

Zur schulpraktischen Nutzung von Daten: Konzepte, Strategien, Erfahrungen

Einleitung

Vergleichsuntersuchungen (Orientierungs-, Parallel- und Vergleichsarbeiten bis hin zu Lernstandserhebungen) dienen nicht nur der Standortbestimmung und Rechenschaftslegung, sondern auch der Verbesserung des Lehrens und Lernens, der Steigerung der Unterrichtsqualität und der Förderung der Professionalisierung der Lehrerschaft. Wenn dieses zweite Ziel nicht zufrieden stellend erreicht wird, „haben die groß angelegten und teuren Evaluationsprojekte einen wesentlichen Zweck verfehlt“ (Helmke 2004, S.1)

In der Praxis zeigt sich allerdings, dass das erste Ziel bisher wesentlich erfolgreicher erreicht worden ist, als das zweite Ziel. Die Rückmeldungen der Leistungsdaten an die Schulen haben zwar zu inhaltlichen Auseinandersetzungen geführt, es gib aber nur wenige Hinweise, dass sie sich auch auf die Gestaltung des Unterrichts beziehen.

Warum führen Rückmeldungen über Leistungsdaten aus zentralen Tests an Schulen kaum zu Bemühungen um die Weiterentwicklung von Unterricht?

Zwei Herangehensweisen an diese Thematik lassen sich unterscheiden.

- (1) Leistungsdaten werden angeboten und es wird erwartet, dass Lehrer/innen daraus Schlussfolgerungen für das Handeln ziehen, d.h. Entwicklungsinitiativen ableiten. Der Fokus der Bemühungen liegt auf der Gestaltung des Angebots. Wenn es gut genug ist sollte der Nutzung nichts mehr im Wege stehen
- (2) Lehrer/innen bemühen sich kontinuierlich individuell und gemeinsam um die Weiterentwicklung und Überprüfung von Unterrichtsqualität. Leistungsdaten werden als eine wichtige Informationsquelle in dieses Bemühen einbezogen. Der Fokus der Bemühungen liegt nicht auf der Gestaltung des Angebots, sondern auf der Schaffung von Bedingungen, unter denen Schulleiter/innen und Lehrer/innen sich um die Qualität von Unterricht bemühen und Entwicklung und Evaluation als selbstverständliche professionelle Tätigkeiten betrachten.

Der Vortrag geht von der These aus, dass die Annahmen, die dem ersten Muster zugrunde liegen etwas zu kurz greifen. Nicht weniger zentral, wenn nicht wichtiger als die Frage, wie sich Leistungsdaten nutzen lassen, erscheinen folgende Fragen zu sein: Warum ist eine kontinuierliche Weiterentwicklung und Evaluation von Unterricht erforderlich? Wie kann sie vor sich gehen und welche Voraussetzungen sind dazu erforderlich? Zu diesen Voraussetzungen gehören extern gewonnene aussagekräftige Leistungsdaten über Schülerleistungen, die allerdings nicht unabdingbar sind, wie gute Schulen ohne externe Leistungsdaten zeigen.

Im Vortrag werden zunächst einige Untersuchungsergebnisse über die Rezeption, Reflexion und Nutzung von Leistungsdaten aus Untersuchungen (u.a. QuaSUM2, VERA3, Ergebnisrückmeldungen zu den österreichischen standardsbezogenen Tests) vorgestellt. Anschließend wird versucht auf die oben genannten Fragen näher einzugehen und diskutiert, weshalb das zweite Ziel der Vergleichsuntersuchungen, die Verbesserung des Lehrens und Lernens, mit den bisherigen Maßnahmen bisher nur sehr eingeschränkt erreicht werden konnte.

1. Zur Ausgangssituation: Einige Befunde zur Nutzung des Angebots an externen Leistungsdaten

Die Mehrzahl der verfügbaren Studien beruht auf Fragebogendaten, deren Aussagewert sich in einem so sensiblen Bereich wie der Evaluation von Unterricht allerdings in engen Grenzen halten dürfte. Kohler & Schrader (2004, S. 12f.) nennen zahlreiche Einschränkungen, vor allem hinsichtlich der Glaubwürdigkeit von Angaben über tatsächliche Veränderungen im Unterricht.

Rainer Peek (2004) berichtet über die Ergebnisse einer postalischen Nachbefragung zur QuaSUM2-Studie. Es zeigte sich starke Auseinandersetzung mit den QuaSUM-Befunden – sowohl seitens der Mathematiklehrkräfte und der Klassenlehrkräfte als auch seitens der Fachkonferenzleiter und der Schulleitungen. Besonders zahlreich waren die Auseinandersetzungen in den Fachkonferenzen. Die Rezeption war primär fachbezogen. Die Auseinandersetzung erfolgte intensiver und länger in Schulen, in denen das Schulklima und die allgemeine wie fachbezogene Zusammenarbeit als gut eingeschätzt wurden (ebenda, S. 81f). Es gab aber nur wenige klar definierbare, unmittelbare Konsequenzen für den Schulalltag und für eine optimierte Unterrichtsgestaltung. Die Zahl der Lehrkräfte, die aus der Lektüre und Reflexion der Ergebnisse für ihre pädagogische Arbeit Konsequenzen gezogen haben, lag jeweils unter 20% und bei den Mathematik-Fachkonferenzen und Schulleitungen unter 30%. Die Nennungen bezogen sich in der Regel auf eine Optimierung von gemeinsamen Standards und auf methodische und didaktische Absprachen. Die Konsequenzen bezogen sich auf Kontextplanung und Auswertung von Unterricht, nicht aber auf die Prozessgestaltung. Aus der Sicht der Lehrer enthielten die Daten dafür keine Ansatzpunkte.

Das Unterstützungsangebot (durch qualifizierte Multiplikatoren und die Schulaufsicht) wurde von den Schulen kaum bis gar nicht angenommen. (ebenda, S. 83).

Grabensberger u.a. (2008) haben bei einer Befragung der an den Pilottestungen für Englisch und Mathematik auf der 8. Jahrgangsstufe in Österreich ähnliche Ergebnisse erhalten:

- Das allgemeine Interesse an den Rückmeldungen war relativ hoch: Etwa 70% haben sich mit den Ergebnissen ihrer Klassen beschäftigt. Etwa 30 % berichten von einer verbesserten Kommunikation und Zusammenarbeit im Kollegium. Angaben über eine Verbesserung der Zusammenarbeit lassen sich auch aus anderen Studien entnehmen. Allerdings hat Kohler in ihrer Studie der Rezeption der TIMSS Studie durch Lehrer/innen, Eltern und Schulaufsicht (2004, S. 33) festgestellt, dass Zusammenarbeit ein sehr schillernder Begriff ist und auch nur das sporadische Ausleihen von Unterrichtsmaterial bedeuten kann.
- Für ca. 60% hat sich „wenig verändert.“ Etwa 40% geben an, dass sie die differenzierten Stärken-Schwächenprofile ihrer Klassen zum Anlass nehmen, „ihren Unterricht besser auf diese Gegebenheiten abzustimmen“ (Grabensberger u.a. S. 71f.). Etwa 50% der Lehrer/innen berichten von einer Verbesserung des diagnostischen Wissens, allerdings nicht auf der Ebene des einzelnen Schülers.
- Deutlich positiver werden die Ergebnisrückmeldungen von den Schulleiter/innen bewertet. Sie halten auch die praktische Nützlichkeit der Ergebnisse für bedeutend höher als die Lehrer/innen.
- Nur wenig weist darauf hin, dass das verbesserte diagnostische Wissen tatsächlich zum Ausgangspunkt von Entwicklung gemacht wird. Nach Grabensberger u.a. (ebenda, S. 75) dürfte dies damit zusammenhängen, dass die Ergebnisse um oder über dem Erfahrungswert der Lehrer/innen liegen und damit keine starken Veränderungsimpulse induzieren könnten.

Als wichtiges Forschungsproblem wird im Resümee u.a. angeführt, dass nur wenig darüber bekannt ist, „ob das Interesse der Lehrkräfte, das in dieser Studie zum Ausdruck kommt, sich auch in realen Handlungsansätzen fortsetzt und welche Formen der Unterstützung Schulen benötigen, um aus Lernstandsdiagnosen Entwicklungsansätze für kompetenzorientierten Unterricht abzuleiten“ (ebenda, S. 76). Die Autoren zitieren auch aus einer Studie über die Situation in England Frankreich und den Niederlanden (van Ackeren 2007, S. 207), dass zu den tatsächlichen Effekten externer Evaluationen wenig bekannt ist und Steuerungsstrategien zu einem großen Teil nicht auf empirischen Belegen sondern auf Überzeugungen beruhen (ebenda, S. 76f.)

Sofern Leistungsdaten überhaupt für den Unterricht genutzt werden, scheint sich das Niveau eher auf oberflächlicher Ebene zu bewegen. Hinweise dazu gibt es in der Interviewstudie (36 befragte HS- und RS-Lehrer/innen) aus Baden-Württemberg, die von U. Maier (2007) zitiert wird: Auf die Frage nach den Auswirkungen der Testrückmeldung auf den zukünftigen Unterricht ergab sich Folgendes:

- | | |
|---|----|
| ➤ Keine nennenswerte Auswirkungen | 20 |
| ➤ Bestimmte Aufgabenstellungen übernommen | 9 |
| ➤ Mehr Wiederholungen | 7 |
| ➤ Arbeitstechniken verstärkt üben | 4 |

Interessant ist die wesentlich positivere Rezeption der Ergebnisse, die für den Grundschulbereich auf Grund einer nicht repräsentativen schriftlichen Befragung von Lehrer/innen aus dem Projekt VERA 3 berichtet werden (Hosenfeld u.a. 2007). 96 % der befragten Lehrer/innen berichten, dass die Ergebnisse von ihren Erwartungen abweichen und nur 10,9 % führen die Abweichungen auf ihren eigenen Unterricht zurück, sondern 50% auf die Schüler, 30,4% auf die Testkonstruktion und 10,7 % auf sonstige externe Bedingungen. Externe Attributionen, die in dieser Studie extrem zum Ausdruck kommen, sind offensichtlich ziemlich verbreitet und wurden in mehreren Studien festgestellt (z. B. in Kohler 2004, S. 20) Trotzdem und das ist bemerkenswert geben mindestens 90 % der Lehrer/innen an, etwas an ihrem Unterricht verändert zu haben:

- Wiederholung oder Vertiefung von Gebieten, in denen es Schwächen gab (75,4 %)
- Entwicklung neuer Übungsaufgaben, angelehnt an VERA-Aufgaben (54,2 %)
- Verwendung von an VERA angelehnte Aufgaben für Klassenarbeiten (50,5 %)
- Ausprobieren neuer Unterrichtsmethoden (49,4 %).

Detlev Leutner u.a. (2007) haben am Beispiel der Schulrückmeldungen über Lernstandserhebungen (Stufe 8) in NRW zwar eine relativ hohe Bereitschaft festgestellt, sich mit den Ergebnissen auseinanderzusetzen und Konsequenzen abzuleiten (bei Schulleitungen höher als bei den Lehrkräften). Die am häufigsten angegebene Konsequenz bezog sich allerdings auf Testinhalte und –formate und nicht auf die Unterrichtsgestaltung.

Julia Schneewind und Harm Kuper (2008) haben die Rezeption der Berliner Längsschnittstudie zur Lesekompetenzentwicklung untersucht und in mehreren Wellen 56 Grundschullehrer/innen schriftlich und 10 Lehrer/innen anhand von fokussierten Interviews befragt. Die Ergebnisrückmeldungen wurden von den Befragten durchwegs als nützlich angesehen, sie haben aber „nicht dazu inspiriert, im Unterricht etwas zu verändern“ (ebenda, S. 25) Aus Sicht der Autor/innen wird die Verwendung der zurückgemeldeten Daten durch zwei Faktoren behindert, durch:

- mangelnde Einbettung des Rückmeldeprozesses in Verfahren zur Qualitätssicherung und durch
- fehlende Ressourcen (vor allem Zeit)

Ohne Anbindung an gemeinsame Vorhaben zur Unterrichts- und Schulentwicklung seien Rückmeldungen wenig hilfreich, weil die Schule nur als Ganzes Veränderungen dauerhaft und effektiv umsetzen könne (ebenda, S. 26).

Fazit: Trotz des in mancher Hinsicht bemerkenswerten Ausreißers, über den Hosenfeld u.a. (2007, s.o.) berichten und der zu einigen Fragen Anlass gibt, besteht eine erhebliche Unzufriedenheit mit den Auswirkungen der externer Leistungsrückmeldungen auf die Gestaltung von Unterricht.

2. Versuche zur Erklärung der Befunde

Ursachen können im Wesentlichen im Wesentlichen in zwei Bereichen liegen: in der Struktur des Angebots an Leistungsdaten und in Merkmalen der Struktur der potentiellen Nutzung der Daten.

2.1 Merkmale der Struktur des Angebots

In ihrer Einführung zum Hefthema der Empirischen Pädagogik 2004 kommen Kohler & Schrader (2004, S. 9) zum Schluss, dass die Rezeption externer Evaluation einerseits und die Gestaltung von Ergebnisrückmeldungen andererseits weitgehend unerforschte Felder sind. In der Zwischenzeit sind einige neuere Studien erschienen. Zumindest was die Gestaltung der Ergebnisrückmeldungen anlangt, dürfte dieses Urteil nicht mehr stimmen, da diese bereits ein relativ hohes Niveau erreicht haben.

Bei der Gestaltung der Rückmeldungen finden folgende Qualitätsmerkmale zunehmende Beachtung (vgl. z.B. Helmke 2004):

- Umfang und Lesbarkeit (möglichst geringer Umfang und übersichtliche Darstellung mit geringen Ansprüchen an statistische Vorkenntnisse.)
- Bezugsnormen (Neben der kriterialen Bezugsnorm bestehen zunehmende Bemühungen um Fairness bei sozialen Bezugsnormen etwa durch Vergleiche mit Klassen ähnlicher Zusammensetzung oder durch Angabe von Erwartungswerten)
- Aggregationsniveau: Die Daten werden in unterschiedlichen Aggregationsniveaus angeboten, in Form von Verteilungen innerhalb einer Klasse, Leistungsprofilen etc.)
- Kontextdaten: Es werden zahlreiche Kontextdaten angeboten (u.a. auch Soziale Schicht, Sprachhintergrund, Motivationslage der Schüler/innen, Unterrichtswahrnehmung der Schüler/innen etc.)
- Zeitdimension: Neben Querschnittuntersuchungen werden auch bereits Längsschnittuntersuchungen durchgeführt.
- Timing: Die Zeit zwischen Datengewinnung und Datenrückmeldung ist wesentlich kürzer geworden (von mehreren Monaten zu wenigen Wochen).
- Zusatzinformationen zu den Daten: Die Daten werden durch unterrichtsbezogene Zusatzinformationen (Fehleranalysen, fachdidaktische Anregungen etc.) ergänzt.

Trotzdem sind Überlegungen zur weiteren Verbesserung des Angebots von Bedeutung: z.B. im Hinblick auf die Fairness bzw. Aussagekraft von Vergleichsdaten (Steffens 2004, S. 6) oder ergänzende diagnostische Materialien¹ und fachdidaktische Anregungen). Dem hohen Niveau der Leistungsrückmeldungen entspricht allerdings nicht das Niveau, auf dem Entwicklung und Evaluation von Unterricht an den Schulen erfolgen. Weil hier eine erhebliche Diskrepanz besteht ist die Nutzung der Leistungsrückmeldungen aus meiner Sicht (siehe obige These) begrenzt und wird wahrscheinlich auch so bleiben, solange auf der

¹ Vgl. z. B. die bei SINUS eingesetzten diagnostischen Materialien für das Fach Mathematik aus Schweden und Australien (vgl. z.B.: <http://www.learn-line.nrw.de/angebote/sinus/projekt5/index.html>).

Nutzerseite Entwicklung und Evaluation kein Merkmal der Arbeitskultur ist. Solange dies nicht der Fall ist muss damit gerechnet werden, dass externe Tests eher als Kontrolle und weniger als Unterstützung verstanden werden. Damit fehlt aber eine wichtige Voraussetzung für eine konstruktive Nutzung der Ergebnisse: die Nachfrage. Trotzdem gibt es in den Daten Hinweise auf positive Effekte, wenn auch die Belege z.T. widersprüchlich sind:

- Sie scheinen Einfluss auf die Auswahl der fachlichen Inhalte zu haben;
- Sie scheinen Lehrer/innen zur Übernahme bzw. Entwicklung komplexer Aufgaben anzuregen;
- Sie dürften Einfluss auf die diagnostischen Fähigkeiten von Lehrer/innen oder zumindest auf das Bewusstsein ihrer Bedeutung haben;
- Sie dürften zum Verständnis der Kompetenzstufen und deren Orientierung an Stärken beitragen;
- Sie dürften die Wahrnehmung der Leistungsunterschiede zwischen den Schüler/innen erweitern.

Auch wenn ein direkter Einfluss der externen Leistungserhebungen auf die Gestaltung von Unterricht kaum nachweisbar ist, dürften sie dennoch zu einem „Paradigmenwechsel von der inhaltlichen Orientierung des Unterrichts zur Kompetenzorientierung“ (Uli Steffens) einen gewissen Beitrag leisten.

Die Festlegung im Positionspapier der letzten EMSE-Tagung (Peek u.a. 2007, S. 1) halte ich allerdings für missverständlich:

„Messen sowie entwickeln werden als zwei Seiten ein und derselben Münze betrachtet. Beide sind aufeinander zu beziehen und bilden einen unauflöselichen Zusammenhang“.

Wenn beide als unverzichtbare Aufgaben des Lehrenden gemeint sind, trifft die These selbstverständlich zu. Wenn die zweite Seite der Münze sich aber auf die Lernstandserhebung bezieht, halte ich sie für falsch oder zumindest für stark überzogen. Lernstandserhebungen sind bestenfalls ein kleiner wenn auch wichtiger Beitrag zur kontinuierlichen Aufgabe des Lehrers, sich um die Überprüfung des Leistungsfortschritts der Schüler/innen zu kümmern. Ich beschäftige mich daher im Folgenden mit strukturellen und individuellen Voraussetzungen der potentiellen Nutzung von Leistungsdaten.

2.2 Merkmale der Struktur der Nutzung

Eine erfolgreiche Rezeption und Umsetzung von Evaluationsergebnissen hängt von einer Vielzahl von Bedingungen ab, die mit der Gestaltung der Rückmeldungen wenig bis nichts zu tun haben: von der einzelnen Lehrkraft (ihren Motiven, Werthaltungen, Kompetenzen), von der Schule (den Führungsqualitäten der Leitung, ihrer Unterstützung und dem Klima an der Schule), von der Schulaufsicht (ihrer Fähigkeit, Spielräume zu geben und unspezialisierten Außendruck zu erzeugen), von regionalen Unterstützungsstrukturen, die Beratung und Fortbildung zugänglich machen. Helmke und Schrader haben bereits 2001 auf die Komplexität dieser Bedingungen hingewiesen. Auch das von Helmke (2001) entwickelte „Makromodell der Einflussfaktoren schulischer Leistungen“ mit seinen proximalen (u.a. Vorkenntnisniveau und Unterrichtsqualität) und distalen Bedingungsfaktoren (z.B. Schulqualität und soziale Schicht) wäre geeignet, zur Vorsicht vor tief greifenden Effekten der Leistungsrückmeldungen zu mahnen.

Zur Untersuchung der wichtigsten potentiellen Gründe für die nicht zufrieden stellende Nutzung der Evaluationsdaten wird im Folgenden eine Modellvorstellung herangezogen werden, die sich an Giddens' (1992) Dualität von Struktur und Aktion (vgl. auch Elliott 1993) anlehnt und in der Governance-Theorie weiter ausdifferenziert wurde (Altrichter/Heinrich 2007). In etwas vereinfachter Darstellung werden dabei strukturelle und individuelle

Voraussetzungen für Innovationen bzw. für die Weiterentwicklung und Evaluation von Praxis (und damit zur Nutzung von Ergebnissen externer Evaluationen) unterschieden. Beide lassen sich weiter untergliedern in Voraussetzungen, die zu Innovationen verpflichten und solche, die Innovationen ermöglichen (ähnlich Strittmatter 2001).

Zu den **strukturellen Voraussetzungen**, die zur Entwicklung und Evaluation *verpflichten* (regulations), gehören u.a. rechtliche Verpflichtungen (z.B. zu einer Form von Evaluation und Entwicklung, zu deren Realisierung Leistungsdaten erforderlich erscheinen), aber auch interne Vereinbarungen und ungeschriebene „Gesetze“ bzw. Selbstverständlichkeiten, Erwartungshaltungen von Seiten gesellschaftlicher Gruppen oder Institutionen. Zu den strukturellen Voraussetzungen, die die Nutzung externer Leistungsdaten *ermöglichen und ihr gleichzeitig Grenzen setzen* (resources), gehören materielle und immaterielle Ressourcen (wie z.B. Zeitressourcen, organisatorischen Maßnahmen, finanzielle Ressourcen etc.).

Zu den **individuellen Voraussetzungen**, die zur Nutzung von Entwicklung und Evaluation *verpflichten* (values), gehört die professionelle Selbstverpflichtung durch entsprechende Werthaltungen und Motive, das Interesse, Bedrohungen auszuweichen (Überlebensmotiv) usw. Zu den individuellen Voraussetzungen, die sie *ermöglichen* (competences), gehören Wissen und Kompetenzen, die einzelne Akteur/innen befähigen, die strukturellen Vorgaben zu realisieren und die verfügbaren Ressourcen zu nutzen.

Wenn man von dieser Modellvorstellung ausgeht, können auf jeder dieser vier Ebenen hypothetisch Bedingungen für eine konstruktive Nutzung von Leistungsdaten lokalisiert werden. Die Modellvorstellung legt auch nahe, dass für feedbackgesteuerte Entwicklungsprozesse Einzelmaßnahmen nur wenig versprechen dürften, sondern ein Ensemble aufeinander abgestimmter Maßnahmen erforderlich ist. Für sämtliche Merkmale gibt es Hinweise unterschiedlicher Qualität aus der empirischen Forschung.

2.2.1 Strukturelle Voraussetzungen, die zu Entwicklung und Evaluation *verpflichten*

Auf Schulebene

- Verpflichtung zur kontinuierlichen Evaluation und Entwicklung und zur periodischen Dokumentation der Ergebnisse (z.B. Schulprogramm). Im Auftrag des Ministeriums wurde in Österreich ein Schulprogrammkonzept ausgearbeitet (Projektgruppe "Qualität im Schulwesen" 2002), das Entwicklung und Selbstevaluation in engen Zusammenhang bringt. Es besteht im Wesentlichen aus der Verpflichtung von Schulen (bzw. von (Fach)gruppen innerhalb der Schulen) in Abständen von ein bis zwei Jahren ihre Bemühungen um Qualität in einem evaluativen Rückblick auf die Vorhaben des vergangenen Zeitraums und einem planerischen Vorausblick auf die Vorhaben samt Erfolgskriterien für den nächsten Zeitraum darzustellen und in einem Dokument (dem Schulprogramm) darüber öffentlich Rechenschaft abzulegen (das Konzept wurde u.a. inspiriert vom Schweizer Konzept der förderlichen Qualitätsevaluation (FQS²) und den schottischen Initiativen zur Förderung von Selbstevaluation an Schulen (How good is our school)³. Über recht erfolgreiche Pilotversuche (Krainz-Dürr u.a. 2002) ist diese Initiative nicht hinausgekommen, da die rechtlichen Voraussetzungen für die Durchsetzung dieses Konzepts wegen des gewerkschaftlichen Widerstands noch nicht geschaffen worden sind. In einigen deutschen Ländern wurde eine Verpflichtung zum Schulprogramm zwar eingeführt allerdings mit einer wesentlich geringeren Betonung systematischer Selbstevaluation.

² Vgl. <http://www.lvb.ch/lvb/cms/upload/fqs/FQSProspekt1.pdf>

³ Vg. Z.B. <http://www.hmie.gov.uk/documents/publication/HMIE%20REPORT%20FINAL%202005C.pdf>

- Verpflichtung der Schulleiter/innen zur Stimulierung und Überprüfung von Entwicklung und Evaluation auf Unterrichtsebene. Interessant finde ich hier die Berichte über das KESS 4 Projekt in Hamburg, bei dem jeder Fachkoordinator veranlasst wurde, die Konsequenzen aus der Analyse der Ergebnisse in der Fachgruppe protokollarisch festzuhalten, die Schulleiter verpflichtet wurden, in Gesprächen mit den Fachgruppenleiter/innen erforderliche fachübergreifende Maßnahmen protokollarisch zu dokumentieren und die Schulaufsicht veranlasst wurde, im Gespräch mit den Schulleiter/innen getroffene Maßnahmen festzuhalten (pers. Mitt. N. Maritzen; vgl. auch Bos/Pietsch 2004).
- Hamburger Straße 31, 22083 Hamburg Generelle Verpflichtung der Schule, alle Schüler/innen zu den Zielen zu führen und bei Leistungsabfall sofort mit unterstützenden Maßnahmen zu reagieren. In Finnland erhielten im Jahre 2006 fast 22 % der Schüler/innen zusätzliche Unterstützung bei Leistungsschwächen, der Median in einem OECD Vergleich von 12 Ländern (2003) lag bei 6% (Laukkanen 2008, S. 313). In Finnland kann diese Unterstützung bis zu Einzelunterricht reichen. Eine gewisse Rolle dürfte dabei der Umstand spielen, dass finnische Pflichtschulen die schwächeren Schüler nicht loswerden können. In dieser Situation entsteht ein gewisser Druck, sich um die Kompetenzentwicklung *aller* Schüler/innen zu bemühen.
- ...

Auf regionaler Ebene:

- Verpflichtung der Schulaufsicht zur Metaevaluation der Selbstevaluation der Schule (z.B. ausgehend vom Schulprogramm). Dies wurde in Österreich bereits im Jahre 1999 als Aufgabe der Schulaufsicht definiert. Sie war allerdings mangels Verpflichtung der Schulen zum Schulprogramm nicht durchsetzbar. Einige Schulaufsichtsbeamte haben allerdings über Zielvereinbarungen mit Schulleiter/innen zumindest Ansätze davon realisiert (Vgl. Liebscher/Posch in Vorb.; vgl. auch die oben angeführten Verpflichtungen der Schulaufsicht im KESS-Projekt.)

2.2.2 Strukturelle Voraussetzungen, die Entwicklung und Evaluation ermöglichen

Auf Schulebene:

- Das Ausmaß an Eigenverantwortlichkeit der einzelnen Schule dürfte eine erhebliche Rolle für die Übernahme von Verantwortung für die Sorge um die Qualität der Arbeit sowohl der Schulleiter/innen als auch der Lehrer/innen spielen. Dazu sind Entscheidungsspielräume unabdingbar. Derartige Spielräume sind in Deutschland (wie auch in Österreich) im Vergleich mit anderen Ländern relativ klein (Arbeitsgruppe „Internationale Vergleichsstudie“ 2007, S. 89). Der Prozentsatz der Schülerinnen und Schüler in Schulen, bei denen die Schulen eine gewisse Verantwortung bei Aspekten der Schulpolitik und -verwaltung tragen (15-Jährige der PISA-Stichprobe) beträgt z.B. bei der Bestimmung des Lehrstoffs 35% gegenüber 91% in Finnland), bei der Entscheidung über Fächer-/Kursangebote 35% (gegenüber 95% in Finnland), bei der Lehrereinstellung 10% (gegenüber 35%) und bei der Festlegung des Schulbudgets 13% (gegenüber 56%).

Eine vermutlich ebenfalls unverzichtbare ermöglichende Struktur innerhalb der Schule sind professionelle Gemeinschaften (z.B. Fachgruppen) die sich um die kleinschrittige (!) kontinuierliche Entwicklung und Evaluation von Unterricht bemühen, sich dabei gegenseitig unterstützen, Erfahrungen austauschen und dabei soziale Verbindlichkeiten herstellen. Gemeinsames Reflektieren über die eigene Situation, die Entwicklung von Handlungsoptionen, gegenseitige Hospitationen, in denen der Besuchte die Agenda definiert, die gegenseitige Unterstützung bei der Überprüfung

der Resultate dürften auch erhebliche Bedeutung für die konstruktive Nutzung externer Leistungsdaten haben. Voraussetzung ist allerdings nach Wiliam (2006, S. 7), dass die Leiter/innen dieser professionellen Lerngemeinschaften die Kompetenzen und das Rollenverständnis haben, um Situationen zu gestalten, in denen Veränderungen möglich sind. Sie bilden damit ein an der Sicherung von Qualität orientiertes mittleres Management an der Schule.

Nach Wiliam haben professionelle Gemeinschaften mehrere Stärken:

- Sie widersprechen der beruflichen Isolation der Lehrer/innen
- Sie reprofessionalisieren Lehrer/innen durch Wertschätzung der Lehrerexpertise
- Sie deprivatisieren den Unterricht, indem Stärken und Herausforderungen bekannt werden
- Sie bieten eine kontinuierliche Unterstützung für Lehrer/innen
- Sie bieten professionelles Wachstum, indem regelmäßig Raum, Zeit und Struktur für systematische Reflexion von Praxis geboten wird
- Sie schaffen eine kollektive Wissensbasis an der Schule.

Auf regionaler Ebene:

- Wenn Schulen Verantwortung für kontinuierliche Weiterentwicklung übernehmen sollen, ist ein flexibles System regionaler Unterstützung unverzichtbar. In Österreich wurde im Vergleich zur Entwicklung von Standards und externen Tests bisher mehr Gewicht auf die externe Stimulierung und Unterstützung von Initiativen zur Weiterentwicklung von Unterricht gelegt. In diesem Rahmen hat sich vor allem das Projekt „IMST⁴ - Innovationen machen Schulen Top“ (Kraimer 2007) profiliert. Es verfolgt primär zwei Ziele: Innovationen an Schulen zu fördern bzw. für deren wissenschaftsgeleitete Analyse und Verbreitung zu sorgen und zweitens zum Aufbau eines nachhaltigen Unterstützungssystems beizutragen, das die eingeleitete Entwicklung strukturell absichern sollte. Im Rahmen dieses Projekts ist der Fonds für Unterrichts- und Schulentwicklung (IMST Fonds) entstanden, der Lehrer/innen finanziell und durch begleitende Fortbildung und Beratung dabei unterstützt, Innovationen zu untersuchen.
- Den professionellen Lerngemeinschaften auf Schulebene müssten auf regionaler Ebene Strukturen entsprechen, die die Nachfrage nach Unterstützung befriedigen können. In Österreich wird dazu ein Netzwerk sog. fachbezogener regionaler Bildungsmanager/innen eingerichtet, die derzeit ausgebildet werden (Posch 2008). Gegenüber den Schulen haben sie vor allem drei Aufgaben:
 - Beratung: Sie soll in enger Zusammenarbeit mit den Lehrer/innen der Anregung und Unterstützung von Entwicklungsprozessen und der Selbstevaluation von Unterricht dienen;
 - Fortbildung: Sie soll durch fachdidaktische Fortbildungsinitiativen unmittelbar an der Weiterentwicklung der Unterrichtsqualität orientiert sein;
 - Vernetzung: Sie soll zu ökonomischen und effizienten Formen des Austausches sowie zur fachdidaktischen und pädagogischen Weiterentwicklung der Praktiker/innen führen.

Die Wirksamkeit einer solchen Unterstützungsstruktur hängt allerdings in hohem Maße von der gezielten Nachfrage aus den Schulen ab und unterscheidet sich dadurch nicht wesentlich vom Angebot an Leistungsdaten großer

⁴ Ursprünglich: "Innovations in Mathematics Science and Technology Teaching". Seit der Ausweitung (vorerst) auf das Fach Deutsch wurde die Bezeichnung "Innovationen machen Schulen Top" gewählt, um das Akronym "IMST" beibehalten zu können.

Vergleichsuntersuchungen. Diese Nachfrage ist in hohem Maße von der Qualität der verpflichtenden und ermöglichenden Strukturen auf Schulebene abhängig.

- Von gewisser Bedeutung dürfte auch die Pflege von Schulnetzwerken sein. Fast alle Modelle der Qualitätsentwicklung von Schulen und Lehrenden arbeiten mittlerweile in irgendeiner Form in institutionalisierten Kooperations- oder Netzwerkstrukturen um professionellen Austausch, professionelles Lernen und die Dissemination von Innovationen zu fördern. Von diesen Netzwerken kann eine erhebliche Dynamik ausgehen, wie einige österreichische Erfahrungen zeigen (z.B. Lerman/Zehetmaier 2008). Allerdings erfordern sie eine permanente professionelle Netzwerkpflge, damit die Kommunikationsstrukturen langfristig aufrechterhalten werden können.

2.2.3 Individuelle Voraussetzungen, die zu Entwicklung und Evaluation verpflichten

Werthaltungen:

Werthaltungen haben zentrale Bedeutung. Die mangelnde Reflexion über die Qualität von Unterricht (und in der Folge die geringe Bereitschaft externe Leistungsdaten zu beachten) dürfte in hohem Maße mit tief verwurzelten Haltungen zusammen hängen. Wesentlich erscheinen u.a.

- Die Sorge um die Gestaltung der Rahmenbedingungen für das Lernen und der Kompetenzentwicklung der einzelnen Schüler/innen muss als bedeutsamer erachtet werden als das Angebot an Wissen. Derzeit ist es in vielen Fällen umgekehrt. Dem entspricht auch die verbreitete Tendenz zur Zuschreibung von Erfolg und Misserfolg an externe Faktoren (vor allem an die Schüler/innen) und die geringe Bereitschaft vieler Lehrer/innen, sich und die eigenen Bemühungen kritisch zu betrachten. Wer sich für ein unerfreuliches Ergebnis nicht verantwortlich fühlt, braucht auch keine Konsequenzen zu ziehen (Kohler 2004, S. 32).
- Das Bewusstsein einer professionellen Verpflichtung, *jedes* Kind auf den Weg durch die Ausbildung mitzunehmen.
Die individuelle Verpflichtung der Lehrer/innen, alle Schüler/innen zu denselben Zielen zu führen und bei Leistungsabfall sofort mit Unterstützungsmaßnahmen zu intervenieren und die Verhinderung der Möglichkeit von Schulen, schwächere Schüler/innen abzuschieben erhöht den Druck auf die Lehrer/innen, den Unterricht auf die unterschiedlichen Voraussetzungen der Lernenden abzustimmen. Dies scheint nach den Erfahrungen mehrerer Länder eine ganz zentrale Bedingung zu sein, um Lehrer/innen zu veranlassen, sich um die Weiterentwicklung von Unterricht zu bemühen und Nachfrage nach externer diagnostischer Unterstützung und nach (fach) didaktischer Unterstützung auszulösen (vgl. Laukkanen 2008, S. 313 und 320; McKinsey-Studie⁵). In Deutschland (wie auch in Österreich) fühlen sich Schüler/innen nur wenig von den Lehrer/innen unterstützt. Der Indexwert der Unterstützung beträgt für Deutschland -0,34, während der OECD Durchschnitt 0,2 beträgt (Arbeitsgruppe „Internationale Vergleichsstudie“ 2007, S. 74).
- Dazu kommen soziale Haltungen wie Orientierung an Stärken statt an Schwächen (Ermutigung statt Suche nach Schwächen beim Lernenden), Erfolgserwartung statt „Vermeidung von Misserfolg; Grundsätzlicher Respekt und Wertschätzung gegenüber den Schüler/innen.

2.2.4 Individuelle Voraussetzungen, die Entwicklung und Evaluation ermöglichen

⁵ Vgl. http://www.mckinsey.com/clientservice/socialsector/resources/pdf/Worlds_School_Systems_Final.pdf

Kompetenzen:

- Die Fähigkeit zur Einbeziehung von Schüler/innen (als Mitproduzent/innen schulischer Leistungen) in die Gestaltung und Evaluation (formative assessment von Unterricht)
- Diagnostische Kompetenzen und Erfahrung mit Methoden der systematischen Reflexion von Unterricht (Aktionsforschung)
- Fähigkeit zum Umgang mit Heterogenität. Derzeit dominiert eine Orientierung des Unterrichts an einem fiktiven Leistungsdurchschnitt in einer Klasse, der dazu führt, dass schwächere Schüler/innen überfordert und begabtere unterfordert werden. Dieses Unterrichtskonzept wird umso problematischer, je heterogener eine Klasse zusammengesetzt ist. In heterogenen Klassen ist der Erfolg des Unterrichts davon abhängig, inwieweit die unterschiedlichen Voraussetzungen der Schüler/innen bei der Gestaltung der Lernsituationen beachtet und genutzt werden. Dabei haben vor allem zwei Lehrerkompetenzen an Bedeutung gewonnen:
 - die Fähigkeit, die individuellen Lernvoraussetzungen von SchülerInnen zu diagnostizieren und bei der Gestaltung der Lernaufgaben zu berücksichtigen und
 - die Fähigkeit, die Zielorientierung des Unterrichts im Bewusstsein der SchülerInnen zu verankern, sodass sie selbst in der Lage sind, sich Ziele zu setzen und zu überprüfen, inwieweit diese jeweils erreicht sind.
- Kompetenz zur Interpretation externer Leistungsdaten.

Wegen der Herausforderungen, vor denen Lehrer/innen stehen (Heterogenisierung, wachsende Bedeutung der Partizipation der Schüler/innen bei der Gestaltung des Lehrens und Lernens, die zunehmende Bedeutung selbständigen Lernens, die Zuweisung von Erziehungsaufgaben etc.) sind tief greifende Änderungen der Unterrichtspraxis und des Lebens an der Schule erforderlich. Diese Ansprüche stehen in einem Spannungsverhältnis mit tief verankerten Werthaltungen, die zudem meist sehr früh, d.h. bereits in der eigenen Schulzeit entstanden sind und teilweise durch die Ausbildung noch verstärkt werden (Sozialisierung in eine Angst vor Verteidigungshaltung gegen Schüler/innen durch frühzeitigen Klassenkontakt). Es ist daher kaum zu erwarten, dass diese sich schnell ändern lassen.

Zugleich müssen LehrerInnen aber grundsätzlich als Professionelle angesehen werden, die wissen, wie sie das Lernen der Schüler/innen verbessern, weil der Unterrichtskontext viel zu komplex ist, um Handlungen einfach anordnen zu können. Lehrer/innen benötigen daher Vertrauensvorschuss und autonome Spielräume, damit sie die Verantwortung für Evaluation und Weiterentwicklung des Unterrichts übernehmen.

Daten über den Erfolg ihrer alltäglichen Handlungsroutinen lösen Irritationen aus, wenn sie ernst genommen werden. Um derartige Irritationen nicht nur zu akzeptieren, sondern im Interesse der Weiterentwicklung des Unterrichts aktiv aufzusuchen und konstruktiv zu verarbeiten, sind Selbstbewusstsein und ein Klima der Akzeptanz erforderlich (Laukkanen 2008, S. 319; Wiliam 2006, S. 7).

Wegen dieser Ansprüche haben Lehrerbildung und Lehrerfortbildung daher enorm an Bedeutung gewonnen. Wiliam (2006), hat dazu einige Schlüsselstrategien formuliert, die in der Lehrerbildung besondere Bedeutung erhalten sollten:

- Lernsituationen (Fragen, Aktivitäten, Aufgaben, Diskussionen) gestalten, die Belege über das Lernen der Schüler/innen liefern.
- Feedback geben, das in der Lage ist, Lernaktivitäten voranzubringen (formative feedback: Diagnosen und Anregungen zur Überwindung von Fehlern).

- Schüler/innen die Lernziele und die Erfolgskriterien deutlich machen.
- Schüler/innen als für das Lernen Verantwortliche aktivieren (d.h. auch: ihnen bewusst machen, dass Kompetenzen sich durch Lernen entwickeln).
- Schüler/innen als gegenseitige Lernressourcen aktivieren.

Zusammenfassung und Schluss

Zu den wichtigsten Maßnahmen, um die Schulen zu Entwicklung und Evaluation des Lehrens und Lernens und des Lebens an der Schule zu veranlassen und damit indirekt auch ihre Bereitschaft zu erhöhen, externe Leistungsdaten für die Verbesserung des Unterrichts zu verwerten sind vermutlich:

- Die Erweiterung autonomer Spielräume an den Schulen und zugleich die Verhinderung der Möglichkeit von Schulen, schwächere Schüler/innen abzuschieben.
- Die Verpflichtung von Schulen zu einer dokumentierten Entwicklung und Evaluation und zur Metaevaluation durch die Schulaufsicht.
- Die Institutionalisierung von „professional communities“, z.B. von Fachgruppen und Koordinator/innen mit einer klaren Aufgabe der kontinuierlichen Evaluation und Weiterentwicklung des Unterrichts.
- Auf regionaler Ebene die Einrichtung einer Unterstützungsstruktur, die vor allem in fachdidaktischer Hinsicht in der Lage ist, Beratungs- und Fortbildungsangebote bereit zu stellen. (vgl. dazu das Netzwerk sog. fachbezogener regionaler Bildungsmanager/innen in Österreich).
- Die Einrichtung von Schulnetzwerken, d.h. institutionalisierten Kooperations- oder Netzwerkstrukturen für den professionellen Austausch zwischen Lehrer/innen auf regionaler und überregionaler Ebene.

Die Abwertung systematischer Reflexion über den eigenen Unterricht und der dazu erforderlichen Methoden der Aktionsforschung dürfte kontraproduktiv sein, das sie eine Voraussetzung für die Nutzung externer Leistungsdaten und verfügbaren theoretischen Wissens im Allgemeinen darstellt (vgl. Altrichter/Posch 2007). Die Förderung einer Kultur der Evaluation und Entwicklung an der Schule kann ein wichtiger Garant dafür angenommen werden, dass externe Leistungsdaten als wertvolle Bereicherung der eigenen Bemühungen wahrgenommen werden.

Literatur

- Altrichter, H./Heinrich, M. (2007): Kategorien der Governance-Analyse und Transformationen der Systemsteuerung in Österreich. In: Altrichter, H./Brüsemeister, T./Wissinger, J. (Hrsg.): Educational Governance – Handlungskoordination und Steuerung im Bildungssystem. Wiesbaden: VS, S. 55–103.
- Altrichter, H./Posch, P. (2007): Lehrer erforschen ihren Unterricht – Einführung in die Methoden der Aktionsforschung“. Bad Heilbrunn: Klinkhardt.
- Arbeitsgruppe „Internationale Vergleichsstudie“ (2007): Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten. Berlin: Bundesministerium für Bildung und Forschung. (<http://www.bmbf.de/pub/pisa-vergleichsstudie.pdf>) (1.6.08).
- Bos, W./Pietsch, M. (2004): Erste Ergebnisse aus KESS 4 – Kurzbericht. Hamburg: Behörde für Bildung und Sport - Amt für Bildung.
- Elliott, J. (1993): What have we learned from Action Research in School-Based Evaluation? In: Educational Action Research, 1, Heft 1, S. 175-186.
- Giddens, A. (1992): Die Konstitution der Gesellschaft. Frankfurt/M.: Campus.
- Grabensberger, E./Freudenthaler, H./Specht, W. (2008): Bildungsstandards: Testungen und Ergebnisrückmeldungen auf der achten Schulstufe aus der Sicht der Praxis – Ergebnisse einer Befragung von Leiterinnen, Leitern und Lehrkräften der Pilotschulen. Graz: bifie.

- Helmke, A. (2001): TIMSS und die Folgen: Der weite Weg von der externen Leistungsevaluation zur Verbesserung des Lehrens und Lernens. In: Trier, U. P. (Hrsg.): Bildungswirksamkeit zwischen Forschung und Politik. Chur: Ruediger, S. 135-164.
- Helmke, A. (2004): Von der Evaluation zur Innovation: Pädagogische Nutzbarmachung von Vergleichsarbeiten in der Grundschule. in: Seminar, 2004, Heft 2, S. 90-112.
- Helmke, A./Schrader, F.-W. (2001): Von der Leistungsevaluation zur Unterrichtsentwicklung. In: Silbereisen, R./Reitzle, M. (Hrsg.): Psychologie 2000. Bericht über den 42. Kongress der Deutschen Gesellschaft für Psychologie in Jena. Lengerich: Pabst, S. 594-606.
- Hosenfeld, I./Groß Ophoff, J./Koch, U. (2007): Vergleichsarbeiten in Klassenstufe 3 („VERA 3“) – Konzept und empirische Befunde zum Umgang mit den Ergebnisrückmeldungen in den Schulen. Landau: Universität Koblenz-Landau. (Präsentation bei der 7. EMSE Tagung in Mainz).
- Kohler, B. (2004): Zur Rezeption externer Evaluation durch Lehrkräfte, Eltern sowie Beamte der Schulaufsicht. In: Empirische Pädagogik, 18, Heft 1, S. 18-39.
- Kohler, B./Schrader F.-W. (2004): Ergebnisrückmeldung und Rezeption: Von der externen Evaluation zur Entwicklung von Schule und Unterricht. In: Empirische Pädagogik, 18, Heft 1, S.3-17.
- Krainer, K. (2008): IMST als Intervention in das österreichische Bildungssystem. In: Journal für Schulentwicklung, 11, Heft 4, S. 36-42.
- Krainz-Dürr, M./Posch, P./Rauch, F. (2002): Schulprogramme entwickeln. Erfahrungen aus einem Pilotprojekt an berufsbildenden Schulen. Innsbruck: Studienverlag.
- Laukkanen, R. (2008): Finnish Strategy for High-Level Education for all. In: Soguel, N.C./Jaccard, P. (eds.): Governance and Performance of Education Systems. New York: Springer, S. 305-324.
- Lerman, S./Zehetmeier, S. (in print): Face-to-face communities and networks of practising mathematics teachers. In: Krainer, K./Wood, T. (eds.): International handbook of mathematics teacher education: Vol. 3, Participants in mathematics teacher education: Individuals, teams, communities and networks. Rotterdam. The Netherlands: Sense Publishers.
- Leutner, D./Fleischer, J./Spoden, C./Wirth, J. (2007): Schulrückmeldungen in landesweiten Lernstandserhebungen – Das Beispiel lernstand 8 in NRW. Präsentation bei der 7. Tagung „Empiriegestützte Schulentwicklung“ in Mainz, 6./7.Dez. 2007.
- Liebscher, M./Posch, P. (in Vorb.): Gesellschaftliche Herausforderungen für das Schulwesen - Antworten von Schulen und Schulverwaltung. In: Krainer, K./Hanfstingl, B./Zehetmeier, S. (Hrsg.): IMST-Buchreihe, Band 4. Innsbruck: Studienverlag.
- Maier, U. (2007): Lehrereinschätzungen zu zentralen Tests und Leistungsrückmeldungen – Ein Vergleich zwischen Baden-Württemberg und Thüringen. Kurzvorstellung erster Ergebnisse auf der 7. Tagung „Empiriegestützte Schulentwicklung“ in Mainz, 6./7.Dez. 2007.
- Peek, R. (2004): Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik (QuaSUM) – Klassenbezogene Ergebnisrückmeldungen und ihre Rezeption in Brandenburger Schulen. In: Empirische Pädagogik, 18, Heft 1, S. 82-114.
- Peek, R./Steffens, U./Köller, O. (2007): Zentrale standardisierte Lernstandserhebungen – Positionspapier des Netzwerks „Empiriegestützte Schulentwicklung (EMSE)“. o.O.
- Posch P. (2008): Aufgaben eines fachbezogenen Bildungsmanagements. Was können Schulen von fachbezogenen Bildungsmanager/innen erwarten? In: IMST Newsletter, Nr. 26 „Fachbezogenes Bildungsmanagement“.
- Projektgruppe "Qualität im Schulwesen" (2002): Vorschläge zur Einführung des Schulprogramms an den Österreichischen Schulen. Wien: Bundesministerium für Bildung, Wissenschaft und Kultur.

- Rauch, F./Scherz, H. (in Vorb.). Regionale Netzwerke: Konzept, Erfahrungen und Chancen. In: Krainer, K./Hanfstingl, B./Zehetmeier, S. (Hrsg.). IMST-Buchreihe, Band 4. Innsbruck: Studienverlag.
- Schneewind, J./Kuper, H. (2008): Rückmeldeformate und Verwendungsmöglichkeiten der Ergebnisse aus zentralen Lernstandserhebungen. In: Dokumentation zum Symposium „Lernen aus Evaluationsergebnissen - Verbesserungen planen und implementieren“ im Rahmen der didacta - die Bildungsmesse vom 20. - 21.02.2008 in Stuttgart, S. 18-29.
- Steffens, U. (2004): Die empirische Wende in der Bildungspolitik und Bildungsplanung – Einleitung zur bundesweiten Fachtagung „Empiriegestützte Schulentwicklung“ am 15.-16. Dezember 2004 in Soest. Ms.
- Strittmatter, A. (2001): Bedingungen für die nachhaltige Aufnahme von Neuerungen an Schulen. In: journal für schulentwicklung, 2001, Heft 4, S. 58 – 66.
- William, D. (2006) Does Assessment hinder Learning? URL: http://www.mission-21.com/ec/images/williams_speech.pdf, Zugriffsdatum:1.6.08.
- Van Ackeren, I. (2007): Nutzung großflächiger Tests für die Schulentwicklung: Exemplarische Analyse der Erfahrungen aus England, Frankreich und den Niederlanden. Bildungsforschung: Bd. 3. Bonn, Berlin: Bundesministerium für Bildung und Forschung (BMBF).