

Franziska Schulze

Folgen zentraler Lernstandserhebungen: Unterscheiden sich Lehrkräfte in der Auseinandersetzung mit Lernstandserhebungen? Empirische Befunde zu Formen unterrichtsbezogener Steuerungslogik

Zusammenfassung

Mittels zentraler Lernstandserhebungen erhalten Schulleitungen und Lehrkräfte Informationen über die Leistungen ihrer Schülerinnen und Schüler. Auf der Grundlage einer Erhebung mit Thüringer Lehrkräften (n=43) wurde der bisher ungeklärten Frage nachgegangen, ob sich individuelle Unterschiede in der Auseinandersetzung mit den Ergebnissen des Thüringer Kompetenztests zeigen. Im Ergebnis können verschiedene Formen der unterrichtsbezogenen Steuerungslogik identifiziert und Unterschiede hinsichtlich der Auseinandersetzung von Lehrkräften mit unterschiedlichen Unterrichtsfächern aufgezeigt werden.

Schlüsselwörter: Neue Steuerung, Lernstandserhebungen, Ergebnismeldung, Lehrkräfte

Consequences of Standardized Performance Tests: Do Teachers Differ in Their Interpretation of Standardized Performance Tests?

Empirical Findings to Types of a Teaching-related 'Control Logic'

Abstract

As a result of standardized performance tests, school administration and teachers receive information concerning student achievement. This study examines how the feedback information, resulting from the Thüringer Kompetenztest, is interpreted and handled by teachers, based on a survey with Thuringian teachers (n=43). Different types of 'control logic' in schools are identified. Additionally, differences in the interpretation of evaluation results can be found between teachers of different subjects.

Keywords: new governance, performance tests, evaluation feedback, teachers

1. Neue Steuerung: Lernstandserhebungen als Ausgangspunkt für die Schul- und Unterrichtsentwicklung

Die Schulsysteme der Länder der Bundesrepublik Deutschland erfahren seit einigen Jahren eine Reihe tiefgreifender Veränderungen. Intendiert ist dabei im Allgemeinen eine Umorientierung von einer input- hin zu einer outputorientierten Form der Steuerung von schulischen Arbeits- und Entwicklungsprozessen (vgl. Oelkers/Reusser 2008). Mit dieser Strategie, auch *Neue Steuerung* genannt, wird vielfach der Anspruch verbunden, Schulleitungen und Lehrkräften ein höheres Maß an Verantwortung für die erfolgreiche Gestaltung der Schul- und Unterrichtsorganisation zu übertragen. Neben der Formulierung bundesweiter Bildungsstandards stellt die Einführung von Lernstandserhebungen¹ ein wesentliches Element der Umorientierung dar. „Lernstandserhebungen beziehen sich auf umfassendere Lern- bzw. Kompetenzentwicklungen (bisher erreichte Fertigkeiten, Fähigkeiten, Kenntnisse) und nicht wie Klassenarbeiten auf das unmittelbar vorangegangene Unterrichtsgeschehen“ (Peek/Dobbelstein 2006, S. 41). Durch die Rückmeldung der Ergebnisse aus diesen Erhebungen erhalten Lehrkräfte und Schulleitungen wissenschaftlich erhobene Informationen über den fachbezogenen Lernstand ihrer Schülerinnen und Schüler sowie indirekt damit auch über das Ergebnis ihrer Arbeit. Diese fachleistungsbezogenen Informationen über die Wirksamkeit des Geleisteten sollen den schulischen Akteurinnen und Akteuren bei der Planung und Gestaltung der Qualitätsentwicklungsarbeit auf der Ebene des Unterrichts helfen (vgl. Altrichter 2010; Netzwerk Empiriegestützte Schulentwicklung 2006). In den letzten Jahren wurde eine Vielzahl von Lernstandserhebungen durchgeführt; mittlerweile gibt es in allen Bundesländern Lernstandserhebungen, die zum Teil vom Institut zur Qualitätsentwicklung im Bildungswesen zentral entwickelt und dezentral in den Ländern eingesetzt und deren Ergebnisse zurückgemeldet werden (**Vergleichsarbeiten** in den Fächern Deutsch und Mathematik in der dritten Klasse der Grundschule [VERA 3], **Vergleichsarbeiten** in den Fächern Deutsch, Englisch und Mathematik in der sechsten Klasse [VERA 6] und **Vergleichsarbeiten** in den Fächern Deutsch, erste Fremdsprache und Mathematik in der achten Klasse [VERA 8]). Darüber hinaus gibt es zusätzlich landesspezifische Verfahren, zu denen z.B. **MARKUS (Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterricht, Schulkontext**; vgl. Schrader/Helmke 2004), **Lernstand 9** bzw. **Lernstand 8** in Nordrhein-Westfalen (vgl. Kühle/Peek 2007) und die **Kompetenztests** in Thüringen (vgl. Nachtigall/Jantowski 2007) zählen.

1 Der Begriff der Lernstandserhebungen wird im weiteren Textverlauf als Synonym für ebenfalls gebräuchliche Begriffe wie Schulleistungstests, Vergleichsarbeiten, Kompetenztests, Gesamterhebungen oder Qualitätsuntersuchungen verwendet. Gemeint sind regelmäßig durchgeführte bundeslandweite Erhebungen zu Lern- bzw. Kompetenzentwicklungsständen der Schülerinnen und Schüler in bestimmten Kernfächern (meist Deutsch, Mathematik und Englisch).

1.1 Nutzung der Ergebnisrückmeldungen aus Lernstandserhebungen

Es ist davon auszugehen, dass die Nutzung von Ergebnisrückmeldungen durch verschiedene Aspekte bedingt wird. Groß Ophoff u.a. (2006) unterscheiden als wesentliche Aspekte erstens die Form und den Inhalt der dargestellten Informationen, zweitens die individuellen Bedingungen der Rezipierenden und drittens schulische und externe Bedingungen (vgl. ebd., S. 32). Die Daten, welche den Adressatinnen und Adressaten der Rückmeldung zur Verfügung gestellt werden, können nach dem Aggregationsniveau in Informationen über das Bundesland, die Schule, die Klasse und über den Lernstand jedes einzelnen Schülers und jeder einzelnen Schülerin unterschieden werden (vgl. Schneewind/Kuper 2009). Außerdem spielt nach Altrichter (2010) die Auswahl der Bezugsnorm (soziale, ipsative/individuelle und sachliche/kriteriale) für die Rückmeldung der Ergebnisse eine Rolle. Dieser erste Aspekt der Informationsqualität ist insofern von Bedeutung, als dass er die folgende Nutzung gewissermaßen präkonfiguriert. Er ist nach Rolff (2002) jedoch nicht hinreichend dafür, dass tatsächlich Veränderungen in Schulen vollzogen werden. Auch wenn davon ausgegangen werden kann, dass die Daten rezipiert werden, stellt sich die Frage, in welcher Weise die Daten in Abhängigkeit der Aspekte individueller und schulischer Bedingungen genutzt werden.

1.2 Modelle zur Ergebnisnutzung

Helmke (2004) hat ein Rahmenmodell vorgelegt, das diesen möglichen Nutzungsprozess in Abhängigkeit von individuellen und schulischen Bedingungen in Form von vier aufeinander bezogenen Phasen beschreibt. Die (1.) *Rezeption* von leistungsbezogenen Daten aus den Lernstandserhebungen beinhaltet sowohl die erfolgreiche Übermittlung und Wahrnehmung als auch „die kognitive Verarbeitung des semantischen Gehalts der zurückgemeldeten Informationen (Ergebnisse)“ (Hartung-Beck/Diemer 2009, S. 216). Dabei sollten die zurückgemeldeten Informationen so aufbereitet und kommentiert sein, dass Fehlinterpretationen und Missverständnisse seitens der Lehrkräfte auf ein Minimum reduziert werden. Bei der (2.) *Reflektion* wird nach den Gründen und Ursachen für gute oder schlechte bzw. erwartete oder erwartungswidrige Ergebnisinformationen gesucht. Dies ist der eigentliche Schwerpunkt für pädagogische Impulse, denn je „intensiver die [...] Ergebnisse innerschulisch für daran anknüpfende Diskussionen genutzt werden [...], desto höher sind die Chancen, dass daraus fundierte und realistische Maßnahmen abgeleitet werden“ (Helmke 2004, S. 102). Die Rezeption und Reflektion kann zu verschiedenen (3.) *Aktionen* in den Bereichen der Schul- und Unterrichtsentwicklung führen. Die Wirkung der eingeleiteten Handlungen und Aktionen wird dem Modell zufolge schließlich durch eine (4.) *Evaluation* überprüft. Die tatsächliche Ausprägung dieser vier Schritte hängt nach dem Modell Helmkes einerseits von individuellen und andererseits von schulischen und externen Bedingungen ab. Bensen und von der Gathen (2004) haben

ein alternatives Rahmenmodell zur Rückmeldungsnutzung entwickelt, welches die Einzelschule fokussiert. Dabei werden neben verschiedenen Arten des Nutzens (instrumenteller, konzeptioneller, symbolischer und strategischer Nutzen) sowie beabsichtigten und unbeabsichtigten Wirkungen von ErgebnISRückmeldungen auch unterschiedliche Faktoren dargestellt, welche die Art und Intensität der Nutzung von ErgebnISRückmeldungen aus Lernstandserhebungen bedingen. „Das ‚Was‘ als Input eines solchen Systems, also Anlage und Inhalt einer Daten-Rückmeldung, beeinflusst in diesem Modell das ‚Wie‘ auf der Prozessebene“ (ebd., S. 247). Beide Modelle stimmen in der Annahme überein, dass die Nutzung von Ergebnissen aus Lernstandserhebungen sowohl von der Gestaltung der Rückmeldungen, von der Reflexionsbereitschaft der Lehrkräfte, vom Schul- und Kooperationsklima als auch von externen Unterstützungssystemen abhängt.

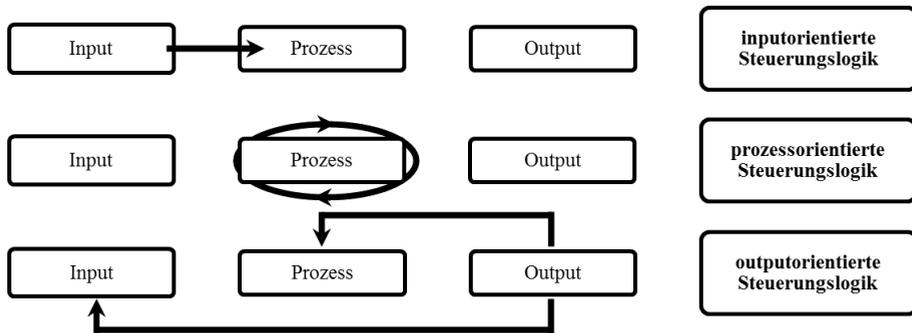
1.3 Rezeptionsstudien

Um die Nutzung von ErgebnISRückmeldungen in Schulen zu untersuchen und um zu prüfen, ob die Ergebnisse aus den Rückmeldungen in den Schulen zu Veränderungen führen, sind zwischenzeitlich bereits verschiedene Rezeptionsstudien durchgeführt worden. Bei diesen Studien steht zunächst die Frage im Vordergrund, welche Bereiche des unterrichtlichen Handelns durch Rückmeldungen beeinflusst werden und welche Kontextfaktoren die Akzeptanz und Nutzung von Lernstandserhebungen bedingen (vgl. Maier 2008). Die Ergebnisse der Rezeptionsstudien machen deutlich, dass die Nutzung von ErgebnISRückmeldungen sehr unterschiedlich verläuft. Während einige Lehrkräfte die Ergebnisse aus den Erhebungen für die Unterrichtsanalyse und -weiterentwicklung nutzen, werden die Ergebnisse in anderen Schulen kaum beachtet (vgl. Kühle/Peek 2007; Nachtigall/Jantowski 2007; Schrader/Helmke 2003). Zudem stellen die Studien heraus, dass die Nutzung von Rückmeldungen aus Lernstandserhebungen durch eine offene Haltung der Lehrkräfte gegenüber externen Evaluationen positiv beeinflusst wird (vgl. Kühle/Peek 2007; Schrader/Helmke 2004). Maier (2008) verweist in diesem Zusammenhang auf den Aspekt der Aufgeschlossenheit der Thüringer Lehrkräfte gegenüber zentralen Tests und der Auseinandersetzung mit den daraus resultierenden Ergebnissen: „Die Kompetenztests in Thüringen werden eher akzeptiert, weniger als Belastung angesehen und haben für die Lehrkräfte einen höheren lerndiagnostischen Hinweischarakter“ (S. 470). In einigen Studien wurden unterschiedliche Rezeptions- und Nutzungstypen beschrieben (vgl. Groß Ophoff/Hosenfeld/Koch 2007; Hartung-Beck 2009; Rossi/Freeman 1993; Stamm 2003). Groß Ophoff, Hosenfeld und Koch (2007) identifizierten beispielsweise aufgrund der Daten aus der Befragung zu VERA 2004 drei Typen von Lehrkräften, die sich unterschiedlich intensiv mit den Daten auseinandersetzten. Der Rezeptionstyp 1 zeichnet sich durch eine intensive positive Auseinandersetzung aus und schätzt die Verständlichkeit und Nützlichkeit der ErgebnISRückmeldungen als hoch ein. Der Rezeptionstyp 2 hin-

gegen ist gekennzeichnet durch eine weniger intensive Auseinandersetzung mit den Rückmeldungen. Außerdem wird hier die Verständlichkeit niedriger als bei Typ 1 eingeschätzt. Der Rezeptionstyp 3 setzt sich intensiv, aber kritisch mit den Ergebnisrückmeldungen auseinander und schätzt die Nützlichkeit als niedrig ein. Diese Rezeptionsprofile unterschieden sich auch signifikant hinsichtlich der durchgeführten Aktionen zur Unterrichtsentwicklung: Der Rezeptionstyp 1 zeichnet sich durch eine aktive Maßnahmenableitung aus, wohingegen der Typ 3 seltener Maßnahmen zur Unterrichtsentwicklung einführt. Der Rezeptionstyp 2 lässt sich hinsichtlich der Maßnahmenableitung zwischen den anderen beiden Typen einordnen.

Erkenntnisse über unterschiedliche Formen der Auseinandersetzung der Lehrkräfte mit den Lernstandserhebungen im Sinne unterschiedlicher auf den Unterricht bezogener Steuerungslogiken liegen bislang in den Studien kaum vor. Dederling (2011) verweist in ihrer systematischen Analyse zur innerschulischen Auseinandersetzung mit Ergebnisrückmeldungen neben unterschiedlichen Rezeptions- und Nutzungstypen auch auf „unterschiedliche Motive der Datenverarbeitung“ (S. 77), welche sie jedoch nicht weiter ausführt. Im Rahmen der Interviewstudie *Folgen zentraler Lernstandserhebungen* (vgl. Diemer/Kuper 2011; Diemer u.a. 2009) wurde erhoben und untersucht, welche Arten von Aktivitäten der Rezeption, der Reflektion, der Maßnahmenableitung und der Evaluation in Schulen im Zusammenhang mit der Nutzung zentraler Lernstandserhebungen realisiert werden. Angelehnt an das Nutzungsmodell nach Helmke (2004) werden diese Tätigkeiten auch in dieser Studie als Phasen dargestellt. Die Ergebnisse aus dem Interviewmaterial zeigen jedoch, dass die Phasen *Rezeption*, *Reflektion*, *Aktion* und *Evaluation* zum Teil unabhängig voneinander ablaufen. So leiten beispielsweise Lehrkräfte Maßnahmen ab, die einerseits zu ihrer Rezeption passen, andererseits jedoch nicht zu den reflektierten Gründen für das Zustandekommen der Ergebnisse. Auch wenn innerhalb der Phase *Reflektion* die Gründe bezüglich des Zustandekommens der Ergebnisse input- und prozessorientiert identifiziert werden, erfolgt die Maßnahmenableitung tendenziell prozessorientiert und ohne Reflexionsbezug. Diese Diskrepanz deutet darauf hin, dass die Ergebnisse aus Lernstandserhebungen nicht ausschließlich als Informationen über den Output wahrgenommen werden: „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“ (Diemer/Kuper 2011, S. 557). Vor diesem Hintergrund lassen sich bezüglich einer steuerungstheoretischen Analyse der Nutzung zentraler Lernstandserhebungen neben einer outputorientierten auch input- und prozessorientierte Steuerungslogiken unterscheiden, welche von den Lehrkräften – in Auseinandersetzung mit den zur Verfügung gestellten Informationen und bezogen auf den Unterricht – realisiert werden (siehe Abbildung 1):

Abbildung 1: Modell zu den unterrichtsbezogenen Steuerungslogiken von Lernstandserhebungen²



Quelle: eigene Darstellung

Eine inputorientierte Steuerungslogik zeichnet sich dadurch aus, dass die Informationen aus Lernstandserhebungen als Lernausgangslage gesehen werden und als solche zu Anfang des Schuljahres in die Unterrichtsplanung eingehen. Lernstandserhebungen stellen hierbei die Grundlage für Lernzielerwartungen dar. Prozessorientierte Steuerungslogik meint, dass die Rückmeldungen als Teil des Unterrichtsprozesses wahrgenommen und die Informationen für die Maßnahmenableitungen weitgehend für die Gestaltung dieses Prozesses genutzt werden. Lernstandserhebungen geben also Orientierungen für die Weiterarbeit im Unterricht. Bei der outputorientierten Steuerungslogik werden die Ergebnisse als abschließende Bewertungen der Leistungen der Schülerinnen und Schüler (Informationen über den Output) mit reflektiertem Rückgriff auf die Eingangsgrößen und die Prozesse betrachtet. Lernstandserhebungen sind dabei ausschließlich Erhebungsinstrumente zur Leistungsüberprüfung. Altrichter (2009) betont in diesem Kontext das Wissen um die unterschiedliche Wahrnehmung der Lernstandserhebungen seitens der Lehrkräfte als Nutzungsvoraussetzung: „Das dahinter liegende Problem besteht darin, dass die *Passung der angebotenen Informationen (und ihre Lokalisierung im Arbeitsprozess)* auf Arbeits- und Denkstile der Lehrpersonen eine Voraussetzung für ihre Nutzung ist“ (S. 219).

2. Fragestellung

Die angeführten Ergebnisse der Interviewstudie *Folgen zentraler Lernstandserhebungen* verdeutlichen, dass es neben der angestrebten und zahlreich diskutierten outputorientierten Steuerungslogik der Rückmeldung von Ergebnissen aus Lernstandserhebungen auch andere Formen der Auseinandersetzung mit Lernstandserhebungen gibt, die

² Dieses Modell wurde im Rahmen der Auswertungen im DFG-Projekt *Folgen zentraler Lernstandserhebungen* (Leitung: Prof. Dr. H. Kuper) von T. Diemer und F. Schulze 2009 entwickelt und in der Qualifikationsarbeit von N. Richter und F. Schulze 2009 erstmals verwendet.

von den Lehrkräften realisiert werden. Bislang gibt es Forschungsarbeiten, in denen verschiedene gegenstandsnahe Nutzungstypologien entwickelt wurden. Dabei spielten die unterschiedlichen Steuerungslogiken kaum eine Rolle. Die vorliegende Untersuchung widmete sich der Frage, in welcher Art und Weise sich Lehrkräfte mit den Ergebnismeldungen aus den Thüringer Kompetenztests auseinandersetzen und welche Steuerungslogiken von den Lehrkräften hierbei im Unterricht realisiert werden. Es sollte überprüft werden, ob sich die in der Interviewstudie *Folgen zentraler Lernstandserhebungen* gefundenen Argumentationslinien auch in der Ergebnisrezeption und -nutzung von Lehrkräften und Schulleitungen zeigen lassen, die mithilfe eines quantitativen Erhebungsinstruments befragt wurden. Diese Untersuchung stellte eine erste Pilotierung eines Fragebogens dar, mit dem unterschiedliche Argumentationsweisen hinsichtlich der unterrichtsbezogenen Steuerungslogiken (input-, prozess- und outputorientiert) abgebildet werden. Ziel der Erhebung war es, weiterführende Informationen darüber zu generieren, wie die Ergebnismeldungen von Lernstandserhebungen in Schulen Verwendung finden.

3. Methode

3.1 Design

Ein wesentlicher Teil der Untersuchung bestand zunächst in der Entwicklung eines standardisierten Fragebogens, der die auf den Unterricht bezogene Steuerungslogik der Lehrkräfte untersucht und der erfasst, als welche Informationsart (input-, prozess- oder outputorientiert) die Lernstandserhebungen und ihre Rückmeldungen wahrgenommen werden. Die Studie diente dazu, im Sinne eines Pretests die Eignung des Untersuchungsinstruments und den reibungslosen Untersuchungsablauf zu testen (vgl. Bortz/Döring 2006). Darüber hinaus konnten gleichzeitig Informationen über die Auseinandersetzung der Lehrkräfte mit Ergebnismeldungen aus zentralen Lernstandserhebungen erfasst und beschrieben werden.

3.2 Stichprobe

Die Erhebung wurde in Thüringen im Schuljahr 2008/2009 durchgeführt. Dieses Bundesland ist aufgrund der dort gegebenen langjährigen Erfahrungen mit Lernstandserhebungen ausgewählt worden. Die *Kompetenztests* werden seit dem Schuljahr 2002/2003 in der dritten Klasse in Deutsch und Mathematik, in der sechsten Klasse in Deutsch, Mathematik und Englisch und in der achten Klasse in Mathematik geschrieben. Diese Tests werden als Diagnoseinstrument eingesetzt, ohne dass eine Verpflichtung zur Notenvergabe besteht (vgl. Diemer/Kuper 2011).

An der Untersuchung nahmen 43 Lehrkräfte teil;³ davon lehrten 13 an Regelschulen, acht an Gesamtschulen und 21 an Gymnasien. In der vorliegenden Stichprobe sind die Lehrkräfte von Gymnasien und Gesamtschulen gemessen an der Gesamtverteilung in Thüringen überrepräsentiert, die von Regelschulen dagegen unterrepräsentiert (vgl. Thüringer Ministerium für Bildung, Wissenschaft und Kultur/Statistikstelle 2010). Das mittlere Alter der Lehrkräfte in der Studie⁴ (MW 49,9 Jahre bei einer Altersspanne zwischen 38 und 61 Jahren) entspricht dem durchschnittlichen Alter der Lehrkräfte in den genannten Thüringer Schultypen von 50,1 Jahren (vgl. ebd.). Von den 43 Lehrkräften sind in dieser Stichprobe 34 Lehrkräfte weiblich und sieben Lehrkräfte männlich.⁵ In der vorliegenden Stichprobe ist der Anteil an weiblichen Lehrkräften etwas höher, spiegelt jedoch die Verteilung der Geschlechter in den Schulen tendenziell wider (vgl. ebd.). Die Unterrichtsfächer der befragten Lehrkräfte sind heterogen.

3.3 Erhebungsinstrument

Der Fragebogen umfasst eine Skala zu den von den Lehrkräften realisierten Steuerungslogiken der Lernstandserhebungen mit drei Subskalen, deren Items mit einer vierstufigen Likert-Skala (1 = stimme voll zu, 2 = stimme eher zu, 3 = stimme eher nicht zu, 4 = stimme gar nicht zu) gemessen wurden. Die unterrichtsbezogene Steuerungslogik kann als Konstrukt bezeichnet werden, das unterschiedliche Orientierungen der Wahrnehmung von Lernstandserhebungen aufweist (input-, prozess- und outputorientierte Steuerungslogik). Die Items beziehen sich an dieser Stelle auf die Wahrnehmung der Lernstandserhebungen und nicht auf die direkte Maßnahmenableitung. Die Formulierung der Items basierte auf dem kategorial ausgewerteten Datenmaterial des Forschungsprojekts *Folgen zentraler Lernstandserhebungen* (vgl. Diemer u.a. 2009). Somit kann von einer inhaltlichen Validität der Items ausgegangen werden. Jedoch zeigt sich aufgrund der mittleren Korrelationen der Subskalen untereinander eine eher geringe diskriminante Validität.

Die in Tabelle 1 aufgeführten Reliabilitäten der Subskalen variieren zwischen $\alpha = .89$ und $\alpha = .92$. Zusammengefasst zeigen die Ergebnisse der Reliabilitätsanalyse gute bis sehr gute Werte für die interne Konsistenz und damit den Grad der Zuverlässigkeit der einzelnen Subskalen. Die ihnen zugrunde liegenden einzelnen Items weisen des Weiteren eine hohe Eigentrennschärfe auf ($r_{ii} > .60$). Sie können demzufolge als prototypisch für die jeweiligen Skalen gelten. Die Überprüfung der Binnenstruktur der Subskalen zu den Formen der Steuerungslogik mithilfe einer Korrelationsanalyse

3 Zur Erfassung der Nutzung von Rückmeldeergebnissen wurden insgesamt 126 Fragebögen verschickt. Davon kamen 43 Fragebögen zurück; die Rücklaufquote beträgt demzufolge 34 Prozent.

4 Acht Lehrkräfte haben dabei keine Angaben zu ihrem Alter gemacht.

5 Zwei Personen haben diese Frage nicht beantwortet.

Tabelle 1: Skalendokumentation zu den unterrichtsbezogenen Steuerungslogiken

Subskalen	Anzahl der Items ¹	Interne Konsistenz der Subskalen (Cronbachs Alpha)	N	Beispielitem
Inputorientierte Steuerungslogik	8	.89	39	„Lernstandserhebungen stellen mir Informationen über die Lernausgangslage in meiner Klasse bereit.“
Prozessorientierte Steuerungslogik	6	.92	42	„Lernstandserhebungen geben mir Rückmeldungen über meinen Unterrichtsprozess.“
Outputorientierte Steuerungslogik	7	.92	38	„Lernstandserhebungen dienen der Zertifizierung von Schülerleistungen.“

¹ Aufgrund mangelnder Trennschärfe und inhaltlicher Differenzen wurden einige Items des Originalfragebogens ausgeschlossen.

Quelle: eigene Berechnungen

zeigt, dass die in den drei verwendeten Subskalen erfassten Steuerungslogiken in einem 2-seitig signifikanten interkorrelativen Zusammenhang stehen (in allen Fällen $p < .01$).

4. Ergebnisse

4.1 Formen der Steuerungslogik von Lernstandserhebungen im Vergleich

Insgesamt ist festzustellen, dass die Items der Subskalen zur input- und outputorientierten Steuerungslogik weder prinzipiell positiv noch negativ beantwortet wurden. Nur die Subskala zur prozessorientierten Steuerungslogik wurde durchweg positiv beantwortet. Es sind Tendenzen sichtbar, welche Aussagen über die Wahrnehmung der Informationen aus Lernstandserhebungen seitens der Lehrkräfte zulassen: Mithilfe der Varianzanalyse mit dem allgemeinen linearen Modell lässt sich in den Ergebnissen zu den Formen der Steuerungslogik zeigen, dass die Lehrkräfte sich mit den Ergebnissen aus Lernstandserhebungen unterschiedlich auseinandersetzen und dabei der prozessorientierten Steuerungslogik (MW 2,37, SD 0,63) tendenziell eher zustimmen.⁶ Auch die Tests der Innersubjekteffekte belegen das Vorhandensein von signifikanten Unterschieden zwischen den Subskalen ($p < .01$). Die Tests der Innersubjektkontraste zeigen, dass es zwischen den Subskalen signifikante Unterschiede auf .01 Niveau gibt: Die inputorientierte Steuerungslogik unterscheidet sich signifikant von der prozessorientierten Steuerungslogik, die prozessorientierte von der outputorientierten

⁶ Die Subskala zur inputorientierten Steuerungslogik der Lernstandserhebungen weist einen Mittelwert von 2,54 auf (SD 0,54); die Subskala zur outputorientierten Steuerungslogik von Informationen aus Ergebnismeldungen hat einen Mittelwert von 2,61 (SD 0,59).

Steuerungslogik. Die Berechnungen im allgemeinen linearen Modell ergaben, dass 21 Prozent der Varianz durch die Formen der Steuerungslogik auf der Stichprobenebene aufgeklärt wurden.⁷ Nach Bortz und Döring (2006) entspricht dies einer mittleren bis großen Effektgröße.

4.2 Gruppenspezifische Unterschiede in der Steuerungslogik von Lernstandserhebungen

Mithilfe des T-Tests bei unabhängigen Stichproben und der Varianzanalyse wurden Unterschiede in den Gruppen Geschlecht, Fächergruppe und Schulart (Gymnasium, Regelschule- und Gesamtschule) hinsichtlich der einzelnen Formen der unterrichtsbezogenen Steuerungslogik untersucht. In den Ergebnissen der Auseinandersetzung mit den Lernstandserhebungen seitens der Lehrkräfte zeigt sich, dass keine Unterschiede hinsichtlich der Steuerungslogik von männlichen und weiblichen Lehrkräften vorhanden sind. Auch gemessen an der Schulart gibt es keine Unterschiede. Dagegen lassen sich Unterschiede bezüglich der inputorientierten Steuerungslogik ($t_{33,88} = -2.54$, $p < .05$, $d = .84$) von Lehrkräften der naturwissenschaftlichen und der sprachlichen Fächer feststellen. Lehrkräfte mit naturwissenschaftlichen Fächern (MW 2,27, SD 0,45) favorisieren im Vergleich zu denen mit sprachlichen Fächern (MW 2,69, SD 0,54) eher die inputorientierte Steuerungslogik. Die Effektgröße von $d = .84$ ist nach Bortz und Döring (2002) als groß zu interpretieren. Lehrkräfte mit diesen Fächern unterscheiden sich auch hinsichtlich der prozessorientierten Steuerungslogik ($t_{33,41} = -2.13$, $p < .001$, $d = .71$). Hierbei bevorzugen Lehrkräfte mit naturwissenschaftlichen Unterrichtsfächern (MW 2,16, SD 0,54) eher die prozessorientierte Steuerungslogik.⁸ Diese Effektgröße ist als mittel einzuschätzen (vgl. Bortz/Döring 2002).

5. Zusammenfassung und Diskussion

Das Ziel dieser Untersuchung war es, die unterrichtsbezogene Steuerungslogik von Lehrkräften dahingehend zu analysieren, ob die Informationen aus Lernstandserhebungen input-, prozess- oder outputorientiert rezipiert werden. Die Einstellungen der Lehrkräfte gegenüber den Lernstandserhebungen werden durch die argumentativ unterschiedliche Auseinandersetzung mit den Erhebungen und ihrer Ergebnismeldung verdeutlicht. Hinsichtlich der Subskalen zu den Formen der Steuerungslogik konnten Unterschiede im Antwortverhalten festgestellt werden: Lehrkräfte tendieren weniger zu den Items der input- und outputorientierten Steuerungslogik und eher zu den prozessorientierten Items und bewerten diese po-

7 Die Voraussetzungen für die Anova im allgemeinen linearen Modell sind erfüllt, da die Sphärizität gegeben ist.

8 Lehrkräfte mit sprachlichen Unterrichtsfächern haben einen Mittelwert von 2,57 (SD 0,6).

sitiv. Die an der Untersuchung beteiligten Lehrkräfte sehen die Informationen aus den Lernstandserhebungen folglich eher als prozessorientierte Information, welche der Zustandsbeschreibung der aktuellen Unterrichtsprozesse dient und welche die Lehrkräfte für die Weiterarbeit im Unterricht nutzen können, und weniger als leistungszertifizierende Information. Auch die Tests der Innersubjekteffekte und der Innersubjektkontraste belegen diesen signifikanten Unterschied. Die gezeigten Unterschiede in der input- und prozessorientierten Steuerungslogik der Ergebnisse aus Lernstandserhebungen von Lehrkräften aus den naturwissenschaftlichen und sprachlichen Fächern weisen daraufhin, dass es Unterschiede in der Auseinandersetzung abhängig von den unterrichteten Fächern gibt. Zu erklären sind diese Unterschiede eventuell mit der erfahrenen naturwissenschaftlichen oder geisteswissenschaftlichen Prägung. Die Gemeinsamkeiten der Personen dieser Fächergruppen hinsichtlich der outputorientierten Steuerungslogik lassen sich durch die von Bildungsadministration und -forschung proklamierte Funktion von Lernstandserhebungen erklären. Diese Unterschiede in den Formen der Steuerungslogik lassen sich jedoch nicht hinsichtlich des Geschlechts und der Schulart zeigen, wobei dies auch auf die unterschiedliche Anzahl der Gruppen in dieser Stichprobe zurückzuführen ist. Die Studie verweist so insgesamt auf Variationen in der auf den Unterricht bezogenen Steuerungslogik, die über die diskutierte outputorientierte Steuerung hinausgeht. Dabei stellt die prozessorientierte und nicht die outputorientierte Steuerungslogik die bei den befragten Lehrkräften verbreitetste realisierte Form der Auseinandersetzung mit den Ergebnismeldungen dar. Dies steht im Einklang mit den ersten qualitativen Ergebnissen der Interviewstudie *Folgen zentraler Lernstandserhebungen* (vgl. Diemer/Kuper 2011). Gründe für diesen Befund könnten zum einen darin bestehen, dass die Lernstandserhebungen in Thüringen nicht bewertet werden und somit kein innerschulisches Instrument der Leistungsüberprüfung darstellen, sondern eher ein Diagnoseinstrument sind. Zum anderen könnte die vergleichsweise geringe Ausprägung einer outputorientierten Steuerungslogik dadurch erklärt werden, dass das Prinzip der Outputorientierung auf Grund der Neuheit der Innovation bislang erst ansatzweise implementiert ist. Inwiefern die unterschiedliche Wahrnehmung der Lernstandserhebungen auch Auswirkungen auf unterschiedliche Maßnahmenplanung und tatsächlich umgesetzte Maßnahmen hat, wurde nicht untersucht, kann jedoch für weiterführende Forschungen interessant sein.

Des Weiteren zielte diese Untersuchung darauf ab, die Datenqualität des entwickelten Fragebogens zu überprüfen, um eine erneute Durchführung bei einer repräsentativen Stichprobe vorzubereiten. Die Subskalen zur input-, prozess- und outputorientierten Steuerungslogik der Ergebnismeldungen weisen eine hohe interne Konsistenz auf. Jedoch korrelieren die Subskalen untereinander stark, wodurch die angestrebte Differenzierung der unterschiedlichen Formen der Steuerungslogik nicht erreicht wurde. Die Gründe hierfür können in der starken Ähnlichkeit der Itemformulierungen liegen. So wurden Items eingesetzt, die zwar ähnlich klingen, jedoch

unterschiedliche Dimensionen der Steuerung und Nutzung abfragen. Voraussetzung für die korrekte Beantwortung dieser Items ist das differenzierte Lesen, Wahrnehmen und Beantworten dieser Fragen durch die Lehrkräfte. Davon kann jedoch nicht ausgegangen werden, sodass die einzelnen Nuancen in der Itemformulierung von den befragten Lehrkräften wahrscheinlich nicht erkannt wurden. Optimal wäre es, den Fragebogen mit den Lernstandserhebungen auszugeben, sodass Lehrkräfte während der Bearbeitungszeit der Schülerinnen und Schüler den Fragebogen beantworten können.

Als Desiderat bleibt die weiterführende Untersuchung der Fragestellung anhand einer repräsentativen Stichprobe. Ziel sollte dabei die Normierung des Tests und die Identifikation von unterrichtsbezogenen Steuerungslogiken sein. Denn nur wenn bekannt ist, wie sich Lehrkräfte mit den Ergebnissen aus Lernstandserhebungen argumentativ auseinandersetzen, kann auch Einfluss auf deren Nutzung genommen werden.

Literatur

- Altrichter, H. (2009): Datenfeedback und Unterrichtsentwicklung – Probleme eines Kernelements im „neuen Steuerungsmodell“ für das Schulwesen. In: Dicke, J.N./Ziegler, H./Böttcher, W. (Hrsg.): Evidenzbasierte Bildung. Wirkungsevaluation in Bildungspolitik und pädagogischer Praxis. Münster u.a.: Waxmann, S. 211-226.
- Altrichter, H. (2010): Schul- und Unterrichtsentwicklung durch Datenrückmeldung. In: Altrichter, H./Maag Merki, K. (Hrsg.): Handbuch neue Steuerung im Schulsystem. Wiesbaden: VS, S. 219-254.
- Bortz, J./Döring, N. (³2002): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Heidelberg: Springer.
- Bortz, J./Döring, N. (⁴2006): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Berlin: Springer.
- Bonsen, M./von der Gathen, J. (2004): Schulentwicklung und Testdaten. Die innerschulische Verarbeitung von Leistungsrückmeldungen. In: Holtappels, H.G./Klemm, K./Pfeiffer, H./Rolff H.-G./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung, Band 13. Weinheim u.a.: Juventa, S. 225-252.
- Dedering, K. (2011): Hat Feedback eine positive Wirkung? Zur Verarbeitung extern erhobener Leistungsdaten in Schulen. In: Unterrichtswissenschaft 39, H. 1, S. 63-83.
- Diemer, T./Kuper, H. (2011): Formen innerschulischer Steuerung mittels zentraler Lernstandserhebungen. In: Zeitschrift für Pädagogik 56, H. 4, S. 554-571.
- Diemer, T./Rucht, S./Schulze, F./Kuper, H. (2009): Zur innerschulischen Nutzung zentraler Lernstandserhebungen – Deskriptive kategoriale Grundauswertungen problemzentrierter Interviews mit Lehrer/innen und Schulleitungen in vier Sekundarschulen in Berlin und Thüringen. URL: http://www.ewi-psy.fu-berlin.de/einrichtungen/arbeitsbereiche/weiterbildung_bildungsmanagement/projekte/folgen_zentraler_lehr/index.html; Zugriffsdatum: 05.04.2011.
- Groß Ophoff, J./Hosenfeld, I./Koch, U. (2007): Formen der Ergebnisrezeption und damit verbundene Schul- und Unterrichtsentwicklung. In: Hosenfeld, I./Groß Ophoff, J. (Hrsg.): Nutzung und Nutzen von Evaluationsstudien in Schule und Unterricht. Landau: Verlag Empirische Pädagogik, S. 411-427.

- Groß Ophoff, J./Koch, U./Hosenfeld, I./Helmke, A. (2006): Ergebnismrückmeldungen und ihre Rezeption im Projekt VERA. In: Kuper, H./Schneewind, J. (Hrsg.): Rückmeldung und Rezeption von Forschungsergebnissen – Zur Verwendung wissenschaftlichen Wissens im Bildungssystem. Münster u.a.: Waxmann, S. 19-40.
- Hartung-Beck, V. (2009): Schulische Organisationsentwicklung und Professionalisierung. Folgen von Lernstandserhebungen an Gesamtschulen. Wiesbaden: VS.
- Hartung-Beck, V./Diemer, T. (2009): Sensemaking durch Outputorientierung. Erfahrungen mit der Nutzung von Lernstandserhebungen in Schulen. In: Göhlich, M./Weber, S.M./Wolff, S. (Hrsg.): Organisation und Erfahrung. Wiesbaden: VS, S. 215-225.
- Helmke, A. (2004): Von der Evaluation zur Innovation: Pädagogische Nutzbarmachung von Vergleichsarbeiten in der Grundschule. In: Das Seminar, H. 2, S. 90-112.
- Kühle, B./Peek, R. (2007): Lernstandserhebungen in Nordrhein-Westfalen. Evaluationsbefunde zur Rezeption und zum Umgang mit Ergebnismrückmeldungen in Schulen. In: Empirische Pädagogik 21, H. 4, S. 428-447.
- Maier, U. (2008): Vergleichsarbeiten im Vergleich – Akzeptanz und wahrgenommener Nutzen standardbasierter Leistungsmessungen in Baden-Württemberg und Thüringen. In: Zeitschrift für Erziehungswissenschaft 11, H. 3, S. 453-474.
- Nachtigall, C./Jantowski, A. (2007): Die Thüringer Kompetenztests unter besonderer Berücksichtigung der Evaluationsergebnisse zum Rezeptionsverhalten. In: Empirische Pädagogik 21, H. 4, S. 401-410.
- Netzwerk Empiriegestützte Schulentwicklung (2006): Zentrale standardisierte Lernstandserhebungen. Positionspapier zur 5. EMSE-Tagung, Berlin 08.12.2006. URL: http://www.isq-bb.de/uploads/media/emse-positionspapier-07-07_04.pdf; Zugriffsdatum: 01.02.2012.
- Oelkers, J./Reusser, K. (2008): Qualität entwickeln, Standards sichern – mit Differenz umgehen. Bonn: BMBF.
- Peek, R./Dobbelstein, P. (2006): Benchmarks als Input für die Schulentwicklung – Das Beispiel der Lernstandserhebungen in Nordrhein-Westfalen. In: Kuper, H./Schneewind, J. (Hrsg.): Rückmeldung und Rezeption von Forschungsergebnissen – Zur Verwendung wissenschaftlichen Wissens im Bildungssystem. Münster u.a.: Waxmann, S. 41-58.
- Richter, N./Schulze, F. (2009): Folgen zentraler Lernstandserhebungen – Entwicklung eines standardisierten Fragebogens zur Erfassung der Nutzung von Ergebnismrückmeldungen. Unveröffentlichte Qualifikationsarbeit. Freie Universität Berlin.
- Rolff, H.-G. (2002): Rückmeldung und Nutzung der Ergebnisse von großflächigen Leistungsuntersuchungen. In: Rolff, H.-G./Holtappels, H.G./Klemm, K./Pfeiffer, H./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung, Band 12. Weinheim u.a.: Juventa, S. 75-98.
- Rossi, P.H./Freeman, H.E. (1993): Evaluation: A Systematic Approach. Newbury Park, CA: Sage.
- Schneewind, J./Kuper, H. (2009): Rückmeldeformate und Verwendungsmöglichkeiten der Ergebnisse aus zentralen Lernstandserhebungen. In: Bohl, T./Kiper, H. (Hrsg.): Lernen aus Evaluationsergebnissen. Verbesserungen planen und implementieren. Bad Heilbrunn: Klinkhardt, S. 113-129.
- Schrader, F.-W./Helmke, A. (2003): Evaluation und was danach? Ergebnisse der Schulleiterbefragung im Rahmen der Rezeptionsstudie WALZER. In: Schweizerische Zeitschrift für Bildungswissenschaft 25, H. 1, S. 79-116.
- Schrader, F.-W./Helmke, A. (2004): Von der Evaluation zur Innovation? Die Rezeptionsstudie WALZER: Ergebnisse der Lehrerbefragung. In: Empirische Pädagogik 18, H. 1, S. 140-161.
- Stamm, M. (2003): Evaluation und ihre Folgen für die Bildung: Eine unterschätzte pädagogische Herausforderung. Münster u.a.: Waxmann.

| Franziska Schulze

Thüringer Ministerium für Bildung, Wissenschaft und Kultur/Statistikstelle (2010): Personal.
URL: <http://www.schulstatistik-thueringen.de>; Zugriffsdatum: 08.12.2010.

Franziska Schulze, M.A. Erziehungswissenschaft, geb. 1984, wissenschaftliche Mitarbeiterin am Institut für Erziehungswissenschaft, Abteilung Schulpädagogik/Schul- und Unterrichtsforschung der Westfälischen Wilhelms-Universität Münster.

Anschrift: Bispinghof 5/6, 48143 Münster
E-Mail: Franziska.Schulze@uni-muenster.de