierter Steuerung darstellen. Bei der zweiten Form wird der Wert der Informationen im Unterschied hierzu abgekoppelt von der reflexiven Relationierung auf Input- und Prozessdimension im Sinne eines reinen Benchmarkvergleichs taxiert.

4.2 *Inwiefern wird im Zusammenhang der Auseinandersetzung mit den Ergebnissen aus Lernstandserhebungen input- und prozessbezogene Ursachenforschung betrieben?*

Die Suche nach Gründen und Ursachen für das Zustandekommen der jeweiligen Ergebnisse gestaltet sich vielfältig. Im Interviewmaterial ließen sich hierzu sechs Kategorien identifizieren, die jeweils intern weiter ausdifferenziert werden konnten. Als Gründe und Ursachen wird demnach verwiesen auf die Schüler (z.B. Leistungsniveau der Klasse, Anstrengungsbereitschaft, Verhältnis zwischen Lehrer und Schüler, Wissen und Fertigkeiten, schulische Voraussetzungen, individuelle Besonderheiten), auf die Lehrer (z.B. allgemeine Arbeitsweise, Beurteilungspraxis, persönliche Schwerpunktbildungen), auf den Unterricht (z.B. methodische Art des Unterrichts, Testvorbereitung, Inhalte des Unterrichts, spezifische Rahmenbedingungen), auf Aspekte der Organisation (z.B. Zusammensetzung der Schülerschaft, Lehrereinsatzplanung, Klassengrößen, Unterrichtsausfall, Zusatzunterricht), auf die Tests (z.B. Schwierigkeitsniveau, Inhalte, Aufgabenformate und Aufgabenstellungen, Testsituation) und auf verschiedene gesellschaftliche Rahmenbedingungen (z.B. soziales Umfeld und außerschulische Sozialisation, Eltern, politische Entscheidungen, gesellschaftliche Entwicklungen).

Inwiefern die Ergebnisse aus Lernstandserhebungen input- und prozessbezogen reflektiert werden, lässt sich darüber beantworten, inwiefern die Themen der Subkategorien als Inhalte der jeweiligen Kategorien Input- oder Prozessvariablen repräsentieren. Input- und Prozessdimension lassen sich in diesem Zusammenhang folgendermaßen unterscheiden: Prozessvariablen zeichnen sich dadurch aus, dass sie für Lehrer als Einzelpersonen in der Interaktion des Lehrens und Lernens im Unterricht direkt verfügbar sind. Inputvariablen hingegen sind entweder nur über koordinierte Kommunikationen und Entscheidungen oder überhaupt nicht verfügbar. Anhand dieser Unterscheidung lassen sich zunächst einmal die drei Kategorien „Aspekte der Organisation“, „Tests“ und „gesellschaftliche Rahmenbedingungen“ geschlossen der Inputdimension zuordnen. Die Kategorien „Schüler“, „Lehrer“ und „Unterricht“ sind darüber hinaus Kategorien, die prima facie beiden Seiten zugeordnet werden können. Betrachtet man die darin jeweils unterschiedenen Subkategorien und ordnet diese den Dimensionen zu, so lassen sich charakteristisch unterschiedliche Ausprägungen der Kategorien feststellen. Von den in der Kategorie „Schüler“ unterschiedenen Subkategorien lassen sich die meisten (Leistungsniveau der Klasse, Anstrengungsbereitschaft, Wissen und Fertigkeiten, schulische Voraussetzungen individuelle Besonderheiten) der Input- und nur eine (Verhältnis Schüler-Lehrer) der Prozessdimension zuordnen. Bei den Kategorien „Lehrer“ und „Unterricht“ verhält es sich hingegen umgekehrt; hier lassen sich die meisten Kategorien der Prozessdimension (Lehrer: Arbeit des Lehrers, Beurteilungspraxis, persönliche Schwerpunktbildungen; Unterricht: methodische Art des Unterrichts, Testvorbereitung, Inhalte des Unterrichts) und nur eine Subkategorie (Unterricht: spezifische Rahmenbedingungen) der Inputdimension zuordnen. Insgesamt ist damit zum einen festzustellen, dass bei der Auseinandersetzung mit den Ergebnissen aus Lernstandserhebungen sowohl input- als auch prozessbezogene Ursachenforschung ohne erkennbare globale

Tendenzen zugunsten einer der beiden Dimensionen betrieben wird. Gleichzeitig ist zum anderen im Hinblick auf die Kategorien „Schüler“, „Lehrer“ und „Unterricht“ im Besonderen aber auch zu bemerken, dass die Rolle der Schüler vorwiegend im Sinne einer Inputvariable, die Rolle des Lehrers sowie des Unterrichts hingegen vorwiegend im Sinne von Prozessvariablen thematisiert werden.

Dieser besondere Befund sei anhand zweier Beispiele etwas konkreter veranschaulicht. Bspw. wird beim Rekurs auf die Kategorie der Schüler das Leistungsniveau der Klasse im Sinne der Zuschreibung einer relativ stabilen, nicht oder kaum beeinflussbaren Eigenschaft thematisiert, die sich aus der jeweiligen Schülerzusammensetzung ergibt (BS in 51 IE).

Also bestimmte Kurse oder Klassen kommen manchmal so zusammen, sind besonders leistungsschwach. Und es gibt auch wieder jetzt Klassen, also da staunt man vom Leistungsniveau her wie gut die sind. ... Also, es hängt schon damit zusammen. (GS1_1_FLD3)

Dass umgekehrt die Rolle des Lehrers sowie des Unterrichts überwiegend im Sinne einer Prozessvariable thematisiert wird, bedeutet hingegen z.B., dass bei der Ursachenanalyse auf die Auswahl der Unterrichtsinhalte durch den Fachlehrer verwiesen wird (BS in 21 IE).

[I]ch konnte eigentlich schon im Vorfeld sagen: Hier wird's gut laufen. Hier wird's eher nicht so gut laufen... Das hab' ich mit ihnen nicht viel gemacht (...). Also wenn sie das nicht konnten, liegt das zum Teil an ihnen, aber eben auch an mir, weil ich das (...) nicht so viel gemacht habe. (RS1_1_FLM5)

Die Beobachtung, dass Schüler tendenziell im Sinne einer Inputvariable und Lehrer und Unterricht tendenziell im Sinne von Prozessvariablen beschrieben werden, erscheint insbesondere insofern bedeutsam, als andere Konstellationen möglich und denkbar sind. Zum Beispiel wäre es im Bereich der Kategorie „Schüler“ möglich, dass die besonderen Lernentwicklungen und -schwierigkeiten der Schüler im Verlauf des Schuljahrs stärker thematisiert würden. Oder im Bereich der Kategorie „Lehrer“ wäre es denkbar, dass über Inputvariablen wie Fortbildung, Methodeninventar, Hintergrundwissen, über Bildungsstandards und kompetenzorientierten Unterricht usw. gesprochen würde. Ferner könnte man sich im Bereich der Kategorie „Unterricht“ vorstellen, dass neben spezifischen Rahmenbedingungen über Inputvariablen z.B. im Sinne des Lernklimas im Klassenzimmer oder der Gestaltung von Lernumgebungen und Gelegenheitsstrukturen gesprochen würde. Dass derlei jedoch kaum zu beobachten ist, ist möglicherweise mit dem Vorherrschen eines eher konventionellen Bildes von Unterricht erklärbar, in dem das Verhalten der Lehrkräfte im Zentrum der Wahrnehmung des Unterrichtsprozesses steht. Das hieße im Umkehrschluss, dass hier Unterricht nicht wie in einschlägigen aktuellen Konzeptionen als „ko-konstruktiver Wachstums- und Entfaltungsprozess“ (Fend, 2001, S. 57) begriffen wird, für dessen Erfolg neben der Lehrperson auch die Ge-

staltung unterrichtlicher Gelegenheitsstrukturen und die längerfristigen individuellen Entwicklungs- und Lernprozesse der Schüler als bedeutsame und gestaltbare Größen erachtet werden (vgl. Dubs, 2004, S. 40).

Zusammenfassend lässt sich so festhalten: Zum einen ist keine globale Tendenz der Suche nach Gründen und Ursachen entweder vor allem im Input- oder im Prozessbereich festzustellen. Gleichzeitig deuten die Befunde zum anderen darauf hin, dass durch die Rückmeldungen der Informationen sowohl input- als auch prozessbezogene Reflexionsvorgänge angeregt werden können. Diese Vorgänge scheinen vorgängig existierenden Reflexionsmustern zu folgen die auf die Erklärung der zurückgemeldeten Informationen übertragen werden.

4.3 Inwiefern wird im Zusammenhang der Auseinandersetzung mit den Ergebnissen aus Lernstandserhebungen input- und prozessbezogenes Steuerungshandeln realisiert?

Während sich im Bereich der Reflexion der Ergebnisse die Bezugnahme auf Input- und Prozessdimension nicht erkennbar ungleich verteilt, ist im Bereich der Ableitung von Maßnahmen eine deutliche Tendenz zur stärkeren Bezugnahme auf die Prozessdimension festzustellen. So stehen im Vordergrund der berichteten Maßnahmen verschiedene Aspekte der Gestaltung des Unterrichts, die gezielte Vorbereitung auf den Test, die stärkere individuelle Förderung der Schüler und die Verwendung einheitlicher Bewertungskriterien – Maßnahmen also, die sich direkt auf den Prozess des Unterrichtens beziehen. Als Aktionen, die sich auf die Dimension des Inputs beziehen, wird über organisatorische Maßnahmen zur Änderung schulinterner Rahmenbedingungen sowie über Veränderungen der Formen schulinterner Kooperation und Kommunikation berichtet.

Wie im Bereich der Rezeption der Ergebnisse, so dominiert auch bei der prozessbezogenen Ableitung von Maßnahmen das Thema der im Test unterdurchschnittlich ausgefallenen Aufgabenbereiche. Die Reaktion besteht in Übereinstimmung mit der im Bereich der Rezeption festgestellten Nutzungspraxis darin, diese Inhalte zu wiederholen und in Zukunft verstärkt im Unterricht zu behandeln (BS in 45 BS). Darüber hinaus werden die Tests vielen Aussagen zufolge zum Anlass dafür, dass verstärkt den Testitems ähnliche oder gleichartige Aufgabenformate im Unterricht praktiziert werden (BS in 21 IE), dass eine Abkehr von stoff- und eine Hinwendung zu kompetenzorientiertem Unterrichten stattfindet (BS in 19 IE), dass auf den Test mittels entsprechender Übungen im Vorfeld gezielt vorbereitet wird (BS in 33 IE) und bisweilen auch, dass die schulinternen Bewertungskriterien den testseitigen Bewertungskriterien angepasst werden (BS in 7 IE). Wie bei der Rezeption der Ergebnisse bei den einzelnen Aufgaben der Tests lässt sich somit feststellen, dass die Realisierung solcher Maßnahmen direkt, ohne Reflexionsschleifen, erfolgt. Die Tests werden zum Anlass für direkte prozessbezogene Steuerungseingriffe und auf diese Weise direkt in die Prozessgestaltung integriert. Bei der übrigen prozessbezogenen Maßnahmenart – der stärkeren individuellen Förderung einzelner Schüler – spielen demgegenüber reflexive Momente eine größere Rolle. Die Entscheidung zur individuellen Förderung basiert in der Regel insbesondere auf der

Feststellung besonders heterogener Leistungsverteilungen in der Klasse. Entsprechend wird eine stärker individualisierende Förderung in diesem Zusammenhang als Mittel betrachtet, um das Niveau leistungsschwacher Schüler zu erhöhen oder auch um die Leistungsspitze zu verbreitern (BS in 16 IE).

Im Bereich der inputbezogenen Maßnahmen fällt des Weiteren auf, dass überwiegend prozessnahe Themen bearbeitet werden. So wird etwa als eine organisatorische Maßnahme zur Änderung schulinterner Rahmenbedingungen über die Gestaltung und Abstimmung schulinterner Curricula gesprochen (BS in 11 IE).

Eine andere bisweilen erwähnte organisatorische Maßnahme besteht in der Abstimmung und Vereinheitlichung didaktischer Vorgaben (BS in 10 IE). Weitere, ebenfalls inputbezogene Veränderungen beinhalten die Verstärkung informeller sowie formeller, auf Jahrgangs- oder Fachschaftsebene angesiedelter schulinterner Kooperation und Kommunikation (BS in 33 IE). Auch diese Folgen beziehen sich auf prozessnahe Themen wie die Schwerpunktsetzung bei den Lehrinhalten oder der Aufgabenanalyse, eine verstärkte Reflexion der Unterrichtsgestaltung oder den verstärkten Austausch von unterrichtsbezogenen Informationen und Materialien.

Über weniger prozessnahe inputbezogene Maßnahmen wird zwar auch gesprochen, dann allerdings im Sinne unverfügbarer, weil extern vorgegebener Strukturen und Bedingungen (BS in 11 IE). Dazu gehören etwa das Verhalten und das Leistungspotential der Schüler, Vorgaben des Schulamtes, Klassengrößen, allgemeine Arbeitsbedingungen, die Schülerzusammensetzung, fehlende finanzielle und personelle Ressourcen usw. Schließlich wird in einzelnen Interviews darauf hingewiesen, dass die Maßnahmenableitung zusammen mit der Ursachenforschung eine komplexe und schwierige Aufgabe darstellt, die mit vielen Unsicherheiten verbunden ist (BS in 4 IE). Insbesondere wird hier festgestellt, dass man sich in einem spekulativen Rahmen bewege und der Test auf Grund des Fehlens entsprechender Kontextinformationen keine hinreichende Ursachenforschung und keine Ableitung weiterreichender Maßnahmen ermögliche.

Insgesamt zeigt sich so hinsichtlich der Frage der Realisierung von Steuerungshandeln im Zuge der innerschulischen Verarbeitung von Lernstandserhebungen wiederum zweierlei: So werden zum einen, korrespondierend zu Befunden im Bereich der Rezeption, prozessbezogene Maßnahmen tendenziell direkt, unter Auslassung des Schrittes reflexiver Ursachenanalysen und damit nicht im Sinne outputorientierter Steuerung realisiert. Ähnliches gilt im Bereich der Realisierung inputbezogener Maßnahmen, die sich sehr prozessnah auf Aufgaben, Inhalte, Bewertungskriterien, auf Unterricht also, beziehen. Zum anderen sind bezüglich weniger prozessbezogener Ansätze im Bereich inputbezogener Maßnahmen eine Konzentration der Aufmerksamkeit auf extern vorgegebene Strukturen sowie Hinweise auf eine grundlegende Verlegenheit zu verzeichnen, die Verarbeitung von Lernstandserhebungen in den Kontext weiterreichender Professions- und Schulentwicklungsprozeduren zu stellen.

5. Diskussion und Fazit

Die präsentierten empirischen Ergebnisse über die Nutzungspraktiken zentraler Lernstandserhebungen weisen auf erhebliche Variationen der Argumentations- und Steuerungslogiken hin, die von den Lehrkräften an den Schulen realisiert werden. Outputorientierte Steuerungsentscheidungen – entsprechend den Intentionen und der verbreiteten Charakterisierung zentraler Lernstandserhebungen als Instrument neuer Steuerung – bilden nur einen, in sich wiederum variablen Typ von Steuerungshandeln, der im Anschluss an die Rückmeldung von Testergebnissen realisiert wird. Empirisch nachweisen lassen sich auch input- und prozessorientierte Steuerungshandlungen. Sofern Lehrerinnen und Lehrer durch die Konzeptionen von Lernstandserhebungen als maßgebliche Steuerungsakteure adressiert werden, liegt demnach eine Übergeneralisierung vor, wenn die Einführung von Verfahren zentraler Leistungsmessung und Ergebnisrückmeldung an die Schulen unter der Prämisse von Outputorientierung reflektiert wird. An diesen Ausgangspunkt lassen sich Überlegungen anschließen, die sowohl die Rezeptionsforschung als auch die praktische Gestaltung von Lernstandserhebungen als Reformmaßnahme im Bildungssystem betreffen.

Auf der Grundlage der vorliegenden Ergebnisse können selbstverständlich keine Aussagen über die Bedeutsamkeit der hier identifizierten input-, prozess- und outputorientierten Steuerungshandlungen im Sinne ihrer Verbreitung oder Auswirkungen auf Bildungsqualität getroffen werden. Wohl aber sind differenzierte und theoretisch begründete Aussagen darüber getroffen, wie Lehrkräfte auf die Ergebnisse aus den Lernstandserhebungen argumentativ Bezug nehmen, um ihren Unterricht betreffende Steuerungsentscheidungen zu begründen. Diese Befunde differenzieren von Lehrerinnen und Lehrern vorgenommene kognitive Konstruktionen, die aus Lernstandserhebungen resultierendem Steuerungshandeln zugrunde liegen. Für folgende Forschungsprojekte schließen sich daran Fragen danach an, wie sich die unterschiedlichen Formen des Steuerungshandelns verteilen und in welchem Zusammenhang sie mit Variablen stehen, die Gegenstand bisheriger Studien zur Verwendung von Ergebnissen aus Lernstandserhebungen waren. Dazu zählen insbesondere allgemeine Einstellungen zur Evaluation und standardbasierter Testverfahren sowie schulspezifische Merkmale, wie etwa die Ausprägung kooperativer Beziehungen. Mit dieser Ausrichtung lässt sich die Verwendung der Ergebnisse aus Lernstandserhebungen unter Bezugnahme auf den Kernbereich professioneller Tätigkeit untersuchen, in dem die kognitive Konstruktion und Begründung von Entscheidungen unter der Bedingung hoher Ambiguität von Entscheidungen ausschlaggebend ist.

Für den praktischen Einsatz sind Erkenntnisse über die Art der Verwendung von Lernstandsergebnissen unter der Maßgabe unterschiedlicher Steuerungslogiken von Bedeutung, weil sie erstens zu einer Evaluation zentraler Leistungsmessung und ihrer Ergebnisrückmeldungen sowie zweitens zu einer Optimierung ihrer Gestaltung beitragen können. Nur wenn bekannt ist, wie die Ergebnisse von Lernstandserhebungen auf ihrem Weg über die Verarbeitung durch die Lehrerinnen und Lehrer und deren Entscheidungen Einfluss auf die Gestaltung von Schule und Unterricht nehmen, lässt sich ihre

Wirkung im Sinne eines Reforminstrumentes nachweisen, das Bildungsqualität möglicherweise sichert, indem es Professionalität im Bildungssystem stärkt. Professionelles Handeln ist auf differenzierte, valide und praktisch anschlussfähige Informationen angewiesen. Kenntnisse über die Argumentations-, Begründungs- und Entscheidungslogiken von Professionellen sind Voraussetzungen, um die verfügbaren Informationsquellen dem Informationsbedarf anpassen zu können. Dabei ist es notwendig, dass die Verwendungsformen, die Lernstandserhebungen praktisch finden, in ihrem Facettenreichtum unterschiedlicher Steuerungsmodelle reflektiert werden, anstatt sie einseitig der Erwartung einer outputorientierten Steuerung auszusetzen.

Literatur

- Bellenberg, G. (2008). Zur Nutzung von zentralen Abschlussprüfungen als Bausteine eines umfassenden Qualitätssicherungs- und -entwicklungskonzepts. Ein Baustellenbericht. In W. Böttcher, W. Bos, H. Döbert & H. G. Holtappels (Hrsg.), *Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive* (S. 223-233). Münster: Waxmann Verlag.
- Bellmann, J. (2006). Bildungsforschung und Bildungspolitik im Zeitalter „Neuer Steuerung“. *Zeitschrift für Pädagogik*, 52, 487-504.
- Bonsen, M., Büchter, A., & Peek, R. (2006). Datengestützte Schul- und Unterrichtsentwicklung. Bewertungen der Lernstandserhebungen in NRW durch Lehrerinnen und Lehrer. *Jahrbuch der Schulentwicklung*, 14, 125-148.
- Bonsen, M., & von der Gathen, J. (2004). Schulentwicklung und Testdaten. Die innerschulische Verarbeitung von Leistungsrückmeldungen. *Jahrbuch der Schulentwicklung*, 13, 225-252.
- Böttcher, W. (2002). *Kann eine ökonomische Schule auch eine pädagogische sein? Schulentwicklung zwischen Neuer Steuerung, Organisation, Leistungsevaluation und Bildung*. Weinheim/München: Juventa Verlag.
- Böttcher, W. (2004). Neuere Entwicklungen der Bildungsreform und ihr Potenzial für den Lehrerberuf und die Lehrerbildung. In U. Beckmann, H. Brandt & H. Wagner (Hrsg.), *Ein neues Bild vom Lehrerberuf* (S. 72-86). Weinheim/Basel: Beltz Verlag.
- Böttcher, W. (2006). „Standard-based reform“ oder: Kann man für die Schulreform von den USA lernen. In F. Eder (Hrsg.), *Qualität durch Standards* (S. 71-84). Münster: Waxmann Verlag.
- Diemer, T., Rucht, S., Schulze, F., & Kuper, H. (2009). *Zur innerschulischen Nutzung zentraler Lernstandserhebungen. Deskriptive kategoriale Grundausswertungen problemzentrierter Interviews mit Lehrer/innen und Schulleitungen in vier Sekundarschulen in Berlin und Thüringen*. Berlin: Freie Universität.
<http://www.ewi-psy.fu-berlin.de> [1.4.2010].
- Dubs, R. (2004). *Qualitätsmanagement in der Schule*. Soest: Landesinstitut für Schule.
- EMSE – Netzwerk Empiriegestützte Schulentwicklung (2006). *Zentrale standardisierte Lernstandserhebungen*. Positionspapier. Berlin: IQB.
http://www.iqb.hu-berlin.de/bista/EMSE/EMSE_Positionsp1_Lernstandserh.pdf [1.4.2010].
- Fend, H. (2001). *Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung* (2. Aufl.). Weinheim/München: Juventa Verlag.
- Glaser, B. G., & Strauss, L. (1967). *The discovery of grounded theory*. New York: Aldine de Gruyter.
- Groß Ophoff, J., Hosenfeld, I., & Koch, U. (2007). Formen der Ergebnisrezeption und damit verbundene Schul- und Unterrichtsentwicklung. *Empirische Pädagogik*, 21(4), 411-427.

- Groß Ophoff, J., Koch, U., & Hosenfeld, I. (2006). Ergebnisrückmeldungen und ihre Rezeption im Projekt VERA. In H. Kuper & J. Schneewind (2006) (Hrsg.), *Rückmeldung und Rezeption von Forschungsergebnissen. Zur Verwendung wissenschaftlichen Wissens im Bildungssystem* (S. 19-40). Münster: Waxmann Verlag.
- Hartung-Beck, V. (2009). *Schulische Organisationsentwicklung und Professionalisierung. Folgen von Lernstandserhebungen an Gesamtschulen*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Helmke, A. (2004). Von der Evaluation zur Innovation: Pädagogische Nutzbarmachung von Vergleichsarbeiten in der Grundschule. *Das Seminar*, 2, 90-112.
- Helmke, A., & Jäger, R. S. (2001). Die Evaluationsstudie MARKUS. *Empirische Pädagogik*, 15(4), 507-512.
- Helmke, A., & Schrader, F.-W. (2001). Von der Leistungsevaluation zur Unterrichtsentwicklung. In R. K. Silbereisen (Hrsg.), *Psychologie 2000. Bericht über den 42. Kongress der Deutschen Gesellschaft für Psychologie in Jena* (S. 594-606). Lengerich: Pabst Verlag.
- Hopf, C. (Hrsg.) (1995). *Qualitative Sozialforschung*. Stuttgart: Klett-Cotta Verlag.
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.) (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: LinkLuchterhand in Wolters Kluwer.
- Koch, U., Groß Ophoff, J., & Hosenfeld, I. (2006). Qualitätssicherung: Von der Evaluation zur Schul- und Unterrichtsentwicklung – Ergebnisse der Lehrerbefragungen zur Auseinandersetzung mit den VERA-Rückmeldungen. In F. Eder (Hrsg.), *Qualität durch Standards?* (S. 187-199). Münster: Waxmann Verlag.
- Kühle, B., & Peek, R. (2007). Lernstandserhebungen in Nordrhein-Westfalen. Evaluationsbefunde zur Rezeption und zum Umgang mit Ergebnisrückmeldungen in Schulen. *Empirische Pädagogik*, 21(4), 428-447.
- Kuper, H., & Hartung, V. (2007). Überzeugungen zur Verwendung des Wissens aus Lernstandserhebungen. Eine professionstheoretische Analyse. *Zeitschrift für Erziehungswissenschaft*, 10(2), 214-229.
- Kuper, H., & Schneewind, J. (Hrsg.) (2006). *Rückmeldung und Rezeption von Forschungsergebnissen. Zur Verwendung wissenschaftlichen Wissens im Bildungssystem*. Münster: Waxmann Verlag.
- Luhmann, N. (1994). *Soziale Systeme: Grundriss einer allgemeinen Theorie*. Frankfurt a.M.: Suhrkamp Verlag.
- Luhmann, N. (2000). *Organisation und Entscheidung*. Opladen: Westdeutscher Verlag.
- Luhmann, N., & Schorr, K. E. (1979). Das Technologiedefizit der Erziehungswissenschaften und die Pädagogik. *Zeitschrift für Pädagogik*, 25, 345-365.
- Maier, U. (2008a). Rezeption und Nutzung von Vergleichsarbeiten aus der Perspektive von Lehrkräften. *Zeitschrift für Pädagogik*, 54, 95-117.
- Maier, U. (2008b). Vergleichsarbeiten im Vergleich. Akzeptanz und wahrgenommener Nutzen standardbasierter Leistungsmessungen in Baden-Württemberg und Thüringen. *Zeitschrift für Erziehungswissenschaft*, 11(3), 453-474.
- Maier, U., & Rauin, U. (2006). Vergleichsarbeiten – Hilfe zur Unterrichtsentwicklung? Zentrale Lernstandserhebungen aus Sicht baden-württembergischer Lehrkräfte. *Die Deutsche Schule*, 98(4), 403-421.
- Mayring, P. (1996). *Einführung in die qualitative Sozialforschung*. Weinheim: Beltz Psychologie-Verl.-Union.
- Merton, R. K., & Kendall, P. L. (1979). Das fokussierte Interview. In C. Hopf (Hrsg.), *Qualitative Sozialforschung* (S. 171-204). Stuttgart: Klett-Cotta Verlag.
- Mitchell, R. (2006). *The nature of assessment. A guide to standardized testing*. Alexandria, VI: Center for Public Education.
<http://www.centerforpubliceducation.org/site/c.lvIXIiN0JwE/b.5056891/apps/s/content.asp?ct=6857851> [1.4.2010].

- Moser, U. (2003). *Klassencockpit im Kanton Zürich. Ergebnisse einer Befragung von Lehrerinnen und Lehrern der 6. Klassen über ihre Erfahrungen im Rahmen der Erprobung von Klassencockpit im Schuljahr 2002-2003*. Zürich: KBL.
- Nachtigall, C., & Jantowski, A. (2007). Die Thüringer Kompetenztests unter besonderer Berücksichtigung der Evaluationsergebnisse zum Rezeptionsverhalten. *Empirische Pädagogik*, 21(4), 401-410.
- Nachtigall, C., & Kröhne, U. (2006). Methodische Anforderungen an schulische Leistungsmessung – auf dem Weg zu fairen Vergleichen. In H. Kuper & J. Schneewind (Hrsg.), *Rückmeldung und Rezeption von Forschungsergebnissen* (S. 59-74). Münster: Waxmann Verlag.
- Orton, D. J., & Weick, K. E. (1990). Loosely Coupled Systems: A Reconceptualization. *The Academy of Management Review*, 15(2), 203-223.
- Schneewind, J. (2007). *Wie Lehrkräfte mit Ergebnisrückmeldungen aus Schulleistungsstudien umgehen. Ergebnisse aus Befragungen von Berliner Grundschullehrerinnen* (Unveröffentlichte Dissertation). FU Berlin.
- Schrader, F.-W., & Helmke, A. (2004). Von der Evaluation zur Innovation? Die Rezeptionsstudie WALZER: Ergebnisse der Lehrerbefragung. *Empirische Pädagogik*, 18(1), 140-161.
- Stanat, P. (2008). Entstehung und Umsetzung von Innovationen im Bildungssystem als Konsequenz aus Bildungsmonitoring, Bildungsberichterstattung und vergleichenden Schulleistungsstudien – Möglichkeiten und Grenzen. In Landesinstitut f. Schule u. Medien Berlin-Brandenburg (Hrsg.), *Bildungsmonitoring, Vergleichsstudien und Innovationen. Von evidenzbasierter Steuerung zur Praxis* (S. 11-24). Berlin: Berliner Wissenschafts-Verlag.
- Tenorth, H.-E. (2006). Professionalität im Lehrerberuf. Ratlosigkeit der Theorie, gelingende Praxis. *Zeitschrift für Erziehungswissenschaft*, 9(4), 580-597.
- Tresch, S. (2007). *Potenzial Leistungstest. Wie Lehrerinnen und Lehrer Ergebnisrückmeldungen zur Sicherung und Steigerung ihrer Unterrichtsqualität nutzen*. Bern: hep Verlag.
- Weick, K. E. (1976). Educational Organizations as Loosely Coupled Systems. *Administrative Science Quarterly*, 21(1), 1-19.

Abstract: Central surveys on student achievement are part of a new paradigm of controlling labeled “new control”. Unlike other instruments associated with this paradigm, such as international comparisons on school achievement or national education reports, central surveys on student achievement bring the new paradigm directly to the schools and teachers. On the basis of a qualitative interview study, the authors examine in how far control processes take place within schools in connection with the feedback from achievement surveys. The results reveal considerable variation in the logics of controlling which are not reduced to the paradigmatic expectation of output-oriented control.

Anschrift der Autoren

Tobias Diemer, Freie Universität Berlin, Fachbereich Erziehungswissenschaft und Psychologie, Arbeitsbereich Weiterbildung und Bildungsmanagement, DFG-Forschungsprojekt „Auswirkungen zentraler Lernstandserhebungen“, Arnimallee 12, 14195 Berlin, Deutschland
E-Mail: tobias.diemer@fu-berlin.de

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