

Empirische Befunde zur Schulstrukturfrage in der Sekundarstufe I – Teil 1

Welche (bildungs-)politischen Optionen lassen sich aus den Befunden ableiten?

Die Diskussion um die Schulstruktur in der Sekundarstufe I ist u. a. auf dem Hintergrund der bildungspolitischen Entwicklung in einigen Bundesländern wie Hamburg und Schleswig-Holstein wieder aktuell. In den kontroversen Diskussionen um diesen Problemkomplex wird immer wieder mit empirischen Befunden insbesondere im Gefolge der großen Leistungsuntersuchungen (TIMSS, PISA) argumentiert.

Gerd Möller

Bochum

Gerhard Orth

Bielefeld

Der folgende Beitrag versucht den Erkenntnisprozess zu den empirischen Befunden zur Schulstrukturfrage seit den siebziger Jahre nachzuzeichnen und dabei zu klären, welche (bildungs-)politischen Optionen sich aus den Befunden ableiten lassen. Er zeichnet dabei die im Zeitablauf durchaus unterschiedlichen Befundlagen bis zu den letzten Auswertungen der Daten von PISA 2000 im Sommer 2006 nach. Er verzichtet bewusst darauf, ein bildungspolitisches Programm aus den beschriebenen empirischen Erkenntnissen abzuleiten. Dies ist aus Sicht der Autoren ein eigenständiges Feld der politischen Diskussion, in dem es unterschiedliche Optionen gibt.

Einführung

Die Frage der Schulstruktur in der Sekundarstufe I ist einmal als der ewige »Wiederkehrer« der deutschen Bildungspolitik bezeichnet worden. Wichtige Stationen sind die Kontroversen im Gefolge des Bildungsgesamtplans aus dem Jahre 1973, der sich für die »Organisationsform der integrierten Gesamtschule« aussprach und die darauf folgende langjährige zum Teil erbitert geführte Schulstrukturdebatte. Diese führte zu einer Etablierung von Gesamtschulen in einer Reihe von Bundesländern, wobei die Gesamtschulen nicht – wie im Bildungsgesamtplan noch vorgesehen – als ersetzenden, sondern als zusätzliche Schulform – etabliert wurden.

Nach einer Phase eines eher ritualhaft ablaufenden bildungspolitischen Grabenkriegs in den siebziger und achtziger Jahren flammte die Diskussion im Gefolge der internationalen Leistungsvergleiche und (TIMSS, PISA, IGLU) seit 1997 wieder auf. Dabei verlagerten sich die Auseinandersetzungen zu einem erheblichen Teil zu einem Streit über die richtige Interpretation der empirischen Befunde, aus denen die Überlegenheit eines integrierten Schulsystems oder eines gegliederten Schulsystems abgeleitet wurde. Zitate von Wissenschaftlern gibt es zuhauf, auf die man sich bei der einen oder anderen Position berufen konnte.

Man muss sich aber an diese Unterscheidung zwischen empirischen Befunden bzw. aus ihnen möglichen Schlussfolgerungen zu den Effekten und Problemen von unterschiedlichen Schulstrukturen und einer – primär normativ und bildungspolitisch argumentierenden – aber auf dieser Basis ebenfalls legitimen Positionsbestimmung zur Schulstrukturfrage gewöhnen. Vor allem ist es gewöhnungsbedürftig, dass sich für dieselben Experten wie für die Forscher aus dem ersten PISA-Konsortium (PISA 2000) um Jürgen Baumert zu unterschiedlichen Zeitpunkten unterschiedliche Aussagen zur Schulstrukturfrage zitieren lassen.

Fortschreitender Erkenntnisprozess der empirischen Bildungsforschung

Dies ist aber nicht auf einen Wankelmut dieser Experten zurückzuführen, sondern auf den fortschreitenden Erkenntnisprozess der empirischen Bildungsforschung, die zu einem bestimmten Zeitraum bewusst nur auf Grund der jeweils ausgewerteten Daten Aussagen machen bzw. nicht machen konnte und wollte.

Im Hinblick auf die Kriterien, nach denen empirische Befunde im Hinblick auf die Schulstruktur untersucht und in den Konsequenzen bewertet werden sollen, orientieren sich die folgenden Ausführungen an dem Konsens, der sich in den letzten Jahren im Hinblick auf zentrale Ansprüche an ein Bildungsmonitoring auf Systemebene weitgehend herausgebildet hat. Es soll vor allem um die erreichten Schülerleistungen (*Kriterium »Leistung«*) sowie um die Abhängigkeit der Leistung von der sozialen Herkunft (*Kriterium »soziale Gerechtigkeit«*) gehen. Im Folgenden sollen die Kriterien »Leistung« und »Soziale Gerechtigkeit« gleichgewichtig berücksichtigt werden.

Erkenntnisse aus der Gesamtschulforschung

Im Vorfeld und im Gefolge des Bildungsgesamtplans hat es für den zunächst als Schulversuch deklarierte integrierte Gesamtschule eine sehr umfassende Forschung gegeben, deren zentrale Ergebnisse von Helmut Fend 1982 zusammengefasst wurden (Helmut Fend, Gesamtschule im Vergleich – Bilanz der Ergebnisse des Gesamtschulversuchs, Weinheim und Basel 1982, dort insbesondere die Zusammenfassung S. 483–507). Im Ergebnis stellt Fend fest, dass die Gesamtschule im Vergleich zu dem gegliederten Schulsystem im Hinblick auf das Kriterium »soziale Gerechtigkeit« zu einer Reduktion, aber nicht zu einer Aufhebung des Zusammenhangs von Bildungslaufbahn (erreichte formale Niveaus, Durchlässigkeit, Abschlüsse) und sozialer Herkunft führte. Allerdings waren die erreichten Schulleistungen für die Gesamtschulen nicht immer überzeugend und verbesserte Abschlussquoten wurden zum Teil durch Senkung der Vergabestandards erreicht (Jürgen Baumert, 1983). Ergab sich bei den eher leistungsschwachen Schülerinnen und Schülern noch ein Gleichstand oder in Einzelfällen eine Verbesserung gegenüber dem gegliederten Schulsystem so waren die Leistungen bei den leistungsstärkeren Schülerinnen und Schülern in den Gesamtschulen nicht besser, bei einem Teil der untersuchten Schulen zum Teil sogar schlechter.



Fend macht ergänzend u. a. folgende Anmerkungen, die auch aus heutiger Sicht noch bedeutsam sind: Er betont, dass Organisationsformen im Schulwesen nicht im Sinne »naturwissenschaftlicher Gesetzmäßigkeiten« wirken. Es gibt sowohl in einem gegliederten wie in einem integrierten System Spielräume für die interne Ausgestaltung. Er verweist ergänzend auf die erhebliche Streuung in den Schulen im Hinblick auf die Effekte zur Chancengleichheit und den Schülerleistungen hin und stellt die Frage nach der Qualität der einzelnen Schulen. Gleichwohl haben schulstrukturelle Rahmenbedingungen aus seiner Sicht schon eine prägende, aber keine deterministische Bedeutung.

Kriterium »Soziale Gerechtigkeit«

Im Ergebnis war im Hinblick auf die Gesamtschulforschung in den siebziger Jahren feststellbar, dass ein integriertes Schulsystem im Hinblick auf das Kriterium »soziale Gerechtigkeit« Vorteile zumindest auf der Ebene der Vergabe von Berechtigungen erbrachte, bei denen allerdings nicht geklärt war, ob sie in einem weiterentwickelten gegliederten Schulsystem nicht in ähnlicher Weise eingelöst werden könnten. Zugleich ergaben sich keine eindeutigen Vorteile für ein integriertes Schulsystem im Hinblick auf das Kriterium »Leistung«. Allerdings waren auch keine systembedingten Nachteile nachweisbar, es gab aber empirisch erkennbare Schwächen zumindest in Teilbereichen der Leistungsentwicklung an den damals untersuchten Gesamtschulen. Der Nachweis einer grundsätzlichen Überlegenheit eines integrierten Schulsystems gegenüber dem gegliederten Schulsystem war damit letztlich nicht erbracht, auch wenn das integrierte System eine in Deutschland ohne Zweifel mögliche und vertretbare politische Option blieb.

Der »TIMSS-Schock«

Seit den 80er Jahren blieb es in Deutschland zunächst bei der kontroversen Bewertung von Gesamtschule und gegliedertem Schulsystem. Bewegung kam in die Diskussion als 1997 die Ergebnisse des ersten in Deutschland ernst genommenen internationalen Leistungsvergleichs vorlagen.

Der »TIMSS-Schock« bestand zunächst darin, dass Deutschland in den dort untersuchten Testbereichen Mathematik und Naturwissenschaften die durchschnittlichen Leistungen im internationalen Vergleich nur im unteren Mittelfeld lagen und wieder einmal hef-

tige Diskussionen über die neue »deutsche Bildungskatastrophe« auslöste. Gleichzeitig wurde die enorme Leistungsstreuung bei den untersuchten Schülerinnen und Schülern der Klassen 7 beklagt.

Im Hinblick auf schulformspezifische Ergebnisse (bundesweit, nicht länderbezogen) bildete TIMSS zum ersten Mal die heute inzwischen aus den verschiedenen PISA-Untersuchungen bekannte Schulformhierarchie auf der getesteten Leistungen (Mittelwerte) ab: an der Spitze das Gymnasium, dann mit einem deutlichen Abstand die Realschule, dann die Gesamtschule und an letzter Stelle die Hauptschule. Diese »Hierarchie« der Schulformen auf der Ebene der erreichten Leistungen bedeutete aber entgegen einem weit verbreiteten Missverständnis keinen Hinweis auf die Effizienz der unterschiedlichen Schulformen, da diese bei ihrer Arbeit von deutlichen Unterschieden im Hinblick auf das für die Schülerleistungen bedeutsame »soziale und kulturelle Kapital« aus den Herkunftsfamilien der Schülerinnen und Schüler ausgehen.

Diese Ergebnisse von TIMSS (siehe dazu im einzelnen Jürgen Baumert und Rainer Lehmann u. a., TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich – Deskriptive Befunde, Opladen 1997) belebten die Schulformdebatte erneut in Deutschland. Von den Gesamtschulbefürwortern wurden zwar auf die Überlegenheit zahlreicher Länder mit Einheitschulen im internationalen Vergleich verwiesen. Auf der anderen Seite nahmen die Vertreter des gegliederten Schulsystems die Ergebnisse von TIMSS als Beleg für die Überlegenheit dieses Systems und verwiesen plakativ auf die geschilderte Schulformhierarchie bei den Schulleistungen sowie die – ein offizieller Ländervergleich fand damals wegen einer zu geringen Stichprobe nicht statt – durchgesickerte Überlegenheit von Bayern gegenüber Nordrhein-Westfalen auf der Leistungsebene.

Grenzen und Möglichkeiten der Erkenntnisse aus TIMSS

Beide Begründungszusammenhänge waren nicht tragfähig und deshalb veranlasste die Diskussion die deutsche Forschergruppe für TIMSS unter Leitung von Jürgen Baumert mehrfach zu klärenden und richtig stellenden Beiträgen (Siehe dazu: Jürgen Baumert/Olaf Köller, 1998), die Aussagen zu Leistungen und Grenzen solcher Studien machten:

- Sie stellten zunächst fest, dass es »keinen einzelnen durchschlagenden Kausalfaktor« gäbe, »der für die teilweisen immensen Leistungsunterschiede zwischen den Schülern der Teilnehmerstaaten verantwortlich wäre«.
- Weiter hieß es: »Aus den Befunden von TIMSS lässt sich kein strukturelles Argument für oder gegen integrierte Schulformen gewinnen. Dies gilt für beide Argumentationsrichtungen gleichermaßen ... Strukturdebatten verfehlen die zentrale Problemstellung der Optimierung vorhandener Bildungsgänge«.
- Gründe für Leistungsunterschiede ließen sich »wahrscheinlich« in der »generellen Wertschätzung von Bildung und schulischem Lernen«, »in der spezifischen Lernkultur eines Schulwesens«, »in der gesellschaftlichen und schulischen Wertschätzung bestimmter Fächer« und »in der Qualität des Fachunterrichts« finden.

Im Ergebnis lässt sich nach TIMSS folgende Zwischenbilanz ziehen: TIMSS bildete zum ersten Mal die inzwischen bekannte, aber häufig falsch interpretierte Schulformhierarchie auf der Leistungsebene (Mittelwerte) ab, betonte dann aber in der Bewertung der entsprechenden Bildungsforscher die Zweifel an der Relevanz unterschiedlicher Schulstrukturen, die bereits Fend formuliert hatte. Damit wurde im Hinblick auf das Kriterium »Leistung« nicht der Vorstellung von der Überlegenheit eines der beiden konkurrierenden Schulstrukturmodelle gefolgt, sondern auf die Vorstellung gesetzt, dass im Hinblick auf das Leistungskriterium die Qualität der Arbeit in den einzelnen Schulformen und Schulen maßgeblich seien. Damit ergaben sich zunächst bezogen auf den Leistungsaspekt keine empirisch abgesicherten Begründungen für eine der traditionellen Positionen im Schulstrukturstreit.

PISA – der internationale Vergleich (PISA 2000, 2003)

Nach TIMSS kam die nächste Welle von Leistungsvergleichen mit PISA. Im Dezember 2001 lagen die Ergebnisse aus dem internationalen Vergleich von PISA 2000 vor, auf den sich die folgenden Aussagen zunächst beziehen. Die Leistungsergebnisse für Deutschland waren für die drei untersuchten Bereiche Lesekompetenz, Mathematik und Naturwissenschaften in ähnlicher Weise ungünstig wie bei TIMSS. Die Leistungsstreuung war hoch mit deutlichen Schwächen im unteren Leistungsbereich (eine relativ



große so genannte »Risikogruppe«). Neu war jedoch gegenüber TIMSS die Feststellung, dass der Zusammenhang zwischen sozialer Herkunft und erreichten Leistungen in Deutschland besonders eng war, was eine sehr schlechte Bilanz im Hinblick auf das Kriterium »soziale Gerechtigkeit« bedeutete.

Der Bericht des deutschen Konsortiums zum internationalen Vergleich von PISA 2000 unter Leitung von Jürgen Baumert (Die folgende Darstellung bezieht sich auf die Berichterstattung der deutschen Konsortien zu PISA 2000 und PISA 2003, die aus Sicht der Autoren ohne Einschränkung tragfähig ist. Die Problemlage der in der Datenlage identischen, in den Bewertungen und der daran zum Teil unterschiedlichen Bewertungen und der folgenden Politikberatung durch den internationalen Koordinator der PISA-Studie Andreas Schleicher kann hier aus Platzgründen nicht dargestellt werden.) griff zwar das Problem der Schul- und Unterrichtsorganisation im deutschen Schulwesen beschreibend in der Form auf, dass es in vielen Staaten im Unterschied zu Deutschland integrierte Schulsysteme bis zum Ende der Sekundarstufe I gebe. Der Bericht stellt aber weiter fest, dass es derzeit nicht möglich sei, eindeutige Aussagen zu der Bedeutung unterschiedlicher Formen der Schul- und Unterrichtsorganisation für die Entwicklung von Schülerleistungen zu machen. Der Bericht stellt vielmehr fest: »Offensichtlich variieren die durchschnittlichen Schülerleistungen in den verschiedenen Teilnehmerländern unabhängig von den ... Formen der Schul- und Unterrichtsorganisation.« (PSA 2000, 2001).

Starker Zusammenhang zwischen sozialer Herkunft und Leistung

Der in Deutschland festgestellte besonders straffe Zusammenhang zwischen der sozialen Herkunft und den erreichten Leistungen entsteht nach dem Bericht in einem erheblichen Maße durch die in Deutschland übliche frühe Verteilung der Schülerinnen und Schüler auf unterschiedliche Schulformen. Der Bericht sieht dies durchaus als Problem, wenn er von einem »unerwünschten Nebeneffekt auf institutionell getrennte Bildungsgänge« spricht. Während in den Gymnasien der Anteil der Schüler aus der Oberschicht hoch sei, »konzentrierten sich in Sonderschulen und einem Teil der Hauptschulen Kinder sozial schwacher Familien, die dann auch das soziale Gefüge einer Einzelschule bestimmen« (ebenda S. 459).

Die zu einem erheblichen Teil leistungsbezogene Schulformgliederung in Deutschland führt also nach dem Bericht unvermeidlich auch zu einer sozialen Segregation in den verschiedenen Schulformen. Der Bericht betont aber ausdrücklich, dass sich nicht feststellen lässt, dass »die soziale Zusammensetzung einer Schule ursächlich für den mehr oder weniger erfolgreichen Kompetenzerwerb der Schülerinnen und Schüler sei« (ebenda S. 466), auch wenn sich in den unterschiedlichen Schulformen unterschiedliche Lernmilieus herausbilden. Dies könne allerdings zu differenziellen Leistungsentwicklungen in den verschiedenen Schulformen führen. Als bedeutsam wird aber letztlich der Einfluss der individuellen sozialen Herkunft auf den Kompetenzerwerb eingeschätzt.

Im Ergebnis ergaben sich aus dem Bericht des deutschen Konsortiums zu den Ergebnisse von PISA 2000 international zwar klare Hinweise eine soziale Segregation in den Schulformen in der Sekundarstufe I, aber keine Belege dafür, dass sowohl nach dem Kriterium »Leistung« wie nach dem Kriterium »soziale Gerechtigkeit« das mehrgliedrige Schulsystem in Deutschland revisionsbedürftig sei. Für beide offensichtlich virulenten Problemkreise ließen sich nach den Hinweisen des Berichts systemimmanente Lösungen im Sinne der Aussagen von Baumert zu den Ergebnissen von TIMSS finden: eine Optimierung der vorhandenen Bildungsgänge vor allem durch einen besseren Unterricht, eine gezielte Förderung insbesondere der leistungsschwächeren Schülerinnen und Schüler und die Sicherung der vertikalen Durchlässigkeit in allen Schulformen. Welche realen Veränderungseffekte es durch solche Maßnahmen geben könne, musste dabei naturgemäß offen bleiben.

Keine Folgerungen für die Schulstruktur

Die Ergebnisse des internationalen Vergleichs aus PISA 2003 wurden von dem neu gebildeten deutschen Konsortium unter Leitung von Manfred Prenzel für die Frage der Schulstruktur in gleicher Weise bewertet wie die Ergebnisse zu PISA 2000. Der Bericht stellt fest, dass im internationalen Vergleich zwischen dem erreichten Kompetenzniveau (Kriterium »Leistung«) und dem Differenzierungsgrad des Schulsystems bzw. dem Alter der Differenzierung kein Zusammenhang bestehe (Prenzel/Baumert, 2003).

Zwischenfazit

Im Ergebnis wurde also aus dem internationalen Leistungsvergleich PISA 2000 bzw. PISA 2003 auf Besonderheiten des deutschen Bildungswesens in der Sekundarstufe I hingewiesen, es ergaben sich aber keine zwingenden empirisch begründeten Notwendigkeiten zu einer Veränderung der Schulstruktur, auch wenn dies weiterhin eine politisch gut begründbare Option blieb.

Baumert bekräftigte diese Position noch einmal in einem Zeitschriftenaufsatz aus dem Jahre 2003: »Ein distanzierter und sachkundiger Betrachter, der keine politischen Botschaften vermitteln will, wird leicht zu dem Schluss kommen, dass die vorliegenden Untersuchungsergebnisse keinesfalls zu eindeutigen Urteilen über Vor- und Nachteile gegliederter oder einheitlicher Schulsysteme führen, vor allem aber keine belastbaren Aussagen über Probleme von Umstellungsprozessen erlauben. ... Berücksichtigt man ferner die Kosten abrupter Traditionsbrüche, mit denen weder Vorstellungswelten ausgetauscht noch eingeschliffene Handlungsroutinen ersetzt werden können, wird man sich hüten, Strukturdebatten aufleben und die Bataillone erneut Stellung beziehen zu lassen.« (Baumert/Artelt, 2003).

Zum Teil vorlaufend, zum Teil parallel zu diesen Auswertungen entwickelten die deutschen Bildungsforscher neuere Erkenntnisse zu der Entwicklung der deutschen Schulstruktur in der Sekundarstufe I und ihren Folgeproblemen, die zu neuen und zum Teil veränderten Bewertungen der Schulstrukturfrage gegenüber dem vorstehend formulierten Zwischenfazit führten. Diese werden im zweiten Teil des Beitrags im nächsten Heft der Schulverwaltung dargestellt.

Jürgen Baumert und Jürgen Raschert, Gesamtschule, in: Erziehung im Jugendalter, Sekundarstufe I. Enzyklopädie Erziehungswissenschaft Bd. 8, Stuttgart 1983, S. 228 – 269

Siehe dazu: Jürgen Baumert/Olaf Köller, Nationale und internationale Schulleistungsstudien – Was können sie leisten, wo sind ihre Grenzen), in: »Pädagogik«, 50. Jg. 1998, Heft 6, S. 12-18

PSA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich, Opladen 2001, S. 427

Manfred Prenzel, Jürgen Baumert u. a., PISA 2003, Ergebnisse des internationalen Vergleichs – Zusammenfassung S. 27

Jürgen Baumert und Cordula Artelt, Bildungsgang und Schulstruktur, in: Pädagogische Führung, Jg. 2003 Heft 4 S. 188

