

Die Klassenlängsschnitt-Analyse „PISA-KLL“ aus PISA 2003

Diese Studie wurde durchgeführt in den Jahren 2003 und 2004 in Klassen des 9. bzw. 10. Jahrgangs im Rahmen der auch „PISA-I-Plus“ genannten „Messwiederholungsstudie (MWH)“ zu PISA 2003.

Ergebnisse der dabei durchgeführten Untersuchungen sind im 3. Berichtsband von PISA 2003 (Waxmann 2006) ausführlich beschrieben worden. Der hat den Titel

„Untersuchungen zur Kompetenzentwicklung im Laufe eines Schuljahres“

Herausgegeben wurde dieser Berichtsband von Manfred Prenzel, Jürgen Baumert, Werner Blum, Rainer Lehmann, Detlev Leutner, Michael Neubrand, Reinhard Pekrun, Jürgen Rost und Ulrich Schiefele.

Begründung der Studie:

„Um förderliche Bedingungen für die Entwicklung von Kompetenzen von Schülerinnen und Schülern zu erhalten, müssen Kompetenzveränderungen empirisch untersucht werden. Wie entwickeln Schülerinnen und Schüler Kompetenzen in ihrer Schulzeit und in welchen Phasen ihrer Schulzeit können sie von welchen Unterstützungsangeboten am meisten profitieren? Zur Beantwortung dieser Fragen ist eine längsschnittliche Erhebung erforderlich.

Erst mit zwei oder mehr Messzeitpunkten lassen sich Veränderungen nachzeichnen. Auch die Bedingungen, die sich förderlich oder erschwerend auf Veränderungen auswirken, lassen sich nur in Veränderungsstudien analysieren. Eine Large-Scale-Studie zu Veränderungen in Deutschland liegt mit PISA-I-Plus erstmalig vor.“ (PISA 2003/06, S.41)

Der letzte Satz muss modifiziert werden, denn er stimmt so nicht.

Schon das Projekt „Schulleistung“ (1968-1970), das erste große Projekt des „Max-Planck-Institutes für Bildungsforschung (MPIB)“, war eine Klassen-Längsschnitt-Analyse. Auch hier wurde an zwei Messzeitpunkten gemessen. Auch hier gab es Auskünfte der Lehrkräfte zu ihrem Unterricht und zu ihren Unterrichtsstrategien. Und sie war doppelt so umfangreich. Aber sie ist kaum bekannt.

Ihr Gegenstand: „die Determinanten von Leistung in der Schule“, also die „Faktoren des Systems“, die das schulische Lernen fördern oder behindern.

Es ging es um die Grundfrage allen Unterrichtens: um die Vereinbarkeit von Leistungsförderung und Leistungsausgleich, genauer: um das Problem „Leistungsunterschiede ausgleichen zu wollen, ohne zugleich leistungsstärkere Schüler zu benachteiligen“ (Jürgen Baumert, Peter Martin Roeder, Fritz Sang und Bernhard Schmitz: „Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen“, 1986, S.639)

Etwa 14.000 Gymnasiasten aus den Ländern der damaligen Bundesrepublik und aus Westberlin haben an der Studie teilgenommen. Darum wird sie gelegentlich auch „Gymnasiastenstudie“ genannt.

Grundlage der Auswertung waren „verarbeitbare Datensätze für 12.594 Schüler aus 427 Klassen und für insgesamt 1.130 Deutsch-, Englisch- und Mathematiklehre“ (Baumert u. a. 1986, S.644).

„Im Anschluss an jüngere Entwicklungen der Lehr-Lern-Forschung wurden auf der Grundlage der Befragung von Fachlehrern zu ihren Unterrichtsstrategien Skalen beziehungsweise Kennziffern zur Qualität und Quantität der Instruktion gebildet.“ (Baumert u. a. 1986, S.644)

„Die Schülerleistungen wurden in den Fächern Deutsch, Englisch und Mathematik mit zwei Tests zu Beginn und zum Ende der 7. Jahrgangsstufe erhoben.“ (Baumert u. a. 1986, S.644)

Es wurden auch Informationen eingeholt zum Bildungsstand der Eltern- und der Großeltern-Generation. Bei den Schülern „wurden die Häufigkeit bildungsrelevanter Tätigkeiten (Lesen, Malen, Musizieren, naturkundliche Beschäftigungen), die tägliche Fernsehzeit sowie die häusliche Belastung

durch schulfremde Aufgaben berücksichtigt. Das allgemeine intellektuelle Fähigkeitsniveau wurde mit dem IST-Amthauer bestimmt.“ (Baumert u. a. 1986, S.644)

Das Projekt „Schulleistung“ war nicht nur die gründlichste, sondern gleichzeitig auch die teuerste Studie, die je in Deutschland durchgeführt wurde. Ihre Auswertung zog sich hin über mehr als 20 Jahre.

Es hatte sich schon damals herausgestellt, dass die Förderung von Leistungen und der Ausgleich von Leistungsunterschieden (auch „Streuung“ oder „Leistungsvarianz“ genannt) eigentlich zwei unvereinbare Zielvorstellungen sind: Das eine geschieht stets auf Kosten des anderen. Eine Lösung gibt es nicht, sondern nur den „ausbalancierten“ Kompromiss.

Roeder und Sang referieren 1991 als Ergebnis: *„Selbst unter den günstigen Bedingungen des Gymnasiums ist es eine schwierige Aufgabe, eine befriedigende Balance zwischen optimaler Förderung und Leistungsausgleich zu finden. Sie wird nur von einer Minderheit der Lehrer wirklich bewältigt.“* (S.164)

Ein Unterricht aber, der auf Leistungsausgleich angelegt ist oder auf Leistungsausgleich angelegt sein muss, führt unvermeidbar zu einer Benachteiligung der leistungsstärkeren Schüler.

„In nach Leistungsgruppen getrennten Analysen konnte gezeigt werden, dass bei streuungsverringendem Unterricht erhebliche Einbußen im Lernfortschritt des oberen Leistungsdrittels relativ schmalen Gewinnen im unteren Leistungsdrittels gegenüberstehen.“ (Baumert u. a. 1986, S.654)

„Beim Vergleich der Leistungsgruppen (drei Niveaus) zeigte sich, dass die Verringerung der Leistungsvarianz innerhalb der Klassen vor allem auf Kosten der Leistungsstarken geht. In Klassen mit Leistungsausgleich in Mathematikunterricht bleiben sie um etwa eine Standardabweichung hinter dem leistungsbesten Drittel der Klassen mit steigender Varianz zurück. Im Englisch- und Deutschunterricht beträgt diese Differenz noch etwa eine dreiviertel Standardabweichung. Dies sind nicht nur statistisch, sondern auch praktisch höchst bedeutsame Unterschiede.“ (Roeder und Sang 1991, S.164)

Nach den vom MPIB eingeführten Maßstäben entsprechen *„0,3 bis 0,5 Standardabweichungen in der Mittelstufe dem Lernfortschritt eines Schuljahres“* (Köller, Baumert, Cortina, Trautwein, Watermann 2004, S.691). Das heißt: Leistungsstarke Schüler, deren Unterricht auf Leistungs-Förderung angelegt ist, erreichen in Mathematik ein doppelt so hohes Lerntempo wie jene leistungsstarken Schüler, deren Unterricht auf Leistungs-Ausgleich angelegt ist.

Wenn aber in einer Klasse die Unterschiede der Vorkenntnisse und der Begabungen allzu groß sind, dann zwingt das die Lehrer zu verstärktem Üben und Wiederholen. *„Diese repetitive Unterrichtsführung nützt wider Erwarten Schülern mit ungünstigen Eingangsvoraussetzungen nur wenig, während die Lernfortschritte des oberen Leistungsdrittels merklich beeinträchtigt werden.“* (Baumert u.a. 1986, S. 655)

Diese Erkenntnisse aus dem MPIB-Projekt „Schulleistung“ hätten mit den Daten der Klassenlängsschnitt-Analyse PISA-KLL und den Daten der COACTIV-Studie aktualisiert werden können.

„Der Ausgangspunkt [...] war die internationale Schultichprobe von PISA 2003 (ohne Förder-/ Sonderschulen und berufliche Schulen). In diesen 198 Schulen wurden per Zufall zwei 9. Klassen zusätzlich zur international vorgeschriebenen Stichprobe der Fünfzehnjährigen gezogen und dann getestet. An der ersten Messung auf der 9. Klassenstufe (im Jahr 2003) mit den üblichen internationalen Erhebungsverfahren nahmen 387 Klassen mit insgesamt 8559 Schülerinnen und Schülern teil. Diese Stichprobe absolvierte ebenfalls einen zweiten Testtag, an dem zusätzliche nationale (stärker curricular orientierte) Tests für Mathematik und Naturwissenschaften sowie ein ergänzender nationaler Schülerfragebogen zu bearbeiten waren.“ (Manfred Prenzel 2006, S.17)

„Außerdem wurden die Lehrkräfte, die in den ausgewählten Klassen Mathematik unterrichteten, nicht nur mit einem Fragebogen befragt, sondern auch gebeten, an der COACTIV-Studie mitzuwirken. Dieses DFG-Projekt (von J. Baumert, W. Blum und M. Neubrand) untersuchte mit unterschiedlichen Verfahren sehr ausführlich und intensiv professionelle Kompetenzen der Lehrkräfte.“ (S.17)

Die Stichprobe der Klassenlängsschnitt-Analyse PISA-KLL enthält nur diejenigen Klassen, „in denen die Jugendlichen zur Erhebung 2004 in demselben Klassenverband wie 2003 unterrichtet worden waren und in denen vollständige Daten für die Jugendlichen und Daten von allen Lehrkräften aus der COACTIV-Studie vorlagen.“ (S.48). Sie umfasst

insgesamt 4.353 Schüler aus 194 Klassen
von 119 Schulen (S.48). Darunter waren

14 Schulen mit mehreren Bildungsgängen,
46 Realschulen,
12 Integrierte Gesamtschulen,
47 Gymnasien

Mit den Daten der Klassenlängsschnitt-Analyse PISA-KLL und den Daten der COACTIV-Studie hätten die bereits erwähnten Ergebnisse des MPIB-Projektes „Schulleistung“ zur Unvereinbarkeit von Leistungsförderung und Leistungsausgleich aktualisiert und bestätigt werden können. Angesichts der Tatsache, dass im Jahre 2003 in vielen Bundesländern von Parteien und Verbänden für die Einführung des „längeren gemeinsamen Lernens“ geworben wurde, wäre eine entsprechende Auswertung der KLL-Daten eine wichtige Aufgabe der Bildungsforschung gewesen.

Der Aufsatz von Martin Senkbeil (S.277-308), in dem diese Problematik hätte angesprochen werden können, steht unter der Überschrift „Die Bedeutung schulischer Faktoren für die Kompetenzentwicklung in Mathematik und in den Naturwissenschaften.“ Da wird anhand von Angaben der Schulleitungen mit viel Aufwand zwischen „aktiven Schulen“ und „passiven Schulen“ unterschieden.

Das letzte Kapitel des Aufsatzes hat die vielversprechende Überschrift „Welchen Einfluss besitzen die Lehrertypen auf die Kompetenzveränderung in Mathematik und in den Naturwissenschaften“ S.301-304). Es wird anhand von Angaben der Lehrkräfte zwischen „aktiven Lehrkräften“, „passiven Lehrkräften“ und „disziplin-orientierten Lehrkräften“ unterschieden und dann die Binsenweisheit mitgeteilt, dass aktive und disziplin-orientierte Lehrkräfte im Unterricht höhere Kompetenz-Zuwächse erreichen.

Im Projekt „Schulleistung“ hingegen wurde unterschieden zwischen Lehrkräften, deren Unterricht auf Leistungssteigerung angelegt war, und Lehrkräften, deren Unterricht darauf angelegt war, die bestehenden Leistungsunterschiede auszugleichen. Den höheren Fördereffekt hatte jener Unterricht, dessen Merkmale **„ein hohes Anspruchsniveau und ein zügiges Fortschreiten im Lernstoff“** waren

Fazit: Das zentrale Problem allen Unterrichtens, „Leistungsunterschiede ausgleichen zu wollen, ohne zugleich leistungsstärkere Schüler zu benachteiligen“, wird in der sehr aufwändigen Klassenlängsschnitt-Analyse PISA-KLL nicht berücksichtigt.

Auch die Frage nach den organisatorischen Rahmenbedingungen effizienten Unterrichtens wurde nicht gestellt, ebenso wenig die Frage, ob es an weiterführenden Schulen bei der Zusammensetzung der Klassen deutliche erkennbare Grenzen der Heterogenität gibt.

Das MPIB-Projekt „Schulleistung“ hatte gezeigt, dass ein über den 4. Jahrgang hinausgehender Unterricht in undifferenzierten oder unzureichend differenzierten Klassen stets zu einer Benachteiligung der leistungsstärkeren Schüler führt (Baumert und Köller im Oerter/Montada 2002, S.770).

Die Klassenlängsschnitt-Analyse PISA-KLL bietet ähnliches Material zur Leistungsentwicklung der Schüler - und mit der COACTIV-Studie noch viel mehr Material zur Qualität des Unterrichtens.

Daher sollte nun von allen Seiten darauf gedrängt werden, dass sowohl die Daten von PISA-KLL 2003 als auch die Daten der COACTIV-Studie recht bald dem „Forschungsdatenzentrum“ des „Institutes für Qualitätsentwicklung im Bildungswesen“ übergeben werden und dort für weitergehende Analysen zur Verfügung stehen. Dann können die Nachteile des „*längeren gemeinsamen Lernens*“ endlich auch anhand von aktuelleren Daten bewiesen - und öffentlichkeitswirksam bekannt gemacht werden.

Ulrich Sprenger, Arbeitskreis Schulformdebatte e.V.

Recklinghausen (2012)

www.schulformdebatte.de – Wissenschaftsorientierte Beiträge zu Fragen der Schulstruktur