
Die Wirksamkeit der neuen Steuerung

Theoretische und methodische Probleme ihrer Evaluation¹

Helmut Fend

Zusammenfassung

Im bildungspolitischen Diskurs der deutschsprachigen Länder über die Konsequenzen nach PISA wird eine neue, „evidenzbasierte Steuerung“ des Schulwesens vorgeschlagen, deren Kern in der Formulierung von Bildungsstandards und deren testbasierter Überprüfung sowie in einem neuen Wechselspiel von lokaler Autonomie und zentralen Kontrollformen besteht. Wer ein evidenzbasiertes Konzept verfolgt, der wird sich um eine empirische Untermauerung der Hypothese bemühen, dass mit der neuen Steuerung auch positive Wirkungen im Sinne verbesserter Lernergebnisse verbunden sind.

Im vorliegenden Beitrag wird deshalb in einem ersten Schritt versucht, diese Erwartungen, insbesondere jene an die Standards und ihre Messung als Instrumente der Steigerung von Lernergebnissen, in den Kontext der „school-effectiveness“-Forschung zu stellen. In einem zweiten Schritt wird geprüft, ob die neuen Konzepte der Steuerung, eingebettet in die politologischen Traditionen von „Governance“-Konzepten, neue Untersuchungschancen der Hypothese enthalten, dass durch Verfahren wie jene der Standardsicherung und Schulautonomie Verbesserungen im Bildungswesen zu erzielen sind. Vorausgeschickt wird eine kurze Klärung, was unter „neuer Steuerung“ heute verstanden wird.

Im konsequenzenorientierten Diskurs nach PISA hat eine als Paradigmawechsel apostrophierte neue Bildungspolitik eine große Rolle gespielt. Sie sollte einen

1 Formal angepasster Nachdruck von: Fend, H. (2011). Die Wirksamkeit der Neuen Steuerung – theoretische und methodische Probleme ihrer Evaluation. *Zeitschrift für Bildungsforschung* 1 (1), 5-24.

grundlegenden Wandel von einer Inputsteuerung zu einer Outputsteuerung vollziehen. Als Kern dieser „neuen Steuerung“ haben sich die Formulierung von Bildungsstandards und ihre testbasierte Überprüfung herauskristallisiert. Gleichzeitig wurden neue Formen der Regulierung des Schulwesens eingeführt. Im Wechselspiel von Autonomie und Kontrolle wurde vielerorts die klassische Form der Kontrolle durch Schulaufsicht ersetzt, durch eine Neugestaltung der Entscheidungsräume auf Schulebene im Sinne einer größeren Autonomie und Entscheidungskompetenz der Schulleitungen bei einer gleichzeitig stärker schulextern organisierten Evaluation (Heinrich 2007).

Diese neue Konfiguration der Steuerung, die an der Oberfläche *Restrukturierungen der Kontrolle des Bildungswesens* enthält und weniger *ressourcenintensive Unterstützungsleistungen*, stand im Kontext einer bildungspolitischen Gesamtstrategie, die diese stärker als bislang „evidence-based“ machen wollte, also solche Maßnahmen als implementationswürdig ansieht, die nachweislich positive Wirkungen haben. Sie sollte aus dem häufig von Glaubensüberzeugungen geprägten Streit über die notwendigen Reformen im Bildungswesen herausführen.

Wer konsequent ein evidenzbasiertes Konzept verfolgt, der wird sich auf eine empirische Untermauerung der Hypothese verlassen wollen, dass mit der neuen Steuerung auch positive Wirkungen im Sinne verbesserter Lernergebnisse verbunden sein werden. Wenn man dies tut, bewegt man sich unwillkürlich in das große Feld der „school effectiveness“-Forschung, das sich seit vielen Jahrzehnten mit den wirksamen Faktoren für die Verbesserung von Schulleistungen beschäftigt (Creemers, Scheerens und Reynolds 2000; Luyten 2003; Scheerens et al. 2006; Sørensen und Morgan 2000; Teddlie und Reynolds 2000; van der Wolf 1988).

Deshalb soll hier in einem ersten Schritt versucht werden, die obigen Erwartungen, insbesondere jene an die Standards und ihre Messung als Instrumente der Steigerung der Lernergebnisse, in den Kontext der „school-effectiveness“-Forschung zu stellen. In einem zweiten Schritt soll geprüft werden, ob die neuen Konzepte der Steuerung, eingebettet in politologische Governance-Traditionen (Benz 2007), neue Untersuchungschancen der Hypothese enthalten, dass durch Verfahren wie jene der Standardsicherung und Schulautonomie Verbesserungen im Bildungswesen zu erzielen sind. Vorausgeschickt sei eine kurze Klärung, was unter „neuer Steuerung“ heute verstanden wird (s. bspw. Altrichter, Brüsemeister und Wissinger 2007; Altrichter und Maag Merki 2010; Berkemeyer 2010).

1 „Neue Steuerung“ und die Überprüfung von „Systemsteuerungen“

Gibt es bereits eine „neue Steuerung“ in der Praxis und welches sind die *Kernelemente der Reformpraxis im Sinne einer neuen Governance*? Es erscheint trivial, wenn festgestellt wird, dass die Bildungsforschung darauf angewiesen ist, dass es diese Interventionen *gibt*, dass Verfahren der Standardbeschreibung und der Standardmessung entwickelt und damit soziale und kulturelle Wirklichkeit werden. Sie zu schaffen erfordert aber viele Ressourcen und je nach Anspruch an Qualität einen langen Zeitraum.

In der Tat sind in allen deutschsprachigen Ländern große Anstrengungen unternommen worden bzw. sind immer noch unterwegs, Verfahren der neuen Steuerung zu entwickeln und zu implementieren. In Österreich wird die Entwicklung von Standards und entsprechenden Testinstrumenten für die vierten und achten Klassen in den Fächern Deutsch, Mathematik und Englisch noch mehrere Jahre beanspruchen und einen zweistelligen Euromillionenbetrag benötigen. Ähnliches gilt für Instrumente zur objektiven Leistungsmessung auf Maturaebene, die testtheoretischen Kriterien der Item-response-Konzepte genügen sollen.

Deutschland und die Schweiz gehen hier etwas andere Wege. In Deutschland sind es sogenannte standardbasierte Vergleichsarbeiten, die zu verschiedenen Zeitpunkten der Schullaufbahn Auskunft geben sollen, was Schülerinnen und Schüler können. Im Hintergrund stehen Einigungen der Kultusministerkonferenz über sogenannte Standards, die erreicht werden sollen. Im Institut für Qualitätsentwicklung in Berlin werden die Testungen in verschiedenen Bundesländern gebündelt, formuliert und für Bundesländervergleiche aufbereitet. Auf Abiturebene sind mit der Ausnahme von Rheinland-Pfalz alle Bundesländer zum Zentralabitur übergegangen. Dabei werden Aufgaben von ausgewählten Expert/-innen der Lehrer/-innenschaft entwickelt, zentral zu gleichen Zeitpunkten gestellt und in einem ausgeklügelten Verfahren von Erstbegutachtung, Zweitbegutachtung durch Lehrpersonen aus anderen Schulen und bei abweichenden Voten durch Dritte bewertet.

In der Schweiz stehen Bemühungen im Vordergrund, die Standards in die Entwicklung neuer Lehrpläne zu integrieren, um Inhalte und Kompetenzen in ihren ergänzenden Funktionen sichtbar zu machen. Um Einblick in das mit anderen Schulklassen vergleichbare Niveau der Schulleistungen einer Schulklasse zu bekommen, können Lehrpersonen auf kostenpflichtige Tests, die über Internetplattformen (Cockpit, Stellwerk) erhältlich sind, zurückgreifen. Die Initiative dazu muss aber von den Lehrpersonen bzw. von der Schulpflege kommen, die schließlich die Kosten zu tragen hat.

An allen Orten werden große Investitionen in die neuen Verfahren getätigt. Dies macht den Wunsch verständlich, eine wissenschaftliche Erfahrungsbasis zur Wirksamkeit dieser Maßnahmen zu schaffen. Die Ausschreibung eines Forschungsschwerpunktes „Systemsteuerung“ durch das BMBF in Deutschland stellt eine erste Anstrengung in diese Richtung dar. Gleichzeitig wird dabei deutlich, was die *Hypothesen* dieser Forschung über Systemsteuerung sein können und was sich als Forschungsgegenstand im Umkreis von Governance herauskristallisiert hat. In der Ausschreibung heißt es:

„Gefördert werden theorie- und hypothesengeleitete empirische Forschungsvorhaben zu Fragen in folgenden Bereichen:

1. Umfassende Steuerungsinnovationen: Steuerungen (z.B. mittels Evaluation, Schulinspektion, stärkerem Management, mehr Wettbewerb, Bildungsstandards und zentraler Abschlussprüfungen) erfordern in vielfacher Weise eine veränderte, oftmals erhöhte vertikale und horizontale Handlungskoordination zwischen Akteuren im Bildungswesen. Insbesondere auch im Zuge stärkerer Eigenverantwortung ist eine verstärkte Rückmeldung der Ergebnisse von spezifischen Steuerungsinstrumenten und -prozessen zwischen verschiedenen Ebenen notwendig ... Untersucht werden soll die Frage, unter welchen Rahmenbedingungen welche Instrumente, die in den letzten Jahren im Bildungsbereich verwendet wurden, welche Effekte für den Erhalt und den Ausbau der Qualität von Bildungssystemen erzielen ...
2. Umstrukturierung der Bildungsverwaltungen: Mit der Umstellung auf output-orientierte Steuerung verändern Bildungsverwaltungen ihre Ausrichtung, Tätigkeit und Funktion, um angesichts neuer Steuerungserfordernisse ihrer koordinierenden Funktion im Bildungssystem gerecht zu werden. In welcher Weise und mit welchem Erfolg werden diese Umstrukturierungsprozesse gestaltet? ...
3. Steuerungswirkung im Hinblick auf Unterricht: Im Mittelpunkt steht die Frage, welche Wirkungen neue Formen der Steuerung im Hinblick auf inner-schulische Koordination und Strukturbildung, auf Unterrichtsgestaltung sowie für Lernergebnisse der Schülerinnen und Schülern haben ...“ (<http://www.bmbf.de/de/13417.php>).

In diesem Rahmenkonzept haben die Antragsteller/-innen jeweils ihre Akzente gesetzt. Die zentralen Projekte, die gefördert wurden, sind dann im Umkreis der neuen Governance angesiedelt (<http://www.empirische-bildungsforschung-bmbf.de/zeigen.html?seite=8832>).

So formuliert eines der geförderten Projekte: „Zentrale standardbasierte Tests von Schülerleistungen bilden das Kernstück einer outputorientierten Steuerung von Bildungsqualität. Sie dienen der empirischen Überprüfung von Bildungsstandards und sollen – vermittelt über die Rückmeldung von Testergebnissen – die Leistungsfähigkeit des Bildungssystems sichern bzw. steigern.“ Hier wird eine klare Wirkungsthese formuliert, die mit der *Intervention*, mit dem *Treatment*, verbunden ist.

In Bezug auf diese großen Interventionen der neuen Steuerung lässt sich in der Tat, z.B. in der Schweiz, in Deutschland und in Österreich, auf Governance-Ebene ein *faktischer Konsens* in der derzeitigen Qualitätssicherung im Bildungswesen beobachten. *Zentrale Prüfungen, überschulische Vergleichsarbeiten*, auf vereinbarte *Standards* und *Kompetenzen* ausgerichtete Bildungspläne werden in allen Ländern akzeptiert. Begleitet sind sie von *Qualitäts-Referenzrahmen*, die in eindrucksvoller Entfaltung *Qualitätsstandards für einzelne Schulen* präzisieren, und von weitgehenden Verfahren der *Autonomie* und *Führungsfähigkeit* von Schulen vor Ort, die nicht selten in einem produktiven Qualitätswettbewerb stehen. Standardsicherungsmodelle und Autonomiemodelle ergänzen sich dabei, um die *Qualitätsentwicklung* und die *Qualitätssicherung* funktionsfähig zu machen.

So scheinen wir auf dem Wege zu einer neuen *Grammatik der Qualitätssicherung*: Auf Systemebene sichert ein *kontinuierliches Monitoring* (Arbeitsgruppe Bildungsberichterstattung 2008; Böttcher, Bos, Döbert und Holtappels 2008) den Wissensstand über Stärken und Schwachstellen im Bildungswesen. Standards und Kompetenzerhebung lassen dies bis auf die Schule- und Klassenebene ausweiten. Da sich die geschlossenen Systeme und die *Rückmeldekultur* als defizitär erwiesen hatten, kann das Qualitätsbewusstsein in Bezug auf die Leistungen durch *überschulische Vergleichs- und Diagnosearbeiten* gestärkt werden. Damit hätten auch die Schulleitungen die Chance, Einblicke in die Qualität der Lehre, aber auch Einblicke in die Lernvoraussetzungen der Schüler/-innen-schaft zu gewinnen. *Qualitätsarbeit* würde erstmals sichtbar und verhandelbar. Damit hätten sich der verobjektivierbare Stand des Wissens über das Bildungswesen und vor allem auch das *Qualitätsbewusstsein* wesentlich verbessert.

In der Summe sind also zwei Kernelemente der Interventionen sichtbar: solche der Standardsicherung und solche der Qualitätssicherung auf schulischer Ebene über Autonomiegewährung und Evaluation, die wiederum im Kontext einer rechtlichen Veränderung von Verantwortlichkeiten steht, die verwaltungstechnisch im Rahmen des New Public Management platziert ist (s. dazu Dubs 1996; Schedler 1995). Hier geht es auch um die Veränderung der Schulaufsicht, um neue Führungsstrukturen, um die Entwicklung eines Qualitätsrahmens für die Schulen und um Verfahren der internen und externen Evaluation. Sie konnten

auf die schon länger laufenden Initiativen und Entwicklungen von „selbständigen Schulen“ aufbauen (Berkemeyer 2010).²

2 **Neue Steuerung im Rahmen des empirisch-analytischen Forschungsparadigmas der school effectiveness-Forschung**

Die neue Steuerung steht im Kontext einer Auffassung, nach der Veränderungen und Entwicklungen im Bildungswesen auf der Grundlage ihrer nachgewiesenen Wirksamkeit, also „evidence-based“ erfolgen sollen. Diesem Anspruch muss sie sich konsequenterweise in gleicher Weise aussetzen. Das Erfolgskriterium ist dabei die Steigerung der Lernergebnisse. Sie ordnet sich damit in den Kontext der school effectiveness-Forschung ein, die sich mit der Prädiktion von Schulleistungen beschäftigt. Im Kontext eines umfassenden Wirkungsmodells müsste dann lokalisiert werden, welche „Wirkkraft“ Verfahren der Standarderreichung über flächendeckende Testverfahren zukommt, um Lernergebnisse positiv zu beeinflussen.

Welche Modelle und Forschungsergebnisse stehen dabei zur Verfügung? Als allgemeines Wirkmodell hat sich das Input-Prozess-Output- und Kontext-Modell der OECD als hilfreich für die Ordnung der zentralen Faktorengruppen erwiesen. Bestmöglicher Input, bestmögliche Prozesse und günstige Kontexte fördern die Lernergebnisse, den „Output“.

Welche Faktoren sind dabei entscheidend? Die Grundlinien moderner Antworten, wie sie z.B. in der holländischen Schulforschung formuliert werden (Creemers, Scheerens und Reynolds 2000; Creemers und Kyriakides 2008; Scheerens und Bosker 1997), sind erstaunlich einfach und einhellig. Auf Seiten der Institution sind vor allem *Zeitinvestitionen* und *inhaltliche Lerngelegenheiten* ausschlaggebend, die auf Makro-, Meso- und Mikroebene angesiedelt sein

2 Fragen der Steigerung von Equity und Equality sind davon allerdings nicht berührt. Dazu sollen andere Instrumente verhelfen, etwa die Veränderung der Schulstrukturen, die Verlängerung gemeinsamen Lernens und Vermeidung von Leistungsgruppierungen, die Abschaffung der Klassenwiederholung, die Abschaffung der Querversetzung in Schulformen, die Stärkung der Elternrechte über freie Schulwahl usw. Um die empirische Untersuchung und Untersuchbarkeit dieser letzten Fragen soll es im Folgenden nicht gehen. Es soll die neue Standardsicherung über die jeweiligen Verfahren der Kompetenzmessung für New Governance als Paradigma herangezogen werden. Sie ist gewissermaßen die große Intervention zur Steigerung der Lernergebnisse in Schulen.

können. Dazu kommen *Qualitätsindikatoren* des Angebotes, die auf Unterrichtsebene (Helmke 2009) über verschiedene Studien, z.B. über Videostudien, immer wieder erarbeitet wurden (Klieme, Lipowsky, Rakoczy und Ratzka 2006; Klieme, Pauli und Reusser 2009): Es sind dies

- ein effizientes und funktionales Klassenmanagement mit möglichst wenig Zeitverlust,
- eine gute Strukturierung des Unterrichtes,
- ein kognitiv aktivierender Unterricht,
- eine gute Abstimmung des Unterrichtes auf die motivationalen und emotionalen Bedürfnisse der Schüler/-innenschaft,
- ein Unterricht, der differenzierend und die Schüler/-innen beteiligend auf die Heterogenität der Lernstände eingeht.

Quantität und Qualität des Angebotes erweisen sich als die zentralen Größen, die die Lernergebnisse bestimmen. Eine zweite große Gruppe von Faktoren, die Lernergebnisse beeinflussen, ist dann auf Schüler/-innenseite zu suchen. Kognitive Lernvoraussetzungen und motivationale Konstellationen stehen dabei im Mittelpunkt. Sie beeinflussen die Qualität der Nutzung von Lernangeboten. An der Schnittstelle von Angebot und Nutzung zeigen sich zusätzliche Qualitätsmerkmale. Sie bestehen in der *Adaption* von gewährter mit benötigter Lernzeit (Carroll 1963). Dieses Prozessmerkmal guten Unterrichts realisiert sich in aufeinander abgestimmten ko-konstruktiven Prozessen zwischen Lehrpersonen und Schüler/-innen (Reusser 1998).

Dieses Modell der Leistungswirksamkeit ließe sich formalisieren (Fend 2008, S. 24 ff.):

1. $LE = f(A, N, A*N)$
LE = Lernergebnisse
A = Angebot
N = Nutzung
2. $A = f(QuantA, QualA)$
QuantA = Allokation von Zeit und Lerngelegenheiten
QualA = Qualitätsmerkmale von Unterricht
3. $N = f(QuantN, QualN)$
QuantN = Zeitinvestitionen in Lernen
QualN = kognitive Lernvoraussetzung, Motivation
4. $A*N$ = Interaktion von Angebot und Nutzung (Adaptivität)

Wenn also mehr und qualitativ bessere Angebote gemacht werden und diese mit mehr Zeitinvestitionen und Motivation auf Schüler/-innenseite genutzt werden, dann resultieren daraus auch Steigerungen der Lernergebnisse.

Die school effectiveness-Forschung, die hier angesprochen ist, konzentriert sich schwerpunktmäßig auf den Kern des schulischen Geschehens, auf Lehren und Lernen. Die Qualität der Mikrostruktur schulischer Lehrprozesse steht damit im Vordergrund (s. dazu Creemers und Kyriakides 2008). Die neue Steuerung scheint von dieser Ebene des Handelns weit entfernt. Welchen Stellenwert können im Rahmen eines mikroorientierten, quantitativ arbeitenden Modelldenkens aus der Makrosteuerung des Bildungswesens kommende Standards und deren Testung für die Steigerung der Lernergebnisse haben? Nach dem obigen Modell wäre dies nur dann der Fall, wenn die Makrosteuerung Kernbedingungen für die Steigerung von Lernergebnissen beeinflussen würde.

Die empirische Forschungslage dazu hat Maag Merki (2010) im Handbuch zur neuen Steuerung referiert. Sie ist nicht so eindeutig, wie man es wünschen würde. Internationale Vergleiche sind eine zentrale Quelle, da in ihnen Variationen in Makrofaktoren des Bildungswesens gefunden werden können. Wößmann (2002) hat die Datenbasis der TIMSS-Studien genutzt, um zu überprüfen, welche Makrofaktoren einen Einfluss auf die Leistungsergebnisse haben. Er glaubte nachweisen zu können, dass jene Länder anderen überlegen sind, die standardisierte Messungen auf Abiturniveau eingerichtet haben. Bei den Maßnahmen zur Stärkung der Schulautonomie war dies nicht der Fall. Doch auch hier kovarierte das Instrument der zentralen Prüfungen als Instrument auf der Makroebene mit anderen, so dass eine exakte kausale Zuordnung schwierig war. Ähnliches gilt für die Übertragbarkeit von amerikanischen Studien zum „high stakes testing“, wonach getestete Schulleistungen gravierende Konsequenzen für Lehrpersonen und auch Schüler/-innen haben. Auf deutsche und österreichische Bildungssysteme sind die Ergebnisse schon deshalb nicht übertragbar, weil ein solches Testen bislang nicht umfassend implementiert ist. Eine Ausnahme bilden jedoch die Abiturprüfungen in den meisten Bundesländern in Deutschland. Da sie noch nicht lange eingeführt wurden, eröffnet sich die Chance, ihre Wirksamkeit zu überprüfen (Maag Merki 2010).

Bei einer *kausalen* Zuordnung müssten sich Faktoren auf der Makroebene auf Quantität oder Qualität des Lernangebotes und die bestmögliche Zeitnutzung auswirken. Sie müssten das Qualitätsbewusstsein in Bezug auf zu erreichende Lernziele stärken und so zu einer Fokussierung der Lehranstrengungen und zu einer Stärkung der Lehrer/-innenmotivation führen. Auf Schüler/-innenseite müsste nachgewiesen werden, dass mehr Motivation und Zeit in die Nutzung der besseren Lernangebote investiert werden.

Kompetenzzuwächse müssten nach den Forschungen zu „school effectiveness“ über die Veränderung jener proximalen Faktoren vermittelt werden, die mit Kompetenzsteigerungen zusammenhängen. Auf diesem Hintergrund hat es die empirisch analytische Forschung schwer, nachzuweisen, dass mit der neuen Standard-Governance auf *Makroebene* Steigerungen in den Lernergebnissen auf der *Mikroebene* erzielt werden können. Sie wird auf dem Hintergrund ihrer Forschungsansätze keine starken *direkten* Wirkungen der neuen Steuerung auf Leistungsergebnisse vermuten und Hypothesen eher für *indirekte* Wirkungen, vermittelt über die proximalen Faktoren, formulieren (Scheerens et al. 2006).

Der Nachweis der Wirksamkeit der neuen Steuerung ist jedoch ein anspruchsvolles Vorhaben und zwar nicht zuletzt deshalb, weil die wissenschaftlichen Kriterien, Wirkungen nachzuweisen, sehr rigoros sind:

1. Ein empirischer Nachweis wäre darauf angewiesen, ein *quasi-experimentelles Design* zu realisieren, das Kontrollgruppen oder zumindest Vergleichsgruppen als Bezugsgrößen erfordern würde und in einem Pre-Post-Testdesign positive Veränderungen auf der Mikroebene des unterrichtlichen Angebotes durch Makrofaktoren nachzuweisen hätte. Wenn ein ganzes Bildungswesen von einer neuen Steuerung betroffen ist, dann fehlen diese Möglichkeiten, es sei denn, die zeitliche Umstellung in verschiedenen Sektoren des Bildungswesens bzw. in verschiedenen Ländern kann dazu klug genutzt werden. Ansonsten setzt dieser Ansatz den Vergleich mehrerer Bildungssysteme mit unterschiedlichen Steuerungsmodalitäten voraus. Durch die internationalen Schulleistungsvergleiche ist allerdings die Datenlage so gut wie noch nie in der Geschichte der Bildungsforschung.
2. Erfahrungsgemäß lässt sich eine Intervention wie jene der „Standardmessung“ nur schwer als „reine“ Variable isolieren. In der Regel sind solche Interventionen *konfundiert* mit anderen Faktoren (curriculare Veränderungen, neue Lehrer/-innenbildung, finanzielle Ressourcen, rechtliche Kompetenzverteilungen). Besonders akut ist dieses Problem bei der Versuchung, Ergebnisse aus anderen Ländern auf das jeweilige nationale Bildungssystem zu übertragen. Einzelne Maßnahmen stehen in der Regel in einem systemischen Kontext, im Kontext einer „stimmigen“ *Grammatik* eines Bildungswesens, der es erschwert, sie als Einzelvariablen zu isolieren und in ihren Wirkungen zu bewerten.
3. Entsprechende Implementationen auf Makroebene – wie z.B. jene der Standardmessung – sind *langfristige* Unternehmungen; dies nicht nur was die Entwicklung, sondern auch was die Zeiträume für Effekte angeht. Fortschritte werden jedoch häufig schnell erwartet, und Reformen können nicht so lange hinausgeschoben werden, bis entsprechende Forschungsergebnisse vorliegen.

Damit besteht für ein Forschungsdesign die Gefahr, dass sich das „Treatment“, also die Intervention in der Gestalt der Steuerung des Bildungswesens, fortlaufend ändert.

Diese Schwierigkeiten der praktischen Realisierung von empirischen quantitativen Studien zur Wirksamkeit unterschiedlicher Makrofaktoren sprechen nicht grundsätzlich gegen den empirisch-analytischen Ansatz. Sie dämpfen lediglich die Erwartung, klare kausale Zuordnungen von Makrostrukturen zu Effekten auf der Lernebene finden zu können und dies in einem Zeitrahmen, der mit den politischen Entscheidungen zu Veränderungen im Bildungswesen kompatibel ist. Es dürfte sich aber lohnen, hier alle Möglichkeiten auszuschöpfen.³

Schwerer als die Schwierigkeiten der praktischen Realisierung klarer Wirksamkeitsdesigns wiegt die *theoretische* Frage, wie die *Mechanismen*, die die erwünschten Steigerungen in den Lernergebnissen bewirken könnten, gefunden und belegt werden können. Die lange *Wirkungskette* von Standardformulierungen und daran gebundenen Testreihen zu besseren Mathematikkenntnissen, Englischleistungen, Lese- oder Orthographieleistungen auf Schüler/-innenebene erfordert eine Spezifizierung der vermittelnden Prozesse zwischen der Makrosteuerung bis auf die Ebene der Kompetenzsteigerungen bei Schülerinnen und Schülern. Dazu bedarf es theoretischer Konzepte, die dies darstellbar machen. Sie müssten dann Hypothesen ermöglichen, wie Unterschiede in der Makrosteuerung zu Unterschieden auf der Mikroebene des Lehrens und Lernens führen. Das zentrale *methodische* Problem ergibt sich aus der Forderung, für solche Vermittlungsprozesse *empirische Belege* zu erbringen.

3 Theoretische und methodische Beiträge von Governance-Konzepten zur Wirksamkeitsanalyse

Bei der Suche nach theoretischen Konzepten, die bei der Aufklärung von Wirkungsketten zwischen der Makrosteuerung des Bildungswesens und Effekten auf der Schüler/-innenebene hilfreich sein könnten, sind Bildungsforscher/-innen auf jene des *Educational Governance* gestoßen (s. bspw. Altrichter, Brüsemeister und Wissinger 2007; Altrichter und Maag Merki 2010; Berkemeyer 2010). Sie weckten die Hoffnung, dass durch sie die Zusammenhänge zwischen mehreren

3 Dass wissenschaftliche Studien mit diesen Problemen auch produktiv umgehen können, zeigt die Gesamtschulforschung der 1970er Jahre (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung, 1982; Fend, 1982).

Educational Governance als Forschungsperspektive
Strategien. Methoden. Ansätze

Maag Merki, K.; Langer, R.; Altrichter, H. (Hrsg.)

2014, VII, 352 S. 9 Abb., Softcover

ISBN: 978-3-658-06442-6