

---

## Vergleichende Lernstandsuntersuchungen, Bildungsstandards und die Steuerung von schulischen Bildungsprozessen

---

Die Einführung von Bildungsstandards wird sowohl in Schulen als auch bei der Bildungsadministration nicht nur positiv bewertet. Bedenken werden verständlich, wenn eine Standardisierung eine Einengung des pädagogischen Handelns nach sich ziehen würde. Dass die Einführung von Bildungsstandards aber auch Chancen für das deutsche Bildungssystem birgt, wird im nachfolgenden Beitrag skizziert. Vor dem Hintergrund der Erfahrungen mit empirischen Studien in Deutschland wird der Blick auf Länder gerichtet, die in internationalen Vergleichsuntersuchungen erfolgreicher als Deutschland abgeschnitten haben, um hier einige Merkmale in Hinblick auf Evaluation und Bildungsstandards vergleichend heraus zu arbeiten. Im Abschluss des Beitrags werden dann mögliche Perspektiven für die Entwicklung des deutschen Bildungssystems beschrieben. Hierbei wird deutlich gemacht, dass eine Erfolg versprechende Implementierung von Bildungsstandards nicht nur die Lehrkräfte sondern konsequenterweise auch die Ausstattung und Organisation der Bildungseinrichtungen tangiert und somit auch die Bildungsadministration mit einbezieht.

### 1 Schattenseiten des deutschen Bildungssystems im Lichte von TIMSS und PISA

Mitte der 1990er Jahre erschütterte ein erster „Schock“ die deutsche pädagogische Fachöffentlichkeit. In der *Third International Mathematics and Science Study* (TIMSS) wurden den deutschen Siebt- und Achtklässlern allenfalls mittelmäßige Leistungen bei mathematischen und naturwissenschaftlichen Kenntnissen bescheinigt (BAUMERT et al., 1997). Den ebenfalls getesteten Jugendlichen, die das deutsche Regelschulsystem demnächst verlassen sollten, konnte kein besseres Zeugnis ausgestellt werden (BAUMERT, BOS & LEHMANN, 2000a, 2000b). Beunruhigt durch diese Befunde entschied man sich auch an den folgenden internationalen Schulvergleichsstudien teil zu nehmen. In der im Jahre 2000 durchgeführten PISA-Studie (*Programme for International Student Assessment*) erreichten die dort untersuchten fünfzehnjährigen Schülerinnen und Schüler in den gestesteten Kompetenzen Lesen, Mathematik und Naturwissenschaften nur ein schwaches mittleres Niveau. Die Ergebnisse der Studie wurden von einer breiten Öffentlichkeit – oft verkürzt und verfälscht – als höchst alarmierend wahrgenommen. In der öffentlichen Diskussion über Bildungsfragen war der „PISA-Schock“ fortan ein wesentlicher Faktor. Doch auch aus wissenschaftlicher Sicht erscheinen die Befunde aktueller Schulleistungsstudien nicht weniger dramatisch. Im Gegenteil: Verlässt man das Niveau des einfachen Länderrankings und misst die Leistungen der Schülerinnen und Schüler bzw. des Bildungssystem mit begründeten Leistungsnormen, wird das Ausmaß der Misere noch sichtbarer (vgl.: BAUMERT et al., 2001). Die PISA 2000 Studie zeigt, dass ca. 20 Prozent der fünfzehnjährigen Jugendlichen nicht über eine Lesekompetenz verfügen, die ausreicht, um ihnen ein selbständiges und selbstbestimmtes Leben in

unserer Gesellschaft zu ermöglichen. Gut vierzig Prozent der Hauptschüler dieser Altersgruppe beherrschen die Mathematik lediglich auf dem Grundschulniveau. Und schließlich verfügt noch nicht einmal die Hälfte der Jugendlichen über Kenntnisse in den Naturwissenschaften, welche über dem normalen Alltagsverständnis liegen. Selbst vor dem Hintergrund, dass es Industrienationen gibt, die noch schlechtere Befunde aufweisen, ist es doch eine Frage der normativen Setzung welchen Anteil von erwartungswidrig schlechten Kenntnissen man in einer Gesellschaft akzeptieren möchte. Aus meiner Sicht scheint es auf keinen Fall akzeptabel nicht weniger als ein Fünftel(!) aller Hauptschüler aus dem Schulsystem mit der Perspektive zu entlassen, später zu funktionalen Analphabeten zu werden.

Wie konnte es zu diesen desaströsen Untersuchungsergebnissen kommen? Zwei Erklärungsansätze sollen im Folgenden kurz dargestellt werden. Zum einen ist es die selbstverschuldete Ahnungslosigkeit über den Zustand des Bildungssystems und zum anderen ist es das falsche Vertrauen auf sich selbst steuernde Bildungsprozesse. Am Beispiel von Ländern, die in PISA-Studien erfolgreicher als Deutschland abgeschnitten haben, können wir Möglichkeiten für die Weiterentwicklung des deutschen Bildungssystems erkennen. Dort zeigt sich, dass die Formulierung von nationalen Bildungsstandards ein Instrument zur effektiven Steuerung von schulischen Bildungsprozessen sein kann. Welche Konsequenzen eine Orientierung an Bildungsstandards für die Wirksamkeit von Schule hat, aber auch welche Perspektive sich für die Ausbildung von Lehrkräften hieraus ergibt, wird zu diskutieren sein.

### **1.1 Das Defizit an gesichertem Wissen über Leistungsstände von Schülerinnen und Schülern**

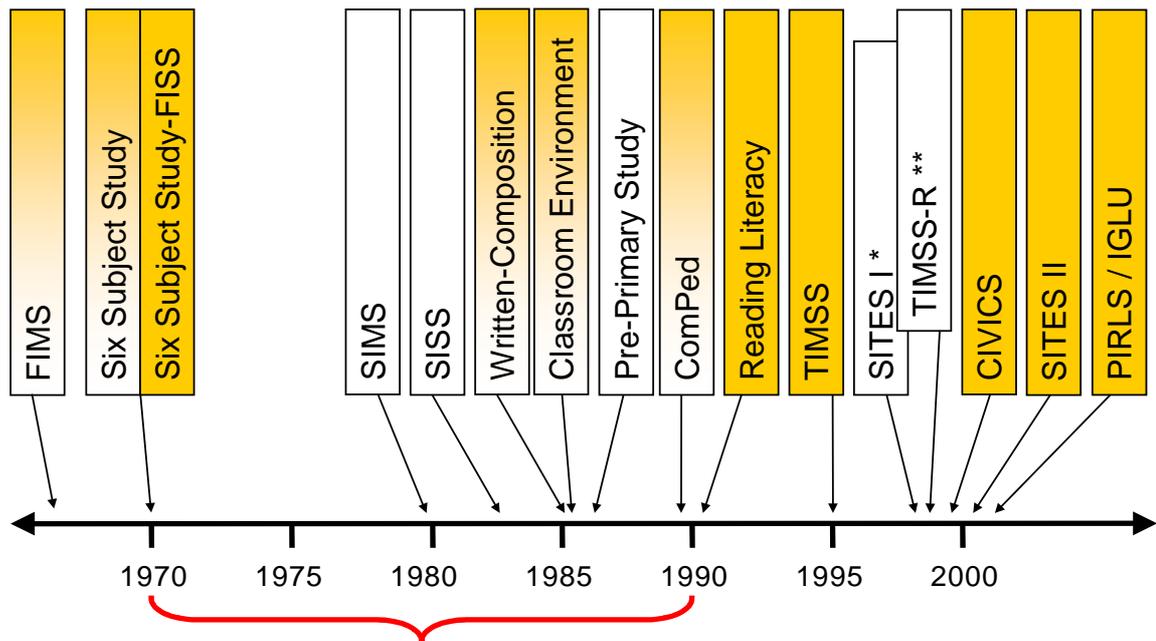
Die deutsche Erziehungswissenschaft ist viele Jahre durch eine geisteswissenschaftliche Tradition geprägt gewesen, in der den Schulen vor allem die Aufgabe gestellt wurde, Bildungsprozesse zu fördern. „Bildung“ umfasste in diesem Verständnis weit mehr als lediglich reproduzierbare Kenntnisse. Im emphatischen Begriff der „Menschenbildung“ klingt der Anspruch an, junge Menschen umfassend, in allen ihren Fähigkeiten zu fördern, so dass sie ein individuell bestimmtes und gesellschaftlich verantwortetes Leben führen können. In dieser traditionellen Denkweise festigte sich das Bewusstsein, dass man das Niveau der „Bildung“, die eine Person erreicht hatte, nicht messen könne. In der Tat entzieht sich dieser Bildungsbegriff weitgehend empirischen Messstrategien, aber er war ein hervorragend geeigneter Gegenstand für philosophische Betrachtungen. Eine übersteigerte Skepsis gegenüber empirischen Untersuchungen von Lernprozessen überhaupt führte zu weitgehender Abstinenz von international verbreiteten empirischen Vorgehensweisen.

Das lässt sich exemplarisch an der westdeutschen Bildungsreformdebatte in den 1970er und 1980er Jahre zeigen. „Bildung für Alle“ oder „gleiche Bildungschancen für Alle“ waren zentrale bildungspolitische Ziele geworden. Man hatte nach PREISERT (1967) das „katholische Arbeitermädchen vom Lande“ als Synonym für im Bildungssystem mehrfach Benachteiligte identifiziert. In dieser Zeit haben sich sowohl einige Politiker als auch verschiedene Wissenschaftler für die Einführung einer Gesamtschule eingesetzt da von dieser Schulform erwartet wurde, dass sie soziale Disparitäten in Bezug auf Bildung kompensieren könnte.

Aus heutiger Sicht ist bemerkenswert, dass viele der in den vergangenen Dekaden durchgeführten Reformen ohne empirische Überprüfung implementiert und tradiert wurden – auch wenn zum Teil Hinweise auf unerwartete, wenn nicht sogar gegenteilige Effekte als die erhofften beobachtet wurden. Wie kam es zu dieser Forschungslücke? In den frühen 1970er Jahren nahmen deutsche Schülerinnen und Schüler an der ersten internationalen Naturwissenschaftsstudie (*First International Science Study*, FISS) teil. Die Befunde dieser Untersuchung stimmten in der Summe mit denen der aktuellen PISA Studie überein: Deutsche Schülerinnen und Schüler erzielten im internationalen Vergleich unterdurchschnittliche Ergebnisse! Zwar hat sich die deutsche Erziehungswissenschaft inhaltlich mit den Befunden auseinandergesetzt, aber sie versäumte es, sich im folgenden an internationalen Vergleichsstudien zu beteiligen.

Die Abbildung 1 gibt einen Eindruck über die Teilnahme Deutschlands an von der *International Association for the Evaluation of Educational Achievement* (IEA) organisierten Systemmonitoring-Studien. Die durchgängig farblich hinterlegten Studien kennzeichnen eine umfassende Beteiligung mit einer für Deutschland repräsentativen Stichprobe. An den nur zum Teil schattierten Studien haben nur Teilgruppen aus Deutschland teilgenommen. Die übrigen Studien wurden ganz ohne deutsche Beteiligung durchgeführt. Wie deutlich hervorgehoben ist, hat es eine Teilnahmelücke von rund 20 Jahren in Deutschland gegeben – 20 Jahre, in denen kaum jemand hinterfragt hat, auf welchem Niveau sich das Deutsche Bildungssystem bei der Vermittlung von zentralen Kulturtechniken und anderen Kompetenzen zur Bewältigung der Anforderungen moderner Gesellschaften befand. In der Tradition intensiver erziehungswissenschaftlicher und bildungspolitischer Diskussionen wurde davon ausgegangen, dass ein internationaler Vergleich – wenn er überhaupt sinnvoll gestaltet werden könne – keine Defizite aufzuzeigen in der Lage sei. Erst als 1991 Rainer H. LEHMANN mit der Tradition der Abstinenz von internationalen Vergleichsuntersuchungen brach, gab es erste Hinweise auf Mängel im Leistungsniveau deutscher Schüler, die jedoch von einer breiten (Fach-)Öffentlichkeit nicht wahrgenommen wurde (LEHMANN, PEEK, PIEPER & STRITZKY, 1991). Trotz der nach wie vor verbreiteten Auffassung, „Bildung“ sei nicht messbar, entschieden sich einige Bildungsforscher, die offensichtlichen Forschungsdesiderate aufzugreifen und in den folgenden Jahren regelmäßig an den internationalen Schulvergleichsstudien teilzunehmen. Aber es brauchte gut eine halbe Dekade bis schließlich der TIMSS-Schock ein Umdenken in der forschenden Erziehungswissenschaft initiierte und die Perspektive der empirischen Überprüfung die Wahrnehmung des deutschen Bildungssystems neu strukturierte.

Aber nicht nur durch die Teilnahme an internationalen Schulvergleichsuntersuchungen zeichnet sich eine reale empirische Wende in der deutschen Bildungsforschung ab. Eine ganze Reihe von bundeslandspezifischen Untersuchungen wurde seit Mitte der 1990er Jahre durchgeführt. Untersuchungen, die nachhaltig auf Defizite der jeweiligen Bildungssysteme hinweisen, haben empirische Evidenz für Fehlentscheidungen und Missstände in die aktuellen Diskussionen einbringen können.



## 20 Jahre ohne systematische empirische internationale Forschung

- \* Deutschland nahm an dem ersten Modul der SITE-Studie nicht teil, da zur gleichen Zeit bereits das inhaltsgleiche Projekt „Schulen ans Netz“ etabliert war.
- \*\* Auf die Teilnahme an der TIMSS-R (TIMSS 1999) Studie verzichtete Deutschland zugunsten der Teilnahme an der PISA-Untersuchung.

Abb. 1: Deutsche Beteiligung an internationalen Vergleichsstudien der *International Association for the Evaluation of Educational Research* (IEA) bis 2001.

Das überzeugendste Argument für die Etablierung von systematischen Untersuchungen im Bildungssystem lässt sich seit kurzem an der Hamburger Studie „Kompetenzen und Einstellungen von Schülerinnen und Schülern“ (KESS) belegen. Diese in der Tradition der Hamburger Untersuchung zur Lernausgangslage (LAU), die in Deutschland die erste flächendeckende Untersuchung eines Bildungssystems im Längsschnitt war, haben sich aufgrund der konsequenten Rückmeldung und Auseinandersetzung mit den Befunden der Studie in Hamburg die Kompetenzen der Viertklässler in allen erfassten Bereichen (Lesen, Mathematik und Naturwissenschaften) signifikant verbessert!

Dennoch erschwert in Deutschland nach wie vor ein evaluationshemmendes Klima die konstruktive Nutzung von Leistungsstudien. So besteht bei vielen Lehrkräften ein latentes Unbehagen gegen eine Evaluation des eigenen Unterrichts. Die Chance, sich auf der Basis empirisch gewonnener und ausgewerteter Informationen selber weiter zu entwickeln und eigene Stärken oder Schwächen auszuloten, wird hierzulande nicht nur auf individueller Ebene nicht erkannt und folglich auch nicht genutzt.

Offenbar bedarf es eines längeren Lernprozesses, um Evaluationen primär als Chance zur Weiterentwicklung für die eigene pädagogische Arbeit zu begreifen. Nur wenn die Schulver-

waltung glaubhaft macht, dass sie ebenfalls die Evaluation nur in einem fördernden Sinne, nicht aber als Ausgangspunkt für Kontrollen und Sanktionen nutzen wird, kann man von Lehrkräften eine positive Einstellungsänderung erwarten. Länder wie die Niederlande oder Kanada hingegen nutzen die verschiedenen Untersuchungen, um Reformen auf ihre Zielerreichung hin zu überprüfen. Hier werden Reformen nicht nur auf Basis guten Glaubens durchgeführt, sondern ihr Ertrag anhand der Lernstände der Schülerinnen und Schüler empirisch überprüft und weiterentwickelt. Für Deutschland besteht neben der Entscheidung für die Organisation einer neuen Evaluationskultur auch die Notwendigkeit zu überlegen, wie weit das Schulsystem reformiert werden soll, um bessere Lernleistungen zu erzielen

## 1.2 Wandel in der Systemsteuerung

Die verstärkte Partizipation an internationalen Vergleichsstudien hat auch zu einem Paradigmenwechsel in der Steuerung von Bildungsprozessen beigetragen. Wie in Abbildung 2 dargestellt, lassen sich im Wesentlichen drei verschiedene Steuerungsmodelle im Bildungssystem unterscheiden, die in der Chronologie des Wandels skizziert werden sollen: Zunächst ist man implizit davon ausgegangen, dass der Lern- oder umfassender formuliert der Bildungserfolg, vornehmlich durch ein ausreichend gutes Angebot von Lerngelegenheiten determiniert ist. Realisiert hat man diese *Inputsteuerung* durch eine – bis heute sichtbare – gute Ausbildung von Lehrkräften und eine umfassende Ausstattung der Bildungseinrichtungen. Ausgangspunkt für diese Überlegung war die Auffassung, dass durch sowohl fachwissenschaftlich als auch erziehungswissenschaftlich professionalisierte Lehrende der Unterrichtsgegenstand gleichsam automatisch bei den Bildungsempfängern „ankommt“. und wenn nicht, lag es nicht an dem guten Angebot, sondern an dem Unvermögen der Bildungsempfänger, die gebotenen Lerngelegenheiten für sich zu nutzen.

Diese „Gießkannenpädagogik“ wurde vor dem Hintergrund wahrgenommener differenzierter Rezeption von Lerngelegenheiten weiterentwickelt. Im Sinne der *Prozesssteuerung* stand jetzt die Unterrichtsqualität im Zentrum der Betrachtung. Kriterien wie „Klarheit des Unterrichts“, „effektive Zeitnutzung“, „Anspruchsniveau“, „Übungsintensität“ oder „Instruktionstempo“ richteten den Blick von den lehrenden Akteuren auf ein optimiertes Lehrangebot (vgl.: EINSIEDLER, 1997). Als Weiterentwicklung der *Inputsteuerung* wurde bei der *Prozesssteuerung* somit nicht mehr nur die Qualität des Angebotes an Lerngelegenheiten sondern auch die Qualität der Vermittlung selber in den Blick genommen. Diese Modifikation kann jedoch nicht verbergen, dass es nach wie vor ein Angebotsmodell bleibt, bei dem der tatsächliche Lernerfolg weiter den individuellen Lernenden zugeschrieben wird. Ob die optimierten Lernangebote tatsächlich aufgenommen wurden, wurde nicht systematisch empirisch untersucht. Die Verantwortung für das Lernergebnis lag nach wie vor voll bei den Lernenden selber.

Erst der Wandel zur Outputsteuerung ließ die konkrete Frage zu „Was kommt tatsächlich bei den Schülerinnen und Schülern an?“ Bereits hier war der Blick „durch PISA“ geschärft. So hat man – bei der ersten Ursachenanalyse – schnell den Blick auf erfolgreiche Länder gerichtet und festgestellt, dass hier eine regelmäßige Kontrolle des tatsächlichen Lernstandes mit

zeitlichem Vorlauf schließlich auch eine individualisierte Optimierung des Lernprozesses möglich macht. Wie in Abbildung 2 dargestellt, lässt erst die empirisch fundierte Diagnose von Kompetenzen, Fähigkeiten und Fertigkeiten der Schülerinnen und Schüler eine wirksame Modifizierung der Lerngelegenheiten zu. Diese *Outputsteuerung* ist jedoch nur dann sinnvoll, wenn rechtzeitig vor Abschlüssen auf den Lernstand geschaut wird, um noch intervenieren und somit die Schülerinnen und Schüler auf optimierten Wegen möglichst zum Erfolg führen zu können. Eine Überprüfung des Lernstandes im Rahmen zentraler Abschlussprüfungen entspricht nicht der Idee dieses Steuerungsprinzips. Nur wenn Raum für Korrekturen bleibt, ist der Blick auf den individuellen Lernerfolg zielführend.

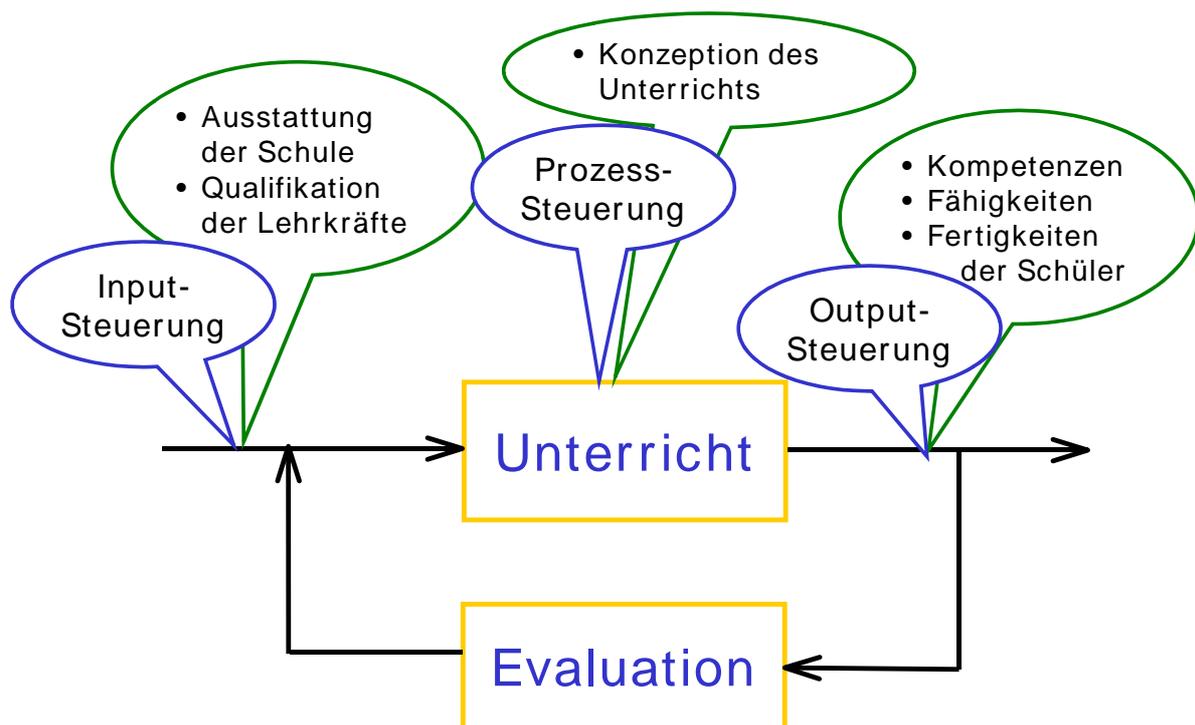


Abb. 2: Exemplarische Darstellung von Steuerungsmodellen am Beispiel von Unterricht.

## 2 Ein Blick über Ländergrenzen hinweg

Ein Vergleich mit anderen Ländern im Rahmen von groß angelegten Schulvergleichsuntersuchungen zeigt nicht nur die Position im Länderranking sondern bietet darüber hinaus auch die Möglichkeit, von erfolgreicheren Bildungssystemen zu lernen. Vor diesem Hintergrund ist im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF) ein vertiefender Vergleich ausgewählter PISA-Staaten in Auftrag gegeben worden, in dem über die bereits vorliegenden quantitativen empirischen Daten hinaus ein qualitativer Vergleich ergänzend vorgenommen wurde (ARBEITSGRUPPE „INTERNATIONALE VERGLEICHsstudie“, 2003). Hierzu wurden Länder ausgewählt, die einen vergleichbaren sozio-ökonomischen Status wie Deutschland aufweisen, in der Regel besser in PISA abgeschnitten haben und

bereits über Erfahrungen mit Systemmonitoring-Studien verfügen. Für die Studie wurden Kanada, England, Finnland, Frankreich, die Niederlande und Schweden ausgewählt. In diesen Ländern wurden Bildungsexperten darum gebeten, einen Katalog von Fragen in Form von Gutachten zu beantworten (DÖBERT, KLIEME & SROKA, 2004). Der Fragenkatalog wurde von einem mit deutschen Wissenschaftlern besetzten Konsortium in Kooperation mit internationalen Partnern erarbeitet. Ein Aspekt der Gutachten war den Zielen der Bildungssysteme gewidmet, wobei aus deutscher Sicht auch der Umgang mit Bildungsstandards von Interesse war. Die internationalen Kooperationspartner sollten deshalb u. a. folgende Fragen beantworten:

*Standards (key words: national understanding of 'standard', dealing with standards):*

- How are educational standards (especially, minimum standards) defined and used in the curriculum, in educational research and in evaluation?
- How are indicators for educational standards at different levels and for different targets of education operationalized (degree of details)?

Im folgenden werden einige zentrale Aspekte dieser Vergleichsstudie vorgestellt.

## **2.1 Sind „Standards“ nicht gleich „Standards“?**

Das deutsche Konsortium ging zunächst davon aus, dass in den Ländern Standards „normativ und verbindlich festlegen, welche Leistungen Schülerinnen und Schüler bis zum Ende einer bestimmen Bildungsstufe erreichen sollten“ (ARBEITSGRUPPE „INTERNATIONALE VERGLEICHsstudie“, 2003, 164). Nach der Durchsicht der Gutachten stellte sich aber bald heraus, dass in den verschiedenen Ländern dieser Begriff uneinheitlich für verschiedene Aspekte des Bildungssystems benutzt wurde. So wurden Standards insbesondere im Zusammenhang mit Lehrplänen bzw. Schulleistungsmessungen genannt.

Im folgenden wird kursorisch das jeweils in den Länder-Gutachten dargelegte Verständnis von „Standards“ vorgestellt.

- Kanada: Im Verlauf einer längeren Fachdiskussion in Kanada bzw. im nordamerikanischen Raum hat sich ein uneinheitliches Verständnis von Standards entwickelt, welches sich grob in drei Verwendungen darstellen lässt: (1) Standards als Lernergebnisse; (2) Standard als akzeptiertes (minimal) Niveau der Leistung von Schülerinnen und Schülern; (3) Standards als differenzierte Niveaustufen. Jeder dieser drei Auffassungen liegt zu Grunde, dass die Lernergebnisse anhand von standardisierten Tests empirisch erfasst und ausgewertet werden können.
- England: Der Begriff „Standard“ ist im aktiven Wortschatz von Beteiligten des Bildungssystems etabliert, wird aber nicht als „normative Setzung“ (167) verstanden sondern stellt vielmehr einen Alltagsbegriff dar, der tatsächlich beobachtete Leistungsstände beschreibt. Eine englische Besonderheit ist das ausgreifende System der „Standard“-Überprüfung durch das staatliche „Office for Standards in Education“ (OFSTED). Die

von dort angestrebte Anhebung von Standards wird als eine Erhöhung der tatsächlich erreichten Lernstände der Schülerinnen und Schüler verstanden. Im Sinne von Kompetenz- oder Leistungserwartungen werden Standards in den Curricula manifestiert. Diese sollen als Vergleichsmaßstab bei der (empirischen) Überprüfung von Leistungsentwicklungen der Schülerinnen und Schüler herangezogen werden. Aus deutscher Sicht ist bemerkenswert, dass Erwartungen in Form von Standards nicht nur als Lernstände der Schülerinnen und Schüler sondern darüber hinaus auch als Merkmale guten Unterrichts von OFSTED beschrieben werden.

- Finnland: Aufgrund einer ganz anderen Tradition wird im tendenziell auf Individualisierung ausgerichteten finnischen Bildungssystem der Begriff „Standard“ vermieden, weil hiermit eher einheitliche Testinhalte assoziiert werden. Dennoch werden in den Lehrplänen konkrete Lernerwartungen beschrieben, die Kinder bzw. Jugendliche verbindlich erfüllen müssen. Hierbei liegt es in der Verantwortung sowohl der Schulen als auch der Schülerinnen und Schüler und – hier wird der hohe gesellschaftliche Stellenwert von Bildung deutlich – auch deren Eltern, die verbindlichen Mindestwartungen zu erreichen. Das Amt für Bildung beschränkt sich darauf, Empfehlungen zur Beurteilung von Wissen und Können vor der Folie guter Leistungen auszusprechen. An diesen Empfehlungen orientieren sich auch stichprobenbasierte nationale Leistungsüberprüfungen an denen Schulen zum einen per Zufallsauswahl verbindlich teil nehmen müssen, nicht ausgewählte Schulen aber auf Eigeninitiative teilnehmen können – wovon auch tatsächlich viele Gebrauch machen.
- Frankreich: Auch wenn es für ein so stark zentralistisch organisiertes Bildungssystem, wie es Frankreich aufweist, nahe liegen würde, hat sich hier ein einheitlicher Gebrauch von Bildungsstandards bisher nicht etabliert. Die Qualität von Lernergebnissen wird zwar regelmäßig anhand nationaler Vergleichsstudien kontrolliert, jedoch hat sich beim Vergleich der Ergebnisse keine kriteriale Bezugsnorm (Orientierung z.B. an Bildungsstandards) sondern eine soziale Bezugsnorm etabliert, bei der die Schulen bzw. Schülergruppen im relativen Verhältnis untereinander verglichen werden. Diese Form des Vergleichs ist überraschend, da die verbindlichen nationalen Lehrpläne und zentralen Abschlussprüfungen eher eine kriteriumsorientierte Überprüfung erwarten ließe. Ein Wandel der Zielfestschreibungen im Bildungssystem zeichnet sich in Frankreich derzeit ab: Während früher insbesondere Lerninhalte in den Curricula beschrieben wurden, werden jetzt zunehmend auch Fähigkeiten und Fertigkeiten mit benannt. (Was – wie auch in Deutschland – auf einen Wandel von der *Prozess-* zur *Outputsteuerung* des Bildungssystems hindeutet.)
- Niederlande: Wie auch in Finnland wird in den Niederlanden die pädagogische Freiheit der Lehrenden als zentrales Merkmal des Schulsystems angesehen. Vor diesem Hintergrund entspricht es den Erwartungen, wenn hier bisher keine (Minimal-)Standards definiert wurden. Dieser Befund überrascht jedoch in Anbetracht der bekannten niederländischen Schulinspektionen und des etablierten Systems zentraler, standardisierter Leistungsüberprüfungen. Bisher festgeschrieben sind allgemeine Zielvorgaben in Bezug auf

Kenntnisse sowie fachspezifische und fächerübergreifende Fähigkeiten und Einstellungen, wobei den Schulen ein Freiraum bei der Ausgestaltung des Unterrichts bleibt. Bedeutsam wird das Erreichen der Zielvorgaben, wenn bei zentralen Abschlussprüfungen das Niveau der Ergebnisse durch die Schulinspektion evaluiert wird. Aber auch Primarstufen stehen verschiedene Tests zur Selbstvergewisserung in den Schulen zur Verfügung, die auch Fächer wie Sport und Musik einbeziehen. Der Begriff Standards ist in den Niederlanden bei diesen Leistungsmessungen nicht gebräuchlich, wohl aber – und auch dies ist wiederum aus deutscher Sicht interessant – in Bezug auf die Kompetenzen von Lehrkräften und die Schulqualität. Diese Standards zielen somit auf Aspekte der *Input*- und auch der *Prozesssteuerung*. Wobei die Überprüfung der Lernergebnisse (*Output*) selbstverständlich auch berücksichtigt wird. Somit wird in den Niederlanden auf ein umfassendes Qualitätsmanagement Wert gelegt.

- Schweden: Auch im schwedischen Bildungssystem wird viel Rücksicht auf die Unterstützung individueller Entwicklung der Schüler genommen. Es hängt damit zusammen, dass auch hier keine expliziten Bildungsstandards definiert werden, obwohl solche in Form zentraler Tests inhaltlich *de facto* vorliegen. Darüber hinaus werden „Ziele“ definiert, die Schülerinnen und Schüler minimal erreicht haben sollen, wenn sie eine Bildungseinrichtung verlassen. Dass im schwedischen Bildungsverständnis in der Kategorie „Ziel“ anstatt von „Standard“ gedacht wird, spiegelt sich auch darin wider, dass Schüler in Tests nicht durchfallen können sondern, dass sie nur „noch nicht so weit sind“ (DAUN, SLENNING & WALDOW, 2004, 463). Wird ein Bildungsziel von Schülern nicht erreicht, wird also zunächst auf ein Förderdefizit aber nicht auf ein Versagen des Kindes bzw. Jugendlichen geschlossen. Die obligatorischen Prüfungen (in den Klassenstufen 5 und 9. Vgl.: DAUN et al., 2004, 454) werden daher nicht zur Evaluation der Schülerinnen und Schüler genutzt, sondern sollen in erster Linie den Lehrkräften als Orientierung für das eigene Handeln dienen. Bemerkenswert ist, dass in Schweden das Lernen nicht nur als Sache der Schülerinnen und Schüler bzw. im weiteren Sinne auch der Lehrkräfte verstanden wird, sondern dass es als gesamtgesellschaftliche Aufgabe verstanden wird. Dies spiegelt sich insbesondere in den zweimal jährlich stattfindenden Gesprächen von Lehrkräften, Schülerinnen und Schülern und deren Eltern wider, in denen der Leistungsstand besprochen und Strategien für die zukünftige Förderung abgestimmt werden.

Wie gezeigt, konzentrieren sich die aufgeführten Länder auf ein in der jeweiligen nationalen Bildungstradition basierendes Selbstverständnis der Bildungsvermittlung, wobei der Lernerfolg jeweils zentral erscheint und konsequenterweise auch fortlaufend überprüft wird und, im Sinne einer Output-Steuerung, als rückwirkende Information in den Bildungsprozess einfließt. Der Weg zur erfolgreichen Zielerreichung wird somit nicht dem Zufall überlassen sondern unter Berücksichtigung verschiedener Rahmendaten systematisch evaluiert und beeinflusst. Besonders fallen hierbei die Niederlande auf, in denen explizit auch Anforderungen an Lehrkräfte und Unterrichtsprozesse genannt werden. Auch wenn die verschiedenen Zielbeschreibungen in den Ländern, ebenfalls wieder aus Gründen unterschiedlicher nationaler Bil-

dungstraditionen heraus, nicht als „Standards“ bezeichnet und verstanden werden, so können sie – durch die deutsche Brille betrachtet – als solche interpretiert werden. Somit wird deutlich, dass eine verbindliche Formulierung von Lernzielen nicht nur wichtig, sondern essenziell bedeutsam ist für eine zielgerichtete Entwicklung von Bildungsprozessen.

## **2.2 Gemeinsamkeiten erfolgreicher Bildungssysteme**

Neben den vorgestellten punktuellen Merkmalen der sechs Länder lassen sich aus den Gutachten in Bezug auf den Umgang mit Bildungsstandards, Evaluationsstudien und Systemmonitoring einige Gemeinsamkeiten herausarbeiten.

Auch wenn in den vorgestellten Ländern der Begriff der Standards bzw. Bildungsstandards nicht einheitlich gebraucht wird, so zeichnet sich doch ab, dass von Seiten der Bildungsadministration Leistungserwartungen „irgendwo auf einem Kontinuum zwischen allgemeinen Bildungs- und Erziehungszielen einerseits und Grenzwerten für einen konkreten Fachleistungstest andererseits“ beschrieben werden (ARBEITSGRUPPE „INTERNATIONALE VERGLEICHsstudIE“, 2003, 177). Dass hierzu nicht wie in deutschen Curricula noch oft verbreitet, sehr kleinschrittig vorgegangen werden muss, lässt sich am Beispiel Finnlands ablesen: Hier ist die Liste von Kriterien für den muttersprachlichen Unterricht auf lediglich ein- einhalb Seiten festgeschrieben. An diesem Beispiel wird konkret deutlich, welche Verantwortung auf die Lehrenden übertragen wird. Sind die Lernziele bzw. Leistungserwartungen festgelegt und werden diese auch regelmäßig überprüft, liegt es im professionellen Handeln der Lehrerinnen und Lehrer, Wege zu finden, auf denen sie möglichst gut erreicht werden. Um tatsächlich nicht erst am Ende einzelner Stationen der Bildungskarrieren von Schülerinnen und Schülern festzustellen, dass sie das gesetzte Ziel nicht erreicht haben, müssen die Lehrkräfte nicht nur eine professionelle Individualdiagnostik beherrschen, mit der sie kontinuierlich den Lernweg der Kinder und Jugendlichen verfolgen, sondern sie müssen auch über eine entsprechende Methodenkompetenz und Verantwortungsbereitschaft verfügen, die eine notwendige individuelle Förderung ermöglicht. Es wird also von den Lehrkräften auf der einen Seite erwartet, dass sie Defizite bei den ihnen überantworteten Schülerinnen und Schülern erkennen, auf der anderen Seite werden aber auch nicht nur die monetären Mittel zur Verfügung gestellt, um hierauf angemessen reagieren zu können.

Die erwähnte Diagnostik ist von den Lehrkräften nicht „nebenbei“ zu machen, insbesondere nicht, wenn das Niveau der gesamten Klasse bzw. der Schule in den Blick genommen werden soll. In den vorgestellten Ländern werden für diese weitergehenden Evaluationen professionelle Agenturen in Anspruch genommen (ARBEITSGRUPPE „INTERNATIONALE VERGLEICHsstudIE“, 2003). Alle vorgestellten Länder haben langjährige Erfahrungen sowohl mit Systemmonitoring-Studien (weshalb in der Vergleichsstudie auch die Niederlande berücksichtigt wurden) als auch mit nationalen und zum Teil mit regional begrenzten Evaluationsstudien. Selbstverständlich fließen die Befunde dieser Untersuchungen sowohl in die pädagogischen als auch in administrative Entscheidungsprozesse ein. Die ermittelten Informationen werden durch regelmäßige nationale Evaluationen und auch durch die kontinuierliche Teilnahme an internationalen Schulvergleichsstudien auf dem aktuellen Stand gehalten.

Bemerkenswert in den hier exemplarisch vorgestellten Ländern ist, dass nicht nur die nationalen Bildungsadministrationen ihre Entscheidungen auf die Ergebnisse aktueller Studien stützen. Auch die Lehrkräfte gebrauchen die Befunde als Hilfe für die eigene professionelle Weiterbildung. Tatsächlich nutzen die einbezogenen Institutionen und Personen die Forschung zur Qualitätsentwicklung von Unterrichts-, Schul- und Bildungssystem.

### **3 Bildungsstandards: Eine Perspektive für Deutschland?**

Zusammenführend kann festgestellt werden, dass es – man könnte fast sagen, dem PISA-Schock sei Dank – Überlegungen zu tiefgreifenden Veränderungen im deutschen Bildungssystem gibt, die, vor dem Hintergrund der beschriebenen Länder, auch zielführend erscheinen. Zum einen orientiert sich in den hier vorgestellten erfolgreicherer Ländern der PISA-Studie die Steuerung des Bildungsprozesses klar am Bildungserfolg der Schülerinnen und Schüler. Diese *Outputsteuerung* wird aber nur dann der Zielsetzung gerecht, wenn die im Rahmen von systematischen Evaluationen erhobenen Daten auch tatsächlich in die Verbesserung der Bildungsprozesse einfließen.

Eine wichtige Voraussetzung für die Akzeptanz der Evaluationen scheint die Transparenz des Evaluationsprozesses von den Anfängen bis zur Bewertung der Ergebnisse zu sein. Das in Deutschland wahrgenommene Unbehagen der in die Evaluation einbezogenen Personen lässt sich dadurch erklären, dass diese den Evaluationsprozess als „*black box*“ erleben. Insbesondere den Lehrkräften sind Nutzen, Möglichkeiten aber auch Grenzen der Evaluation nicht vertraut – so dass von den durchführenden Evaluatoren darauf geachtet werden sollte, diese Aspekte deutlich herauszustellen. Allerdings sollte von den Lehrkräften auch eine gewisse Offenheit gegenüber Evaluationen – die sie in zunehmendem Maße mit gestalten werden – erwartet werden dürfen.

Sowohl die flächendeckende Überprüfung von Kompetenzständen unterschiedlicher Schülerjahrgänge als auch das Umdenken in den Steuerungsprozessen scheint wie vorgestellt, in die richtige Richtung zu weisen. Der systematische Versuch, beide Ansätze zusammenzuführen ist die Einführung von Bildungsstandards. In der Expertise zu Bildungsstandards formuliert die interdisziplinär besetzte KLIEME-KOMMISSION folgende Empfehlung:

„Bildungsstandards benennen präzise, verständlich und fokussiert die wesentlichen Ziele der pädagogischen Arbeit, ausgedrückt als erwünschte Lernergebnisse der Schülerinnen und Schüler. Damit konkretisieren sie den Bildungsauftrag, den Schulen zu erfüllen haben. Für die Gestaltung von Bildungsstandards werden in dieser Expertise folgende Vorschläge gemacht: Bildungsstandards greifen allgemeine Bildungsziele auf. Sie legen fest, welche Kompetenzen die Kinder oder Jugendlichen bis zu einer bestimmten Jahrgangsstufe mindestens erworben haben sollen. Die Kompetenzen werden so konkret beschrieben, dass sie in Aufgabenstellungen umgesetzt und prinzipiell mit Hilfe von Testverfahren erfasst werden können.“ (KLIEME et al. 2003).

Die flächendeckende Überprüfung von Bildungsstandards in den Schulen soll als groß angelegte Schulvergleichsuntersuchung durchgeführt werden. Es ist besonders zu betonen, dass die Überprüfung der Bildungsstandards, wie sie die KLIEME-KOMMISSION empfiehlt, nicht zur individuellen Diagnostik oder Erteilung von Bildungszertifikaten geeignet ist – auch wenn Gegenteiliges zur Zeit diskutiert wird. In der Expertise wird explizit von einer Verwendung der Standards bzw. standardbezogener Tests für Notengebung und Zertifizierung abgeraten.

Was sind die Vorteile bzw. Chancen, die sich bei der Einführung von Bildungsstandards einem Bildungssystem eröffnen? Die Überprüfung von Bildungsstandards soll den Bildungsträgern und den beteiligten Schulleitern und Lehrkräften Rückmeldungen über den Ausbildungsstand der überprüften Kompetenzen bieten. Auf deren Grundlage können die Vermittlungsprozesse optimiert werden, so dass möglichst viele Kinder und Jugendliche wenigstens die Mindeststandards erreichen. Bildungsstandards übernehmen somit eine wichtige Orientierungsfunktion nicht nur für Schüler und Lehrer sondern auch für die Bildungsadministration und die Eltern.

### **3.1 Mögliche Konsequenzen für die Schule und den Unterricht**

Die Befunde nationaler Vergleiche müssen nicht unbedingt, wie es für Schweden vorgestellt wurde, den Unterricht für die jeweils getesteten Schülerinnen und Schüler beeinflussen, sondern sie können, nach entsprechender Auswertung und Implementierung in der Schule, zukünftigen Schülergenerationen zu Gute kommen. An dieser Stelle ist noch einmal auf die Forderung der KLIEME-KOMMISSION zu verweisen, dass Evaluationsdaten zur Überprüfung der Bildungsstandards (oder im weiteren Sinne des Bildungssystems) nicht in die individuelle Benotung der Kinder und Jugendlichen einfließen dürfen. Jede Art der Diagnostik, ob auf System- oder auf Individualebene verlangt entsprechend angepasste Testverfahren. Für alle Evaluationen im Bildungsbereich gilt jedoch, unabhängig von der untersuchten Ebene, dass es nicht um eine schlichte Dichotomisierung in gute oder schlechte Lehrer, Schüler oder Bildungsadministrationen geht, sondern um die kontinuierliche Weiterentwicklung des Bildungsprozesses auf der Basis empirischer Evidenzen. Gerade die systematische Überprüfung von Lernerfolgen kann den Lehrkräften einen weiten Rahmen geben, etablierte Unterrichtskonzepte oder auch innovative, noch wenig bekannte Ansätze zielorientiert zu entwickeln, mit der Gewissheit, dass bei der Identifizierung ungeeigneter Unterrichtsansätze rechtzeitig reagiert werden kann. Jedoch liegt es in der Verantwortung der Lehrkräfte, diese auch tatsächlich regelmäßig und systematisch zu hinterfragen, oder mit anderen Worten zu evaluieren. Hier könnte ein Blick nach Frankreich lehrreich sein: Dort wird neben zentralen nationalen Tests den Lehrkräften über das Ministerium ein kontinuierlich aktualisiertes Angebot von individual-diagnostischen Tests angeboten von denen rege Gebrauch gemacht wird.

Die Erweiterung der Entscheidungsspielräume der Lehrkräfte bzw. des gesamten Kollegiums in Verbindung mit der Einführung von Bildungsstandards setzt voraus, dass auch tatsächlich Wahl- und Entscheidungsfreiheit in der Gestaltung des jeweiligen Bildungs- und auch Betreuungsangebotes gewährleistet werden, um das tatsächliche Potenzial, welches in der

Einführung der Bildungsstandards steckt, ausschöpfen zu können. Wird dieser Weg gegangen, zählt der Erfolg und nicht die konsequente Einhaltung eingefahrener und eventuell inzwischen ausgefahrener, wenig effektiver Wege.

### **3.2 Mögliche Konsequenzen für die Aus- und Weiterbildung der Lehrenden**

Die zuvor vorgestellte Vision der Reform des Bildungssystems zugunsten zunehmender Orientierung an Leistungserwartungen bei Schülerinnen und Schülern, einer zunehmenden Mobilisierung pädagogischer Freiheit bei den Lehrkräften, einer fortschreitenden Autonomisierung der Schulleitungen und einer wachsenden Akzeptanz regelmäßiger Evaluationen als Mittel der Schul- und Unterrichtsentwicklung gegenüber scheint mit der Einführung von Bildungsstandards kurz- bis mittelfristig realisierbar. Doch was bedeutet diese Umstellung für die Lehrkräfte und deren Aus- bzw. Weiterbildung? Vor dem geschichtlichen Hintergrund der traditionell geisteswissenschaftlich geprägten Lehrkräfteausbildung ist zu bedenken, dass ein großer Teil von Lehrerinnen und Lehrern, insbesondere in den alten Bundesländern, dem Testen von Lernerfolgen mit Vorbehalten gegenüber stehen. Diese Vorbehalte sind einerseits ideologisch begründet, basieren aber häufig auch auf fehlender Fachkenntnis, wobei letzteres weniger den Lehrkräften selber anzulasten ist. Die praktisch zwanzigjährige empirische Forschungslücke schlägt sich auch in Defiziten in Bezug auf empirisches Basiswissen der Pädagogen nieder. Um also die Akzeptanz gegenüber empirischen Untersuchungen zu erhöhen, sollten Lehrkräfte systematisch sowohl in der universitären Ausbildung als auch bei obligatorischen Weiterbildungsmaßnahmen entsprechend auf den aktuellen Stand der empirischen pädagogischen Forschung gebracht werden. Diese Aus- und Weiterbildung sollte auch die Vermittlung der oben skizzierten diagnostischen Fertigkeiten umfassen, ohne die eine individualisierte Förderung von Schülerinnen und Schülern im System eines output-orientierten Unterrichts nicht auskommt. Tatsächlich kann man sich an dieser Stelle auch an den Niederlanden orientieren, die bereits Standards für die Unterrichtsorganisation und Unterrichtsgestaltung zur Sicherung einer einheitlichen Unterrichts(mindest)qualität entwickelt haben.

Denkt man diese Gedanken konsequent weiter, kommt man schließlich auch auf Standards der Lehrerausbildung, die an Universitäten oder pädagogischen Hochschulen einzurichten wären. Aber auch hier – folgt man den Argumenten von TERHART – darf die Überprüfung der Lehrkräfte (oder deren Ausbildung) nicht an Prozessmerkmalen guten Unterrichts oder effizienter Schulorganisation stehen bleiben – schließlich ist Ziel jedes Unterrichts Schülerinnen und Schüler oder Studierende in ihrer Entwicklung zu unterstützen (TERHART, 2002). Langfristig muss die Implementierung und Evaluation von Bildungsstandards – verstanden als eine Qualitätsfeststellung des geleisteten Unterrichts – auch zu einer Optimierung der Rahmenbedingungen der Lernprozesse führen.

## **Literatur**

ARBEITSGRUPPE „INTERNATIONALE VERGLEICHsstUDIE“ (2003): Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten. Bonn: Bundesministerium für Bildung und Forschung (BMBF).

BAUMERT, J., BOS, W. & LEHMANN, R. H. (Hrsg.) (2000a): TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Bd. 1. Opladen: Leske + Budrich.

BAUMERT, J., BOS, W. & LEHMANN, R. H. (Hrsg.) (2000b): TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Bd. 2. Opladen: Leske + Budrich.

BAUMERT, J., KLIEME, E., NEUBRAND, M., PRENZEL, M., SCHIEFELE, U., SCHNEIDER, W., STANAT, P., TILLMANN, K.-J. & WEIß, M. (Hrsg.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.

BAUMERT, J., LEHMANN, R. H., LEHRKE, M., SCHMITZ, B., CLAUSEN, M., HOSENFELD, I., KÖLLER, O. & NEUBRAND, J. (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Budrich.

DAUN, H., SLENNING, K. & WALDOW, F. (2004). Sweden. In: H. DÖBERT, E. KLIEME & W. SROKA (Hrsg.): Conditions of school performance in seven countries. Münster: Waxmann, 430-483

DÖBERT, H., KLIEME, E. & SROKA, W. (Hrsg.) (2004): Conditions of School Performance in Seven Countries. Münster: Waxmann.

EINSIEDLER, W. (1997): Unterrichtsqualität und Leistungsentwicklung: Literaturüberblick. In: F. E. WEINERT & A. HELMKE (Hrsg.), Entwicklung im Grundschulalter. Weinheim: Beltz, 225-240.

LEHMANN, R. H., PEEK, R., PIEPER, I. & STRITZKY, R. v. (1991): Lesefähigkeit und Lesegewohnheiten von Achtklässlern. Deutschunterricht (44), 410-417.

PREISERT, H. (1967): Soziale Lage und Bildungschancen in Deutschland. München: Piper.

TERHART, E. (2002): Standards für die Lehrerbildung. Eine Expertise für die Kultusministerkonferenz. Münster: Universität Münster. (ZKL-Texte Nr. 24).