

**Norbert SOMMER, Osnabrück**

**„Bremens Gymnasiasten hinken ein Jahr hinterher.“<sup>1</sup>**

Die Schlagzeile stellt keine journalistische Überzeichnung des in Form von Punktdifferenzen angegebenen Leistungsunterschieds verschiedener Schülergruppen dar, sondern sie deckt sich von Intention und Inhalt her mit Aussagen internationaler Vergleichsuntersuchungen. In TIMSS II wird die Zuordnung zwischen Rasch-skalierten Punktwerten und Schulzeit erstmals hergestellt: „Im schulischen Bereich bieten sich als mögliche Vergleichsmaßstäbe die fachlichen Lernfortschritte an, die normalerweise in einem Schuljahr erzielt werden ... Die mittleren Fähigkeitsdifferenzen zwischen einer Jahrgangsstufe zur nächsten betragen im Durchschnitt ungefähr 1/3 Standardabweichungen.“ (Baumert/Lehmann u.a., 1997, S.67). In diesem Maßstab erweist sich der Unterschied zwischen Bundesländern mit hohem und niedrigem durchschnittlichen Kompetenzniveau in der PISA-E 2003 Studie über alle Schüler hinweg sogar als noch größer als in der Schlagzeile für Gymnasien berichtet: „Die innerhalb der Länder erreichten Kompetenzniveaus variieren zwischen 471 Punkten (Bremen) und 533 Punkten (Bayern). Diese Spannweite von 62 Punkten entspricht dem Kompetenzzuwachs von etwa 1,5 Schuljahren.“ (Ehmke, 2005, S. 3)

Schon in diesen Deutungen von Leistungsdifferenzen aus zwei Vergleichsstudien fallen Abweichungen in der Höhe des Zusammenhangs zwischen Punktdifferenz und Schulzeit auf: Bei einer Skalierung auf den Mittelwert 500 und die Standardabweichung 100 sind 1/3 Standardabweichung ca. 33 Punkte. Die in der deutschen Längsschnittstudie innerhalb von TIMSS II beobachtete Differenz in den Mathematikleistungen 13jähriger in den Klassenstufen 7 und 8 ist mit 25 Punkten deutlich geringer; die in der PISA-E Studie beschriebene Differenz von 62 Punkten liegt nach diesem Maßstab näher an zwei Schuljahren als an eineinhalb.<sup>2</sup>

Die Frage der Variabilität des Zusammenhangs zwischen Punktdifferenzen in Mathematikleistungstests und Schulzeit soll näher untersucht werden, wobei theoretische und empirische Problemfelder angesprochen werden.

- Lassen sich Zuwächse in Längsschnittstudien allein auf Unterrichtszeit zurückführen (u.a. Reifung, Retesteffekte)?
- Liegen die Ergebnisse von Querschnittstudien (PISA) in derselben Größenordnung wie Längsschnittstudien?

---

<sup>1</sup> Schlagzeile der Süddeutschen Zeitung vom 2.11.2005 aus Anlass der Veröffentlichung der zweiten PISA-E-Studie 2003

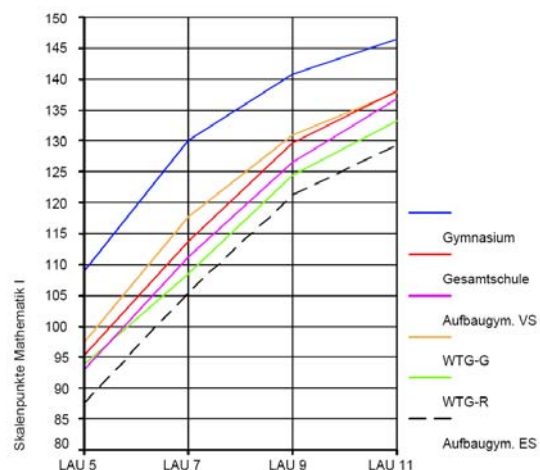
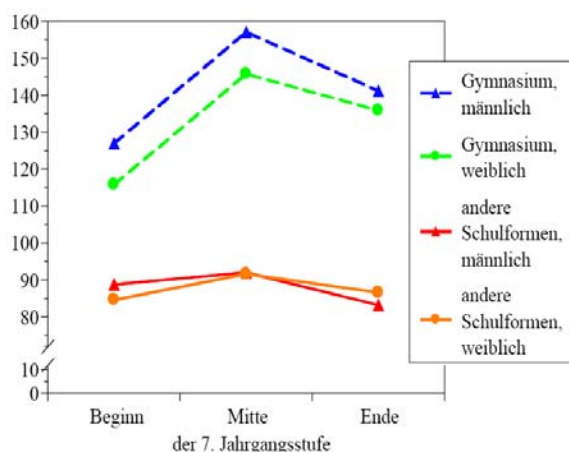
<sup>2</sup> In Veröffentlichungen finden sich auch grobere Angaben für ein Schuljahr, z.B. „etwa 30 bis 40 PISA-Punkte“ (Baumert u.a., 2005), „zwischen einer Drittel und einer halben Standardabweichung“ (Blum/Neubrand u.a., 2004).

- Handelt es sich bei TIMSS-13- bzw. PISA-15jährigen in verschiedenen Jahrgangsstufen um vergleichbare Schülergruppen (Kohorteneffekte)?<sup>3</sup>
- Besteht ein proportionaler Zusammenhang zwischen der Höhe der Punktdifferenz und der Länge der Schulzeit (u.a. Latenzphasen)?
- Ist die Annahme der Eindimensionalität und die Differenzwertbildung bei Leistungen, die zu verschiedenen Zeitpunkten erhoben werden, sinnvoll (u.a. Änderung des Konstrukts Schulleistung über die Zeit; vgl. Köller/Baumert 2002, Lind 2003)?
- Welchen Einfluss hat die Itemzusammenstellung auf die Höhe des Zusammenhangs (u.a. floor-/ceiling Effekt für die jüngere/ältere Untersuchungsgruppe, Abdeckung mathematischer Kompetenz)?
- Welche Beziehung muss zwischen Unterrichts- und Iteminhalten bestehen um einen Lernzuwachs erwarten zu können (u.a. Behandlung abgefragter neuer und Wiederholung abgefragter „alter“ Inhalte zwischen den Messzeitpunkten; vgl. Baumert/Lehmann u.a., 1997)?

Die Untersuchung der Fragen geschieht anhand von

- Veröffentlichungen zu größeren Schulleistungsuntersuchungen,
  - Zusammenstellungen der mitgeteilten Daten der Aufgabenschwierigkeit von 106 veröffentlichten TIMSS II Items,
  - vertieften Analysen des KMK Datensatzes zu PISA-E 2000,
- eigenen Untersuchungen in verschiedenen Jahrgangsstufen mit veröffentlichten TIMSS und PISA Aufgaben.

Die Stabilität der Beziehung „1/3 Standardabweichung entspricht einem Schuljahr“ erweist sich in verschiedenen Längsschnittuntersuchungen als eher gering. Auch die Annahme einer Proportionalität bzw. der Übertragbarkeit von Aussagen eines Abschnitts der Schullaufbahn auf einen früher

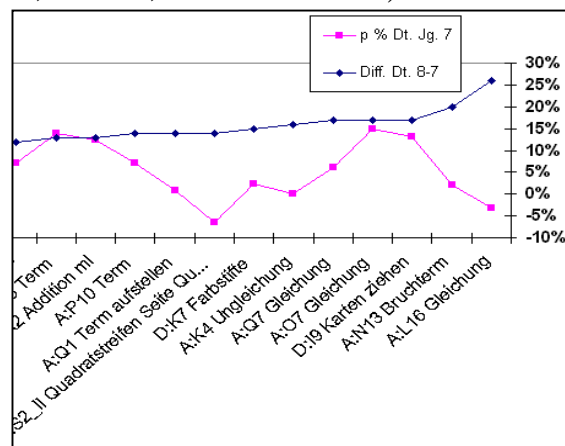


3 Nach eigenen Analysen der KMK-Daten zu PISA-E 2000 unterscheiden sich die Klassenstufen im Alter in Monaten, der Verteilung auf die Schulformen, den Anteilen von frühzeitig bzw. verspätet eingeschulten Schülern und Klassenwiederholungen sowie der allgemeinen kognitiven Fähigkeit.

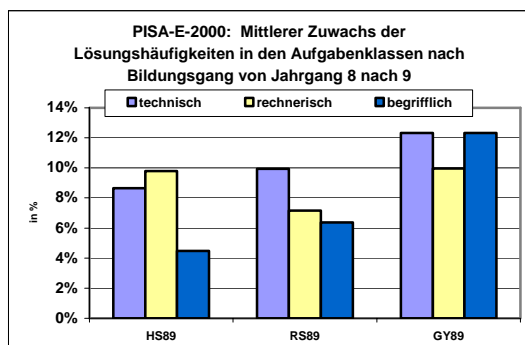
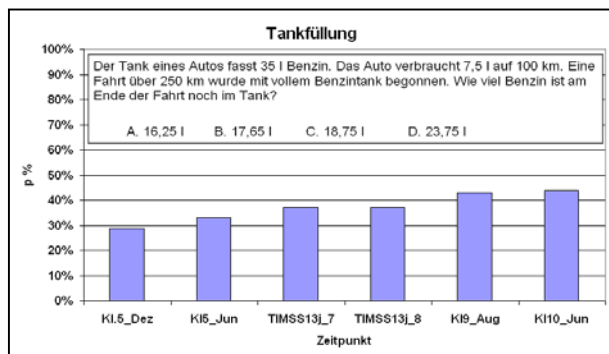
oder später in der Schullaufbahn gelegenen oder eine kürzere oder längere Phase ist nicht in jedem Fall gerechtfertigt. Während die rechte Grafik (s.o.) aus der BIJU-Studie (Hosenfeld, 2000) Schwankungen der Leistungsentwicklung innerhalb eines Schuljahres und Effekte in einer Größenordnung zeigt, die schon im ersten Halbjahr den angenommenen Jahreswert überschreitet, spricht die rechte aus LAU 11 (Lehmann u.a., 2004) für ein Abflachen der Zuwächse in älteren Jahrgangsstufen.

Die Angaben zu veröffentlichten TIMSS II Aufgaben ergeben, dass die auf Schülerebene ermittelten Leistungsdifferenzen zwischen Klasse 7 und 8 von 25 Punkten in Deutschland und 29 Punkten international auf Itemebene einen durchschnittlichen Zuwachs der Lösungshäufigkeit um 4,8 % (dt.; min. - 6,0 % max. 26 %) bzw. 5,7% (int.; min. 0,0 % max. 17 %) bedeuten.

Die nebenstehende Grafik zeigt, welche zehn Items in TIMSS II von Klasse 7 auf 8 den höchsten Zuwachs aufweisen.<sup>4</sup> Sieben der zehn sind dem Bereich Algebra zugeordnet, einem Unterrichtsinhalt, in den schwerpunktmäßig in den Klassen 7 und 8 eingeführt wird. Diese Tendenz lässt sich an den Items aus PISA-E 2000 bestätigen.



In eigenen Untersuchungen fallen Lernzuwächse bei mehrschrittigen „Problemaufgaben“, für die nicht explizit ein Lösungsverfahren im Unterricht erarbeitet wird, wie der Aufgabe „Tankfüllung“, gering aus.

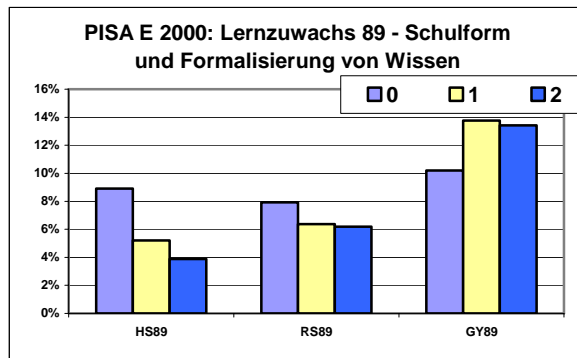


Aber nicht nur curriculare Erklärungen tragen zur Aufklärung von Lernzuwachsmustern bei. Für die PISA-2000 Aufgaben haben Neubrand u.a. (2002), Blum u.a. (2004) und Cohors-Fresenborg u.a. (2004) die Bedeutung von Aufgabenmerkmalen nachgewiesen, die

<sup>4</sup> In dem Diagrammausschnitt bezieht sich die Skalierung rechts nur auf Diff. Dt. 8-7.

kognitive Prozesse bei der Aufgabenlösung betreffen. Bei rechnerischen Modellierungsaufgaben ist der Zuwachs von Klasse 8 auf 9 bei Hauptschülern ähnlich groß wie am Gymnasium (bei unterschiedlichem Ausgangsniveau), der Leistungsvorteil der Gymnasialschüler zeigt sich besonders beim begrifflichen Modellieren.

Auch bei einer Klassifizierung der Aufgaben nach Grundvorstellungsintensität (Blum u.a., 2004), Sprachlogischer und Kognitiver Komplexität, Formalisierung von Wissen (s. Grafik) und Formelhandhabung



(Cohors-Fresenborg u.a., 2004<sup>5</sup>) zeigen sich schulformspezifische Unterschiede im Kompetenzzuwachs. Schon bei einem geringen Anspruch an Formalisierung (Grafik rechts; Stufe 1) fällt der Lernzuwachs von Haupt- und Realschülern gegenüber Schülern aus dem Gymnasium deutlich ab.

Die Ergebnisse sind vermutlich Ausdruck unterschiedlicher Schwerpunktsetzungen im Unterricht der Schulformen, wie sie sich in Lehrerbefragungen im Rahmen von PISA und COAKTIV zeigen (vgl. Neubrand, 2006).

Die Beziehung „1 Schuljahr = 1/3 s“ erweist als relativ stabil, wenn:

- Schüler des mittleren Sek I - Bereichs untersucht werden,
- die Stichprobe ein repräsentatives Bild der Schülerpopulation eines Jahrgangs ist,
- der eingesetzte Test die mathematische Kompetenz umfassend abbildet (Multi-Matrix-Design),
- die Aussagekraft der Testergebnisse am oberen Ende der Leistungsfähigkeit nicht durch Deckeneffekte eingeschränkt ist.

Die Beziehung wird beeinflusst durch:

- die Leistungsfähigkeit der Untersuchungsgruppe (allgemeine kognitive Leistungsfähigkeit; Vorwissensstand; Schulform),
- das Alter der Untersuchungsgruppe (höhere Zuwächse durch schmales Kompetenzfeld bei jüngeren Schülern; bei älteren Schülern Abflachen des Zuwachses),
- die Unterrichtsnähe der in den Testaufgaben berührten Inhalte bezogen auf die Zeit zwischen den Messungen,
- die spezifischen kognitiven Anforderungen der Aufgaben.

<sup>5</sup> Das Literaturverzeichnis kann beim Autor abgerufen werden.