

Hans Brügelmann

LUST auf IGLU-E: Bewertung nach Leistungsstand oder nach Lernzuwachs?

http://www.gew.de/wissen/zeitschriften/e-w/2004/2004-3/texte/d_s-13.htm).

[Manuskript für e&w, H. 3 (Frühjahr) 2004, 13-14]

Die wichtigsten Ergebnisse der internationalen Grundschul-Lese-Untersuchung (IGLU) sind¹: Die Leseleistung deutscher ViertklässlerInnen liegt - international gesehen - im oberen Drittel (Platz 5-12) unter 35 Ländern, also relativ gesehen deutlich höher als die der 15-Jährigen bei PISA. Die Bandbreite der Leistungen beträgt mehrere Schuljahre, ist aber deutlich geringer als in anderen Ländern (und als in Deutschland bei PISA). Nur 10% der Kinder werden als RisikoschülerInnen klassifiziert (bei PISA 25%), wiederum deutlich weniger als in den Grundschulen anderer Länder. SchülerInnen mit Migrationshintergrund und Kinder aus den Familien mit niedrigem sozioökonomischem Status liegen umgerechnet rund 10 bzw. 17 Monate hinter den deutschsprachigen bzw. sozioökonomisch privilegierten AltersgenossInnen zurück - wiederum erheblich weniger als am Ende der Sekundarstufe.

Trotzdem wird weiter beharrlich aus den schlechten Ergebnissen der *15-Jährigen* in PISA (2000) gefolgert, man müsse den Leseunterricht in der *Grundschule* verändern. Nun liegen die Befunde des innerdeutschen Vergleichs von IGLU-E in erweiterten Stichproben aus sechs Bundesländern vor².

„Ein Jahr liegt zwischen Baden-Württemberg und Bremen“, titelte die Berliner Zeitung, als am 28.1. die Ergebnisse veröffentlicht wurden. Und im Untertitel heißt es: „Deutsche Grundschüler gehören zur Spitze Europas. Doch das Gefälle zwischen den Ländern ist groß.“

¹ Bos, W., u. a. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Waxmann: Münster u. a.

² Bos, W., u. a. (Hrsg.) (2004): Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich. Waxmann: Münster.

Keiner ist für Rankings, aber alle gieren nach den Ergebnistabellen der Leistungsstudien wie nach dem 100-m-Lauf bei olympischen Spielen. Fatal ist das weniger wegen des offensichtlichen Unterhaltungswerts einer so ernststen Angelegenheit. Fatal sind vereinfachende Folgerungen, die in Merkmalen des jeweiligen „Systems“ die entscheidende Bedingung für gute Leistungen sehen. Die Unterschiede *innerhalb* der Bundesländer jedenfalls finden viel weniger Aufmerksamkeit - obwohl sie wesentlich größer sind als die Differenzen zwischen den Mittelwerten, so dass sich z. B. die Leistungsverteilungen von Baden-Württemberg und Bremen stark überlappen.

Als Rezept gegen Uneinheitlichkeit und gegen unzureichende Leistungen werden sog. Bildungs- (redlicher: *Leistungs-*) Standards empfohlen. In der zur Zeit von den Bundesländern praktizierten Form handelt es sich um fachliche Leistungsniveaus, die zu bestimmten Terminen (z. B. im 4. Schuljahr) für alle SchülerInnen als Lernziel definiert werden.

Aber wie kommt man zur Bestimmung dieser Niveaus? Für sog. *Regelstandards* kann man auf empirische Daten zurückgreifen. Da die deutschen SchülerInnen in IGLU im oberen Leistungsdrittel lagen, glaubt man, den Mittelwert zur Norm erklären zu können. Allerdings lässt sich schon die Frage, wie man lesen können muss, um den Anforderungen in Beruf und Alltag zu genügen, kaum beantworten³. Wir haben in einer Pilotstudie den von Wilfried Metzke für die Grundschule entwickelten „Stolperwörter-Lesetest“⁴ Stichproben aus zwei Berufsfeldern vorgelegt: MeisterschülerInnen einer Handwerkskammer und LehrerInnen. Im Mittel bearbeiteten die *GesellInnen* pro Minute 11 Sätze richtig, die *LehrerInnen* 19. In der Grundschule waren es 4 Sätze Mitte zweiter und 8-9 Sätze Mitte vierter Klasse. Das heißt konkret: 19% der ViertklässlerInnen erreichen das Durchschnittsniveau der HandwerksschülerInnen und bereits 67% können mit den unteren 10% dieser beruflich erfolgreichen Gruppe mithalten, sogar 98% mit dem schwächsten Meisterschüler.

³ Vgl. die Daten in: OECD & Statistics Canada (ed.) (1995): Grundqualifikationen, Wirtschaft und Gesellschaft. Ergebnisse der ersten internationalen Untersuchung von Grundqualifikationen Erwachsener. Paris/ Ottawa.

⁴ In vorgegebener Zeit müssen in Sätzen nicht passende Wörter identifiziert werden, so dass sowohl Lesegeschwindigkeit und Lesegenauigkeit als auch Textverständnis abgefragt werden.

Regelstandards erscheinen aber noch aus einem zweiten Grund problematisch. Nicht nur die Leistungen der Erwachsenen streuen erheblich, auch unter den ViertklässlerInnen finden wir eine Streuung von +/- zwei Schuljahren. Welchen pädagogischen Sinn haben dann gleiche Anforderungen für alle?

Vor allem der Befund aus unserer Studie LUST⁵, dass fast alle Gruppen von SchülerInnen von ihrem jeweiligen Entwicklungsstand aus vergleichbare Fortschritte machen, spricht dagegen, Leistungshürden an bestimmte Termine zu binden. Das gilt auch für Migrantenkinder. Zwar schneiden sie auf allen Klassenstufen schlechter ab als deutschsprachige Kinder. Aber beide Gruppen machen vergleichbare Fortschritte. Sowohl die Anforderungen an die SchülerInnen als auch die Erwartungen an LehrerInnen müssen sich also auf die unterschiedlichen Voraussetzungen beziehen, von denen her sich die Einzelnen bzw. die Gruppen auf den Weg gemacht haben.

Aber auch die Vorstellung, über die Definition „normaler“ Lernzuwächse lasse sich Unterricht beurteilen und steuern, kann leicht zu mechanistischen Modellen des Lehrens und Lernens verführen. Das wird deutlich an dem Konzept sog. *Mindeststandards*. Die Daten aus den großen Studien helfen da *per se* nicht viel weiter, denn die Leistungen im unteren Bereich befriedigen ja gerade nicht und können deshalb nicht zum Maßstab für die schwächeren SchülerInnen gemacht werden. Vor allem liegen keine Daten dazu vor, welche Leistung Ende vierter Klasse als „Risiko“ für den weiteren Lernweg zu betrachten sind. Aber selbst wenn dieser Datenmangel behoben wäre: Studien zur Leistungsentwicklung von Kindern sprechen dagegen, dass sich Lernschritte so verlässlich vorhersagen lassen. Der zukünftige Leistungsstand von verschiedenen Leistungsgruppen kann relativ gut prognostiziert werden, aber die individuelle Entwicklung *innerhalb* dieser Gruppen ist nicht vorhersagbar. Weder Schuleingangstests noch Screenings in bestimmten Leistungsbereichen noch breiter angelegte Längsschnittstudien konnten RisikoschülerInnen sicher vor-

⁵ „Leseuntersuchung mit dem Stolperwörter-Test“ mit rund 25.000 SchülerInnen in den Klassen 2 bis 4 → www.uni-siegen.de/~agprim/lust

hersagen⁶. So warnen auch die Autoren von PISA und IGLU, Baumert und Bos, vor dem Glauben, man könne die Übergangentscheidung ins dreigliedrige System der Sekundarstufe technisch perfektionieren, z. B. indem man das Urteil der LehrerInnen durch punktuelle Tests ersetzt.

Das heißt aber nicht, LehrerInnen könnten sich zurücklehnen und allein auf ihre Intuition vertrauen. PISA und IGLU bestätigen, was wahrnehmungspsychologische Studien immer wieder zeigen: Auch die Urteile von Fachleuten unterliegen Verzerrungen. Bei der Bewertung von Schülerleistungen spielen soziale Herkunft und Migrationshintergrund eine große Rolle. Manchmal werden schwache Leistungen nicht erkannt oder gute nicht anerkannt. Vor allem aber unterscheiden sich die Maßstäbe, nach denen (zutreffend festgestellte) Leistungen *bewertet* werden, erheblich.

Aus allen diesen Gründen ist es wichtig, dass LehrerInnen *neben* ihren kontinuierlichen Beobachtungen und der Auswertung informeller Aufgaben auch standardisierte Tests einsetzen, um ihre Urteilsbasis zu erweitern. Solche Verfahren müssen inhaltlich aussagekräftig sowie in der Durchführung robust und unaufwändig sein. In unserer Studie LUST haben wir gute Erfahrungen mit dem Stolperwörter-Lesetest (s. Anm. 4) gemacht. Die Durchführung der Aufgabe kostet nur wenige Minuten, sie ist leicht auszuwerten und kann vor allem vom Ende Klasse 1 bis zu Erwachsenen hin eingesetzt werden. Die Originalaufgabe kann abgerufen werden über www.lesetest1-4.de⁷.

Wer auf diese Weise die Leistungsentwicklung seiner SchülerInnen systematisch im Blick behält, kann ihnen im Unterricht mehr Freiraum gewähren als in Deutschland sonst üblich⁸. IGLU-E belegt das verbreitete enge Verständnis von Leseförderung mit dem Befund, dass „...der Leseunterricht in weiten Tei-

⁶ Vgl. Brügelmann, H. (2003): Fördern: nur das Reparieren einer unzulänglichen Technik? In: Grundschule, 35. Jg., H. 10, 54-57. s. a. www.uni-siegen.de/~agprim → "Projekte" → "LOGIK"

⁷ Inzwischen arbeiten wir an einer differenzierteren Version des Stolperwörter-Lesetests, von dessen Profilen wir uns mehr Aufschluss über die Entwicklung der Leseleistung erhoffen, vgl. Brügelmann, H./ Backhaus A. (2003): Leseentwicklung von Grundschulkindern in verschiedenen Leistungsgruppen. Lese-Untersuchung mit dem Stolperwörter-Test (LUST-3). Antrag an die Deutsche Forschungsgemeinschaft. FB 2 der Universität: Siegen.

⁸ Brügelmann, H. (2000): Wie verbreitet ist offener Unterricht? In: Jaumann-Graumann, O./ Köhnlein, W. (Hrsg.) (2000): Lehrerprofessionalität --Lehrerprofessionalisierung. Jahrbuch Grundschulforschung, Bd. 3. Klinkhardt: Bad Heilbrunn, S. 133-143.

len im Klassenverband und eng an didaktisch vorgefertigten Materialien durchgeführt [wird]. Insgesamt wird weniger gelesen, vor allem auch deutlich weniger durch die Lehrkraft vorgelesen als in Staaten wie Schweden oder England. Auch individuelle Förderung wird seltener berichtet.“ Hier ist der Ansatzpunkt für eine Reform des Unterrichts zu suchen, die Lesen als individuelle und persönlich bedeutsame Aktivität in einer anregenden Schreib- und Lesekultur des Klassenzimmers entwickelt⁹.

⁹ Bambach, H. (1989): Erfundene Geschichte erzählen es richtig. Lesen und Leben in der Schule. Libelle Verlag: CH-Lengwil (2. Aufl.).