

Geringfügig überarbeitet für NeSa (GB/ 6.1.05 & Brü 8.1.05)

Text: 37.650 (incl. Leerzeichen)

Literatur (kleiner zu setzen): 6877 (incl. Leerzeichen)

Anmerkungen/Fußnoten (kleiner zu setzen): 8.390 (incl. Leerzeichen)

Der Karawaneneffekt

Eine Zwischenbilanz des Projekts LUST zum Lesenlernen¹

Von Hans Brügelmann

1. Ziele und Kontext des Projekts LUST-1

Am 8. April 2003 wurden die Ergebnisse der Internationalen Grundschul-Lese-Untersuchung veröffentlicht. Von einem Tag auf den anderen entzogen die Befunde von IGLU dem bildungspolitischen Konsens der Diskussion nach PISA den Boden: Waren die 15-Jährigen bei PISA noch im unteren Drittel des internationalen Vergleichs gelandet, erreichten die 10-Jährigen einen respektablen 5.-11. Platz². Monatlang hatten Ministerien, Parteien und Verbände einen Vorschlag nach dem anderen aus dem Boden gestampft, was in Vor- und Grundschule geändert werden müsste, um das „blamable Abschneiden“ bei PISA auszumerzen. Mit IGLU konnte die ebenso weit verbreitete wie abenteuerliche „Logik“ außer Kraft gesetzt werden: Wenn 15-Jährige nicht gut genug lesen können, dann muss sich der Anfangsunterricht ändern, denn Lesen lernt man im ersten und evtl. noch im zweiten Schuljahr. Dass Lesenlernen ein lebenslanger Prozess ist und dass zum Lesenlernen mehr gehört als Wörter zu entziffern und geübte Texte sinngerecht intoniert vorzutragen, dämmerte Vielen erst, als deutlich wurde, dass bereits Viertklässler Texte nutzen können, um komplexe Fragen und Aufgaben inhaltlich zu bearbeiten. Wozu neben IGLU also noch LUST?

In der aktuellen Diskussion über Evaluation von Schule und Unterricht dominieren Studien zum *System Monitoring*. Das Ziel von TIMSS, PISA und IGLU ist eine Bestandsaufnahme „des

¹ LUST steht für: Lese-Untersuchung mit dem Stolperwörter-Test (s. Anm. 6). Die Studie LUST-1 wurde durchgeführt und ausgewertet von Hans Brügelmann, Siegen, in Kooperation mit Wilfried Metze, Berlin (Autor der Tests) und mit Erika Brinkmann, Schwäbisch Gmünd (Parallelstudie Baden-Württemberg). Die Auswertung von LUST wurde finanziell gefördert von der Zukunftsstiftung Bildung, Bochum, und vom Grundschulverband – Arbeitskreis Grundschule e.V., Frankfurt. Beiden Förderern danken wir für sehr rasche und unbürokratische Hilfe, als uns der unerwartet hohe Rücklauf aus den Schulen zu überwältigen drohte. Die Nachfolge-Studien LUST-2 (2003/2004) und LUST-3 (2004-2006), die echte Längsschnitte enthalten werden, sind zur Zeit in Vorbereitung (Koordination: Axel Backhaus). Eine Vorfassung ist erschienen in: Grundschulverband aktuell Nr. 84 (November 2003, 19-25). Ein ausführlicherer Bericht mit differenzierten Tabellen ist zum Download zugänglich unter → www.uni-siegen.de/~agprim/lust/index.htm. Ganz herzlich danke ich Katrin Belz, Sina Gerlach, Oliver von Keutz, Steffi Maxa, Verena Reich, Sara Roth, Markus Spannan und Petya Todorova, die gemeinsam mit mir das fast endlose Geschäft des Kodierens übernommen haben.

² Vgl. Bos u. a. (2003).

Systems“. Adressaten sind die Bildungspolitik und andere Funktionsträger, die *allgemeine* Entscheidungen zu treffen haben.

Um die Qualität von Unterricht in der einzelnen Klasse zu verbessern, brauchen Lehrerinnen und Lehrer vor Ort aber Informationen über die Leistung ihrer Klasse. Der Grundschulverband hat deshalb bereits vor einigen Jahren ein mehrstufiges Modell der Rechenschaftspflichten und der Evaluationsformen vorgeschlagen (vgl. Bartnitzky u. a. 1999). In diesem Modell hat eine interne Evaluation, die externe Kennwerte zur Einordnung der eigenen Daten nutzt, Vorrang (vgl. zur Begründung ausführlicher Brügelmann 2000; 2005; Rolff 1999). Dieser Anspruch hat Folgen für Design und Methoden der Untersuchung:

- Statt landesweit repräsentativer Stichproben sind Vollerhebungen von ganzen Klassen in einer Schule bzw. Region erforderlich.
- Statt zeitlich und methodisch hohen Aufwands für Durchführung und Auswertung werden robuste und unaufwändige Verfahren benötigt, die von den Lehrern und Lehrerinnen im Alltag mit Gewinn für ihre Alltagsarbeit eingesetzt werden können.

Da die Schulministerien zur Zeit Großstudien favorisieren³ haben der Grundschulverband und die Zukunftsstiftung Bildung ein Projekt der Universität Siegen unterstützt, in dem in drei Schulamtsbezirken Nordrhein-Westfalens⁴ allen LehrerInnen der 2., 3. und 4. Klassen ein einfacher Lesetest⁵ und – für die Erstdurchführung – auch dessen Auswertung angeboten wurde.

Im Folgenden werden die Daten von 18.083 Kindern der Klassen 2 (6.039), 3 (6.198) und 4 (5.846) berichtet. Ihre Lehrerinnen und Lehrer haben den Test im Januar bzw. Februar 2003 durchgeführt, also zur Schuljahresmitte⁶. Die Leistung im Test kann zusammenfassend durch die Zahl der richtig bearbeiteten Sätze beschrieben werden („Rohpunkte“ in der Anweisung des Testautors Wilfried Metze). In dieses Maß gehen zwei zentrale Teilleistungen ein: die Geschwindigkeit (Zahl der *bearbeiteten Sätze*) und die Genauigkeit (Zahl der *Fehler*). Da die Schüler und Schülerinnen auf den einzelnen Jahrgangsstufen unterschiedlich viel Zeit hatten, um die 60 Sätze der Aufgabe zu bearbeiten, sind die Ergebnisse nicht direkt vergleichbar. Wir haben deshalb zusätzlich errechnet, wie viele Sätze die Schüler im Mittel *pro Minute richtig* bearbeitet haben. Dies ist der zentrale Kennwert der folgenden Auswertungen. Da er hoch

³ Vgl. neben TIMSS und PISA speziell für den Grundschulbereich IGLU.

⁴ Ich danke den beteiligten Schulämtern für ihre Unterstützung, insbesondere meinen Kontaktpersonen SAD'in Karin Brügelmann (Rhein-Sieg-Kreis), SAD'in Heidemarie Goßmann (Märkischer Kreis) und SAD Wolf Kuhnke (Kreis Siegen-Wittgenstein) sowie den Schulleiterinnen, den Lehrern und Lehrerinnen, die in unerwartet großem Umfang an der Untersuchung und zum Teil mit zusätzlichem persönlichen Engagement an der Auswertung mitgewirkt haben.

⁵ Dieser Test ist von Wilfried Metze, Berlin, für 1. bis 4. Klassen entwickelt worden (vgl. zu den Aufgaben, ihrem Hintergrund und ersten Ergebnissen → www.wilfriedmetze.de/lesetest oder www.lesetest1-4.de). Gefordert ist, in 60 einzelnen Sätzen mit 4 -15 Wörtern jeweils ein nicht passendes Wort zu streichen. In einer Reihe von Zusatzstudien haben wir sowohl für die Reliabilität (.80**-.90**) als auch – durch die Korrelation mit Wort-, Satz- und Textverständnis-Tests – für die Validität (.60** bis .90**) sehr zufriedenstellende Werte erhalten (vgl. Backhaus u. a 2004).

⁶ Insgesamt haben über 25.000 SchülerInnen an der Erhebung teilgenommen, ein Teil aber außerhalb des Zeitraums Januar/ Februar 2003 oder aus anderen Klassenstufen.

korreliert mit anderen Maßen des Wort-, Satz- und Textlesens korreliert, ist er ein guter Indikator für die grundlegende Lesefähigkeit (vgl. Backhaus u. a. 2004).

2. Ergebnisse der Jahrgänge im Vergleich⁷

Zwei Bezugswerte sind zur Einschätzung der Daten über die Grundschulzeit hinweg interessant: zum einen das Niveau von SchülerInnen am Ende der ersten Klasse, sozusagen als „Ausgangswert“ von Leseanfängern, zum anderen die Leistungen von lesekundigen Erwachsenen, sozusagen der „Zielwert“. Für Daten zu den ersten Klassen konnten wir auf eine Erhebung von Metzke am Ende des ersten Schuljahres zurückgreifen⁸. Um einen Maßstab für die Zielperspektive zu gewinnen, haben wir den Test mit Studierenden und Lehrerinnen und Lehrern durchgeführt⁹ – ein sehr anspruchsvoller Maßstab, wenn man bedenkt, dass immer noch weniger als 30% eines Jahrgangs die allgemeine Hochschulreife erwirbt. Die folgende Tabelle 2.1a zeigt die Bezugswerte im Überblick:

Altersstufe	Bearbeitete Sätze pro Minute	Anteil Fehler / bearbeitete Aufgaben	Richtige Sätze pro Minute
Ende 1. Klasse	2.2	19.7 %	1.8
Akademikerinnen / Akademiker	18.9	0.8 %	18.7

Vor diesem Hintergrund von „Ausgangs- und Zielwerten“ lassen sich die Leistungen der Zweit- bis Viertklässler nicht nur quantitativ, sondern auch qualitativ einschätzen.

2.1 Unterschiede in den Leseleistungen über die Grundschulzeit hinweg

Eine Übersicht über die erreichten Leistungen zu verschiedenen Zeitpunkten gibt die folgende Tab. 2.1b¹⁰:

	Bearbeitete Sätze	Anteil Fehler /	Richtige Sätze

⁷ Bei unserer Studie handelt es sich um drei gleichzeitig erhobene Querschnitte. Insofern kann man im strengen Sinne nicht von einer „Entwicklung“ (einzelner Kinder) sprechen. Die Daten zeigen aber, welches Niveau Schülergruppen verschiedenen Alters unter gleichen Bedingungen zu den genannten Zeitpunkten erreichen, so dass mit großer Plausibilität eine Durchschnittsentwicklung erschlossen werden kann. wie auch die ersten Befunde aus unseren Längsschnitten bestätigen, vgl. Brinkmann u. a. (2004).

⁸ Die Daten wurden errechnet aus Angaben in dem Bericht auf seiner Homepage www.lesetest1-4.de [Zugriff auf das Manuskript 10.8.2003]. Für das erste Schuljahr handelt es sich um 1.536 Kinder aus 72 Klassen.

⁹ Stand der Auswertungen von N = 181 : 5.7.2004

¹⁰ Alle hier und im Folgenden berichteten Unterschiede sind statistisch signifikant. Der Übersichtlichkeit halber sind die statistischen Kennwerte nur in den Volltabellen im technischen Tabellenanhang enthalten (Download s. Anm. 1). Auch unter Berücksichtigung der Standardfehler sind alle Mittelwertunterschiede hoch signifikant.

Jahrgangsstufe	pro Minute	bearbeitete Aufgaben	pro Minute
Mitte 2. Klasse	4.8	12.7 %	4.1
Mitte 3. Klasse	6.6	6.3 %	6.1
Mitte 4. Klasse	8.7	4.5 %	8.1

Mitte 2. Klasse ist in den vielen Klassen der Leselehrgang „abgeschlossen“, wie es in den meisten Lehrplänen heißt. Dies ist der früheste Termin unserer Erhebungen (Januar/Februar 2003).

Die **Zweitklässler** schaffen im Durchschnitt in sechs Minuten 25 Sätze *richtig*. Damit verdoppeln sie nur sechs Monate später das Ergebnis der Vergleichskinder Ende erster Klasse, sie brauchen aber vier- bis fünfmal so lange wie die von uns getesteten Erwachsenen. Wie auch die Fehlerquote zeigt, liegt für den durchschnittlichen Zweitklässler die Hauptschwierigkeit nur noch teilweise im Lesen-*Können*: Knapp 13% der Sätze werden falsch bearbeitet, gegenüber immerhin noch fast 20% Ende erste Klasse. Die besondere Anforderung liegt also in der *Lesegeschwindigkeit*.

Drittklässler bearbeiten in fünf Minuten¹¹ durchschnittlich 30 Sätze *richtig*. Sie brauchen etwa ein Drittel weniger Zeit als die Zweitklässler pro Satz. Zusätzlich halbiert sich die Fehlerquote. Damit bearbeiten sie pro Minute fast 50% mehr Sätze richtig als die ZweitklässlerInnen.

Im Durchschnitt bearbeiten die **Viertklässler** unserer Stichprobe fast 9 Sätze pro Minute. Sie erreichen damit schon die Hälfte dessen, was erwachsene Vielleser bewältigen. Zugleich sinkt die Fehlerquote noch einmal um ein Drittel, so dass am Ende der vierten Klasse durchschnittlich 8 Sätze pro Minute *richtig* bearbeitet werden – gegenüber 18 richtigen Sätzen bei den mindestens 15 Jahre älteren Erwachsenen. Für die im „Stolperwörter-Test“ geprüfte *grundlegende* Leseleistung ist das ein gutes Ergebnis, zumal wenn man die Differenz von rund drei Jahren Lesepraxis der Grundschülerinnen und Grundschüler vs. 20 bis 50 Jahren bei den Erwachsenen bedenkt.

Fazit zu 2.1 :

Im Durchschnitt brauchen Kinder der 2. Klassen 15 sek. pro *richtig* bearbeitetem Satz. Drittklässler benötigen nur noch 10 sek. und Viertklässler im Durchschnitt 7-8 sek. Bezogen auf die durchschnittlich 3 sek., die von akademisch gebildeten LeserInnen benötigt werden, bedeutet das einen rapiden Lernfortschritt in nur zwei Schuljahren und ein hohes Niveau grundlegender Teilleistungen des Lesens zum Ende der Grundschule.

Dabei verdoppelt sich die Lesegeschwindigkeit von Klasse 2 bis Klasse 4 (nur noch 7 sek. statt 14 sek. pro bearbeitetem Satz). Zusätzlich schrumpft die Fehlerquote

¹¹ Nach der Testanweisung des Autors Wilfried Metzke wird den Kindern von Klassenstufe zu Klassenstufe weniger Zeit zur Bearbeitung des Tests eingeräumt.

von rund 13% auf 4-5 %.

2.2 Vergleich der oberen und der unteren Leistungsgruppen

Die Mittelwerte geben nur einen sehr verdichteten Eindruck vom Leistungsniveau auf den verschiedenen Klassenstufen. Nicht weniger wichtig ist die Bandbreite, innerhalb derer die Leistungen von diesen Mittelwerten abweichen. Die hohen Standardabweichungen zeigen, dass die Leistungen auf den einzelnen Jahrgangsstufen erheblich streuen.

Alle Leistungsgruppen werden in der untersuchten Aufgabe von Jahr zu Jahr besser, wobei die Fortschritte (in absoluten Zuwächsen richtiger Sätze pro Minute) in den oberen Prozentranggruppen etwas höher sind.

Dieses allgemeine Ergebnis soll zunächst in einzelnen Schritten für die drei Jahrgänge entfaltet werden. Dabei können zusätzlich die Fehlerquoten berücksichtigt werden.

In den **zweiten** Klassen bewältigen die besten fünf Prozent mindestens zehnmal so viele Aufgaben wie die schwächsten fünf Prozent. Beachtlich ist, dass die besten 15% Ende der zweiten Klasse nicht weit vom Lesetempo der Erwachsenen entfernt sind. Für die Grundleistungen „rasches und genaues Wortlesen“ sowie „Integration der Wortbedeutungen auf Satzebene“ scheint in dieser Teilgruppe der Mastery-Level bereits fast erreicht. Die leistungsstärksten Zweitklässler könnten unter diesem Aspekt selbst im vierten Jahrgang noch im oberen Leistungsdrittel mithalten! Abgesehen von einer weiteren Steigerung der Geläufigkeit durch häufiges Lesen werden diese Kinder eher in anderen Dimensionen bzw. auf anderen Ebenen der Lesefähigkeit (Textverständnis) Fortschritte machen müssen.

In **Klasse 3** ist der absolute Zuwachs in beiden Extremgruppen mit zwei Sätzen pro Minute etwa gleich groß. Relativ gesehen haben sich die besten fünf Prozent allerdings nur wenig gegenüber der vergleichbaren Gruppe ein Jahr vorher verbessert (Zuwachs um 10-20%). Schüler und Schülerinnen auf den Prozenträngen 5-15 lesen jetzt (mit 16-23 sek. pro richtigem Satz) fast genauso schnell und genau wie der Durchschnitt der Zweitklässler (15 sek.). Dazu dürfte neben einem höheren Lesetempo vor allem die Halbierung der Fehlerquote in dieser Gruppe beigetragen haben. Der Rückstand der unteren 15% ist damit qualitativ geschrumpft, wenn man einmal von den allerschwächsten Leserinnen und Lesern absieht.

Und die Streuung in **Klasse 4**? Relativ gesehen schaffen die oberen 15% jetzt nur noch viermal so viele Aufgaben wie die unteren 15% (gegenüber einer Relation von 10:1 in Klasse 2). Absolut gesehen hat sich der Abstand aber nicht wesentlich verändert.

Wieder sind die oberen 15% um etwa 20% besser als ein Jahr vorher. Sie erreichen sogar das Niveau der Erwachsenen in unserer kleinen Zusatzerhebung. Zwei Drittel der Stichprobe liegen zwischen 5 und 11 Sekunden pro richtigem Satz. Das ist eine deutliche Differenz – aber die Leistungsunterschiede sind gegenüber der dritten und erst recht der zweiten Klasse qualitativ

nicht mehr so bedeutsam. Das zeigt vor allem die erneute Abnahme der Fehlerquote im unteren Prozentrangbereich.

Vergleicht man die äußersten Extreme, lässt sich das Bild noch einmal präzisieren. Mit den oberen bzw. unteren 5% erfassen wir näherungsweise „den besten“ bzw. „den schlechtesten“ Schüler einer durchschnittlichen Klasse.

„Der schwächste Schüler“ einer durchschnittlichen 4. Klasse schafft mit 3,3 richtigen Sätzen pro Minute nur ein Fünftel der Leistung seines besten Mitschülers. Er bearbeitet zwar knapp vier Sätze pro Minute. Damit benötigt er nur drei- bis viermal so lange wie der beste Mitschüler. Gleichzeitig macht er aber neunmal so viele Fehler (15.2 vs. 1.7).

Andererseits erscheint eine Quote von 15% falschen Urteilen in der Gruppe besonders leistungsschwacher Schüler und Schülerinnen vergleichsweise gering, wenn man an die dramatischen Meldungen über angeblich wachsende „Analphabeten“- und „Legastheniker“-Zahlen denkt. Dennoch zeigen sich auf beiden Indikatoren – Lesetempo und Lesegenauigkeit – deutliche Unterschiede zwischen den Leistungsextremen. Deren alltags- und unterrichtspraktische Bedeutung muss noch gesondert diskutiert werden (s. dazu unten 3).

Fazit zu 2.2 :

Die leistungsstärksten 15% brauchen bereits Mitte der 2. Klasse nur 8-9 sek. für einen richtigen Satz, erreichen also zu diesem Zeitpunkt schon das Durchschnittsniveau der 4. Klasse. Ihre Leistung verbessert sich bis Mitte 3. Klasse um 10-20% auf 6-7 sek., bis Mitte 4. Klasse um weitere 20% auf 4.5-5.5 sek. Damit erreicht diese Teilgruppe bereits vor dem Ende der Grundschulzeit das Niveau der akademisch gebildeten Leser und Leserinnen. Diese Kinder dürften in einem herkömmlichen Leseunterricht massiv unterfordert sein.

Die leistungsschwächsten 15% brauchen Mitte der 2. Klasse mehr als 30 sek., um einen Satz zu erlesen und das nicht passende Wort zu finden, teilweise schaffen sie nur Zufallstreffer (0-2 Sätze in 6 min.). Auffallend ist die hohe Streuung selbst innerhalb dieser Teilgruppe. Damit brauchen diese Jungen und Mädchen das 5- bis mehr als 10-fache der Zeit, die die leistungsstärksten aufwenden. Mitte 3. Klasse benötigen die unteren 15% mindestens 18 sek., Mitte 4. Klasse mindestens 11 sek. Allerdings gibt es immer noch Kinder, die nur Zufallstreffer landen. Damit nimmt die Streuung *innerhalb* der Teilgruppe der unteren 15% von Jahr zu Jahr zu.

Die schwächsten 5% bearbeiten auch am Ende der Grundschulzeit noch jeden siebten Satz falsch, außerdem brauchen sie drei- bis viermal so viel Zeit wie die leistungsstärksten 5% .

3 Die Entwicklung der Leistungen im unteren Bereich

Die durchschnittlichen Leistungen und erst recht das Niveau der Spitzengruppe werfen nach den bisher berichteten Befunden keine besonderen Probleme auf. Genauer zu prüfen sind aber die Ergebnisse im unteren Leistungsbereich. Damit stellt sich die Frage nach den sog. „Mindeststandards“. Konkret: wie viele Sätze muss ein Kind am Ende der Grundschulzeit pro Minute richtig bearbeiten, damit man von einer „tragfähigen Grundlage“ für seine weitere Entwicklung im Lesen sprechen kann?

3.1 Die Leistungen der zwei, drei schwächsten SchülerInnen pro Klasse

Solange es keine Längsschnitte über die Entwicklung von Schülerinnen und Schülern auf verschiedenen Niveaus in den anschließenden Schuljahren gibt, kann jede Entscheidung zwischen verschiedenen Leistungsanforderungen am Ende der Grundschule nur Plausibilitätscharakter beanspruchen. Dies umso mehr, als aus der Forschung sehr unterschiedliche Entwicklungsverläufe des Schriftspracherwerbs bekannt sind (vgl. u. a. Brinkmann 1997).

Damit sich Leserinnen und Leser ein eigenes Urteil bilden können, berichten wir im Folgenden Daten zu vier verschiedenen Kriterien, die unterschiedlich hohe Anforderungen an eine „grundlegende Lesefähigkeit“ definieren:

- 0-2 Sätze/min. kann nicht lesen (30-60% falsch, richtige Lösungen allenfalls durch Zufallstreffer bei einer Ratewahrscheinlichkeit von durchschnittlich 10-15% pro Satz);
- 2-3 Sätze/min. kann Sätze zwar eigenständig, aber nur sehr langsam (20-30 sek. pro Satz) und nur mit großen Schwierigkeiten (etwa 20-30% Fehler) erlesen;
- 3-4 Sätze/min. kann Sätze erlesen, aber immer noch langsam (15 sek.) und mühevoll (im Durchschnitt noch mehr als 10% Fehler);
- 4-6 Sätze/min. kann Sätze zunehmend fehlerfrei (6-8%) und zügiger, aber auch immer noch relativ langsam (etwa 12 sek. pro Satz) erlesen.

In der folgenden Tabelle haben wir die Anteile der Jahrgänge in allen vier Kategorien ausgewiesen, so dass man sich je nach Maßstab ein eigenes Bild vom Niveau der Schüler und Schülerinnen im unteren Leistungsbereich machen kann:

Tab. 3.1a: Jahrgangvergleich der Anteile (Rohdaten) im unteren Leistungsbereich				
richtige Sätze pro Minute Jahrgang	0-2	0-3 ¹²	0-4	0-6
2. Klasse	14.8 %	32.1 %	51.4 %	80.5 %
3. Klasse	1.5 %	6.8 %	17.4 %	53.0 %
4. Klasse	0.1 %	1.2 %	4.0 %	20.3 %

Diese Übersicht vermittelt auf den ersten Blick ein sehr positives Bild: Die Anteile in den unteren Leistungsgruppen nehmen von Jahrgang zu Jahrgang rapide ab. Die Gruppe der Leseunfähigen liegt schon Mitte 3. Klasse, bei einem härteren Maßstab Mitte 4. Klasse unter 2% .

Allerdings ist daran zu erinnern, dass wir es nicht mit einem echten Längsschnitt zu tun haben. So können wir nicht davon ausgehen, dass die drei Querschnitte in ihrer Zusammensetzung identisch sind. Da es sich um Erhebungen in denselben Schulbezirken handelt, können wir zwar davon ausgehen, dass die Zusammensetzung der Schülerschaft auf den verschiedenen Jahrgangsstufen vergleichbar ist. Wir haben aber beim Vergleich der unteren Leistungsbereiche selektive Abgänge durch Sitzenbleiben und durch den Wechsel auf Sonderschulen zu berücksichtigen. Zur Überprüfung dieser Unsicherheit sind unbedingt echte Längsschnittstudien erforderlich, die wir zur Zeit in Nordrhein-Westfalen und Baden-Württemberg durchführen

Vorläufig kann man mit folgenden Schätzwerten arbeiten: pro Jahr bleiben in NRW in der Grundschule etwa 1.5% der Schülerinnen und Schüler sitzen (ddp 2003). Da es sich dabei um besonders leistungsschwache Kinder handelt, muss statistisch gesehen der Anteil der leistungsschwachen Gruppe in Klasse 2 gegenüber Klasse 4 um bis zu 3% höher angesetzt werden. Zusätzlich wechseln in der Grundschulzeit etwa 2% in die Schulen für Lernbehinderte, so dass insgesamt rund 5% „Dropouts“ in den unteren Leistungsgruppen zu vermuten sind und die dortigen Werte entsprechend zu erhöhen sind, um einen fairen Längsschnitt zu konstruieren. Die Verteilung der Korrekturwerte auf die Kategorien im unteren Leistungsbereich muss nach persönlicher Einschätzung erfolgen. Sie ist so vorgenommen worden, dass sie das Ausmaß der Leseschwierigkeiten eher über- als unterschätzt.

¹² Die Werte sind kumulativ berechnet, die höheren schließen also die jeweils vorhergehende Gruppe mit ein. Als „Lesehilfe“: In der Klasse 3 erreichen 1,5% nur die Stufe 0-2, 6,8% nur die Stufe 2-3, 53,0% nur die Stufe 4-6. Das heißt: 47,0% erreichen bessere Leistungen als Stufe 4-6.

Tab. 3.1b: Jahrgangvergleich der Anteile (korrigiert) im unteren Leistungsbereich				
richtige Sätze pro Minute Jahrgang	0-2	0-3 ¹³	0-4	0-6
2. Klasse	14.8 %	32.1 %	51.4 %	80.5 %
3. Klasse	2.5 %	8.8 %	19.9 %	55.5 %
4. Klasse	2.1 %	5.2 %	9.0 %	29.3 %

In dieser Übersicht wird das Bild differenzierter. Zum einen ist immer noch erkennbar, dass der Anteil der leistungsschwachen Schüler und Schülerinnen deutlich abnimmt, in den untersten Kategorien schon von Klasse 2 auf 3, bei schärfer formulierten Anforderungen (nicht mehr als 12 sek pro Satz) auch bis Klasse 4. Gleichzeitig ist aber festzustellen, dass im untersten Bereich (0-2) von Klasse 3 auf 4 mit einer Stagnation zu rechnen ist, die für die betroffenen Kinder erhebliche Schwierigkeiten in der Sekundarstufe befürchten lässt. Erste Befunde aus einer Ergänzungsstudie von Venn-Brinkmann in niedersächsischen Klassen 9 und 10 belegen dieses Problem für die SchülerInnen und Schüler der Sonderschule für Lernbehinderte.

Aber: Alle Gruppen machen Fortschritte¹⁴, und proportional am stärksten die unteren Leistungsgruppen. Das bedeutet: Das Verhältnis der Leistungen der oberen zu den unteren 15% beträgt in Klasse 2 noch 1:9.4, in Klasse 4 dagegen nur noch 1:3.8 .

In der Perspektive der einzelnen Lehrperson „bleiben die Schwachen schwach“. Durch den Vergleich allein mit der Bezugsgruppe, die sich ja auch weiter entwickelt, werden aber die individuellen Fortschritte unterbewertet. Das nennen wir den „Karawaneneffekt“ des Lernens in Klassen¹⁵. Die unteren 5% oder 15% einer Gruppe sind definitionsgemäß *immer* schlechter als der Durchschnitt. Mit der Fixierung des Blicks auf ihren Platz in der Bezugsgruppe wird aber übersehen, dass *alle* Jungen und Mädchen von Jahr zu Jahr Fortschritte machen. Pädagogisch gesehen sind die *Fortschritte* jeder *Teilgruppe* bedeutsamer als die *Abstände* innerhalb der *Gesamtgruppe*. Konkret: Wer in der vierten Klasse auf Prozentrang 15 pro Minute 5.2 Sätze

¹³ Vgl. Fußnote 12!

¹⁴ Auch hier ist allerdings wieder zu betonen: Diese Folgerung hat solange nur hypothetischen Charakter, wie sie nicht durch entsprechende Belege aus einem echten Längsschnitt abgesichert ist. Erste Längsschnittdaten aus einer Teilstichprobe von 198 Kindern sprechen allerdings mit Korrelationen von durchschnittlich .83** pro Klasse schon jetzt dafür, dass sich Ränge der einzelnen Kinder kaum verändern.

¹⁵ Analog hat May (1995) für die Rechtschreibentwicklung festgestellt, dass sich die Fortschritte der leistungsstarken und der leistungsschwachen SchülerInnen auch qualitativ gleichen – wenn man sie auf die jeweilige Ausgangsposition bezieht. Am Vergleich von deutschsprachigen Schülern und Schülerinnen und Kindern anderer Muttersprache können wir zeigen, dass dieses Muster auch in der Sekundarstufe gilt (Sekundärauswertung der NRW-Kids-2001-Studie, Brügelmann 2003).

schaft, hat sich gegenüber Klasse 2 um 3.4 Sätze verbessert – sein Abstand zum Durchschnitt beträgt aber nur 2.9 Sätze. Die pädagogische Frage ist, welcher Bezugspunkt bei der Leistungsbewertung dominieren sollte.

Bei einer Karawane verwundert es niemanden, wenn der, der zuletzt auf die Reise gegangen ist, auch als letzter ankommt. Bedeutsamer ist der Weg, den die Karawane als *ganze* geschafft hat. Man muss bedenken, dass sich schon die Schulanfänger in ihren schriftsprachlichen Voraussetzungen bis zu drei, vier Jahren durchschnittlicher Entwicklung unterscheiden. Dann überrascht der Unterschied von rund drei Entwicklungsjahren in der Leistung eines Viertklässlers auf Prozentrang 10 (4.9 Sätze pro Minute; korrigiert um die Abgänge: 4.0) und einem auf Prozentrang 90 (11.7 Sätze pro Minute) nicht. Damit stellt sich aber die Frage, wie sinnvoll eigentlich die aktuell diskutierten „Bildungsstandards“ sind, wenn sie als verbindliche Niveaus für alle zum gleichen Zeitpunkt formuliert werden.

Mit diesem Hinweis sollen keinesfalls die in der untersten Leistungsgruppe beobachteten Schwierigkeiten verdrängt werden. Auch wenn die Größenordnung von rund 5% besonders gefährdeten Schüler und Schülerinnen¹⁶ nicht mit jenen 25% zu vergleichen ist, die von den PISA-Interpreten als „Risikogruppe“ bezeichnet werden, stimmt die Art der Probleme besorgt, da es sich beim Stolperwörter-Lesetest anders als bei IGLU (und erst recht bei PISA) um grundlegende Teilleistungen handelt.

Ebenso deutlich muss man andererseits sagen, dass Maßnahmen zur Behebung dieses Problems nicht Inhalt und Stil des Unterricht für die übrigen 95% bestimmen dürfen¹⁷. Diese Warnung ist deshalb wichtig, weil sich die in PISA in einer größeren Breite, d. h. bei ca. 25% der Schülerinnen und Schüler, beobachteten Schwächen auf das Textverständnis beziehen. Dass die Kinder der Grundschule auch in dieser Hinsicht relativ besser abschneiden, haben die IGLU-Ergebnisse gezeigt.¹⁸

3.2 Die Leistungsverteilung in Schulen in sozialen Brennpunkten

Da unsere drei Stichproben aus eher ländlichen und kleinstädtischen Regionen stammen, haben wir ergänzend eine großstädtische Zusatzstichprobe von 25 vierten Klassen aus dem Kölner Raum einbezogen.¹⁹ Bei dieser Sonderstichprobe handelt es sich um Schulen in sog. sozialen „Brennpunkten“, deren Schülerinnen und Schüler in der Regel ungünstigere

¹⁶ Auf die Altersgruppe insgesamt bezogen: plus vermutlich weiteren 5%, die den Jahrgang oder die Grundschule verlassen haben.

¹⁷ Mehr noch: Es kann aus unseren Ergebnissen auch für die gefährdete Teilgruppe der unteren 5% nicht geschlossen werden, eine effektive Förderung müsse sich auf einfache Teilleistungen konzentrieren. Intensivstudien (vgl. Peschel 2002) zeigen, dass die Förderung anspruchsvollen Lesens, also des selbstständigen Umgangs mit Texten vom ersten Schuljahr an, gerade bei schwachen Schülerinnen oder Schülern ein hohes Niveau auch in den Basisleistungen sichern kann.

¹⁸ Vgl. Bos u. a. (2003).

¹⁹ Ich danke den KollegInnen SAD'in Monika Baum und SAD Hans Wielpütz vom Schulamt der Stadt Köln für ihre rasche und unbürokratische Hilfe.

Lernvoraussetzungen und Unterrichtsbedingungen mitbringen und die häufig auch während der Schulzeit zusätzliche Belastungen zu verarbeiten haben.

Im Februar und März 2003, also einen Monat nach unserer Hauptstichprobe, wurden die Daten in 8 Schulen mit insgesamt 539 Kindern erhoben. Diese Teilstichprobe weicht in zwei zentralen Merkmalen tatsächlich von unserer Hauptstichprobe ab:

- der Anteil der Kinder von Alleinerziehenden ist doppelt so hoch (14.8% statt sonst 7.3%);
- der Anteil der Kinder mit anderer Muttersprache ist dreimal so hoch (64.2% statt 21.3% sonst) .

Erwartungsgemäß schneiden diese Klassen In den Durchschnittsleistungen (richtig bearbeitete Sätze) insgesamt schwächer ab:

- arithmetisches Mittel 7.6 (SD 2.5) vs. 8.7 (SD 2.7) in der Hauptstichprobe.

Anders als nach den Umfeldbedingungen erwartet sind allerdings die Leistungen im unteren Bereich nur wenig schlechter:

- 7.0 % der Kinder unter Prozentrang 5 der Gesamtstichprobe.
- 1.6 % haben weniger als 3 Sätze pro Minute richtig (vs. 1.2% in der Gesamtstichprobe);

Insbesondere ist festzuhalten:

- Die deutschsprachigen Kinder schneiden um weniger als 10% schlechter ab als in der Hauptstichprobe (8.1 vs. 8.7 richtige Sätze pro Minute), obwohl sie aus einem sozialen Brennpunkt kommen und in den Klassen mit einem Anteil von einem Drittel (gegenüber mehr als drei Vierteln in der Hauptstichprobe) die Minderheit darstellen.
- die Kinder anderer Muttersprache schneiden unter diesen Bedingungen mit durchschnittlich 7.2 richtigen Sätzen (SD 2.5; SE 0.15) sogar geringfügig besser ab als sonst (6.9 richtige Sätze [SD 2.4; SE 0.08] in der Hauptstichprobe).

Der schlechtere Gesamtwert dieser Sonderstichprobe ergibt sich also nicht aus schlechteren Leistungen ihrer Teilgruppen. Er ist vielmehr rein rechnerisch bedingt und resultiert daraus, dass die in *beiden* Stichproben schwächer abschneidenden Migrantenkinder in den Brennpunkt-Klassen einen deutlich höheren Anteil darstellen (ausführlicher unten 4.2). Bedenkt man die zusätzlichen sozio-ökonomische Belastungen in den sog. „Brennpunkten“, fällt selbst der Unterschied in der deutschsprachigen Gruppe erstaunlich gering aus.

Diese Befunde sind ein wichtiger Hinweis darauf, dass es einer Reihe von Grundschulen gelingt, Kindern mit anderer Muttersprache auch unter schwierigen Bedingungen gerecht zu werden. Dabei könnten objektive Rahmenbedingungen (z. B. die Anzahl der Förderstunden), aber auch pädagogische und didaktische Konzepte der Lehrerinnen und Lehrer eine bedeutsame Rolle spielen.

Fazit zu 3:

Aus den bisher berichteten Befunden folgt: der Abstand zwischen leistungsschwachen und leistungsstarken Schülerinnen oder Schülern verringert sich von Klassenstufe zu Klassenstufe deutlich – von etwa 1:8 in Klasse 2 auf 1:3 in Klasse 3 und Klasse 4.

Dieser Befund darf aber nicht als „Aufhol“-Effekt missverstanden werden, da rund 5% der Zweitklässler durch Nichtversetzung oder Wechsel in die Sonderschule den Anteil leistungsschwacher LeserInnen verringern. Berücksichtigt man diese rein „statistische Verbesserung“ speziell der untersten Teilgruppen, so öffnet sich die Schere zwischen Leistungsstarken und Leistungsschwachen über die Grundschulzeit hinweg (sog. „Matthäus-Effekt“: *Wer hat, dem wird gegeben*).

Andererseits machen *alle* Gruppen deutliche Fortschritte. Bedenkt man, dass sich bereits Schulanfänger in ihren schriftsprachlichen Voraussetzungen um rund 3-4 Jahre durchschnittlicher Entwicklungszeit unterscheiden, überrascht nicht, dass auch Mitte 4. Klasse die leistungsstarken Grundschülerinnen und Grundschüler auf Prozentrang 90 den leistungsschwachen auf Prozentrang 10 um rund drei Jahre durchschnittlicher Leseentwicklung voraus sind²⁰. Wir interpretieren beide Befunde zusammen als „Karawaneneffekt“: Im Vergleich zum Durchschnitt bleiben die Leistungsschwachen am Ende der Gruppe – bezogen auf ihre je unterschiedlichen Voraussetzungen machen sie bedeutsame Fortschritte. Das gilt allerdings nicht für die leistungsschwächsten 5-10%. Diese bleiben immer mehr zurück – vermutlich auch eine Folge der ständig entmutigenden Rückmeldung durch eine vergleichende Leistungsbewertung und der fehlenden Förderung von ihrem individuellen Entwicklungsstand aus.

4 Muttersprache der Eltern

Die Muttersprache der Eltern ist ein besserer Indikator für abweichende Sprachvoraussetzungen als die früher erhobene Staatsangehörigkeit. Wie in anderen neueren Studien auch fasst das Merkmal „Sprache der Eltern“ die Sprachlernsituation der Kinder spezifischer als rechtliche Kategorien. Auf diese Weise werden auch Kinder aus „deutschen“ Aussiedlerfamilien und ebenso eingebürgerte Kinder aus Arbeitsmigrantenfamilien erfasst.

²⁰ ... sofern man Lernen nicht als einen mechanischen Prozess missversteht, in dem sich Entwicklungszeit beliebig verkürzen lässt, wenn man die Intensität des Unterrichts entsprechend steigert. Insbesondere ist zu bedenken, dass entwicklungsförderliche bzw. hinderliche Bedingungen auch parallel zur Schule wirken. Eine US-amerikanische Studie zeigt, dass leistungsstarke Leserinnen und Leser außerhalb der Schule rund 100-mal so viel Text lesen wie leistungsschwache (Anderson u. a. 1988). Diesen beiläufigen Übungsvorteil kann kein Förderunterricht ausgleichen. Vielmehr kommt es darauf an, auch die leseschwachen Schüler und Schülerrinnen „zum Lesen zu verlocken“ – in ihrer Freizeit.

4.1 Grunddaten für die Klassenstufen 2 bis 4

	Klasse 2	Klasse 3	Klasse 4
beide Eltern Muttersprache Deutsch ²¹	4.3**	6.3**	8.5**
ein Elternteil andere Muttersprache ²²	2.9	5.1	7.5
beide Eltern andere Muttersprache	3.1	4.8	6.9**

Mitte 2. Klasse macht die Muttersprache der Eltern einen beträchtlichen Unterschied für die Lesefähigkeit aus: Kinder aus Elternhäusern, in denen eine andere Sprache gesprochen wird, schaffen mit 3.1 bzw. 2.9 vs. 4.3 deutlich weniger richtige Sätze als die deutschsprachigen Kinder (auch statistisch nicht nur ein signifikanter, sondern auch ein quantitativ bedeutsamer Unterschied: $ES^{23} = .55$).

Der Unterschied resultiert sowohl aus einem langsameren Lesetempo (4.1 vs. 5.0 Sätze) als auch aus einer höheren Fehlerquote (21% vs. 11% der bearbeiteten Sätze falsch).

Rund 30% Differenz bedeuten einen erheblichen Abstand. Die umgerechnet 1.2 Sätze pro Minute entsprechen einer durchschnittlichen Entwicklung von etwa einem halben Schuljahr.

Die Unterschiede sind ein Jahr später, also Mitte der **3. Klasse**, noch deutlicher (auch qualitativ bedeutsam: $ES = .68$). In absoluten Zahlen entsprechen sie weiterhin etwa einem halben Schuljahr, aber relativ sind sie von rund 30% auf etwa 20% geschrumpft:

In beiden Gruppen hat sich die Fehlerquote gegenüber den Zweitklässlern etwa halbiert und das Arbeitstempo deutlich beschleunigt.

Kinder mit Eltern, die eine andere Muttersprache sprechen, lesen auch am Ende der Grundschulzeit schlechter als deutschsprachig aufgewachsene Kinder: Sie brauchen Mitte der **4. Klasse weiterhin** gut 20% länger für einen richtigen Satz ($ES = .59$).

²¹ 78-80% der Gesamtgruppe in den Jahrgängen 2 bis 4

²² Die Gruppe der Migranten und Migrantinnen ist in den Tabellen zwar detaillierter aufgegliedert, aber die Teilgruppe der Kinder mit nur einem Elternteil anderer Muttersprache macht je nach Jahrgang nicht mehr als 4-8% der Kinder mit Migrationshintergrund aus, so dass beide Teilgruppen in der Interpretation zusammengefasst werden, zumal sich ihre Leistungen nur in Klassenstufe 4 statistisch signifikant unterscheiden.

²³ Die Effektstärke („ES“) wird berechnet aus der Differenz der Mittelwerte, geteilt durch die Streuung der Kontrollgruppe – hier der deutschsprachigen Kinder.

Auf den ersten Blick ist das Bild analog wie bei den leseschwächeren Kindern insgesamt: Auch die Migrantenkinder machen deutliche Fortschritte, aber da die deutschsprachigen Kinder ebenfalls besser werden, bleibt der Abstand erhalten.

Korrigiert man die Anteile um die (analog zu oben 3.1) geschätzten unterschiedlichen Abgängen in den einzelnen Leistungsgruppen, so sind im unteren Bereich die Zuwächse absolut gesehen geringer, aber proportional gesehen teilweise höher als im oberen Bereich. Im Vergleich mit den deutschsprachigen Kindern sind die Zuwächse im unteren Bereich absolut gesehen geringer, proportional gesehen etwas höher (abgesehen von der untersten Leistungsgruppe).

Insgesamt ist die Verteilung der Kinder anderer Muttersprache um fünf bis fünfzehn Prozentränge nach unten verschoben, d. h. ein Kind anderer Muttersprache auf Prozentrang 20 bringt etwa die Leistung, die ein deutschsprachiges Kind auf Prozentrang 10 zeigt.

Es muss noch einmal betont werden, dass es sich bei dieser Berechnung von Lernzuwächsen aus Querschnittsdaten um Schätzungen handelt. Teilt man unsere Annahmen, so sieht man, dass die Entwicklung der unteren Leistungsgruppen in beiden Gruppen zu erheblicher Sorge Anlass gibt.

Im Übrigen aber gilt auch hier: Alle Kinder machen deutliche Fortschritte²⁴. Dabei ist der proportionale Lernzuwachs höher in den unteren, der absolute höher in den oberen Leistungsgruppen. Damit öffnet sich die Leistungsschere in beiden Gruppen.

4.2 Durchschnittsleistungen von Klassen in Abhängigkeit vom Anteil der Kinder mit anderer Muttersprache

Aus PISA-3 wird in der Presse berichtet, die Leseleistungen würden bei einem Anteil von Kindern mit anderer Muttersprache über 20% „sprunghaft“ schwächer, ab 40% ergäben sich aber erstaunlicherweise keine weiteren Veränderungen. Schon der Urtext (Baumert u. a. 2003, 256) ist in seinen Aussagen missverständlich. Gesprochen wird dort von einem „Anteil von Schülerinnen und Schülern mit Migrationshintergrund in Schulen“, also nicht in der einzelnen *Lerngruppe*. Dennoch wird der unzulängliche „Umgang mit Heterogenität“ im *Unterricht* beklagt.

Nach unserer Auswertung von 696 Klassen ergibt sich in LUST für die Grundschule ein etwas anderer Befund. Zwar nimmt die durchschnittlich Leseleistung der Klassen mit wachsendem Anteil an Kindern anderer Muttersprache ab. Gliedert man die Leistungen aber nach Teilgruppen auf, differenziert sich das Bild.

Es sind – bezogen auf den wachsenden Anteil von Kindern anderer Muttersprache – keine linearen Veränderungen zu erkennen. Zwei Punkte lassen sich festhalten, wobei diese

²⁴ Auch May (2000, 179-180) berichtet aus dem Hamburger PLUS-Projekt, dass Migrantenkinder vergleichbare Fortschritte machen wie die deutschsprachigen Kinder (dort in der Rechtschreibung).

Aussagen vorläufig wegen der zum Teil zu geringen Zellenbesetzung im Extrembereich auf die Schwellenwerte 0-80% bzw. 0-60% Anteil von Migrantenkinder begrenzt werden müssen:

- *Deutschsprachige* Kinder schneiden in Klassen *ohne* Kinder anderer Muttersprache *nicht* besser ab als in Klassen *mit* Kindern anderer Muttersprache.
- Für die Leseleistung von Kindern mit *anderer Muttersprache* macht es ebenfalls keinen Unterschied, wie groß ihr Anteil in der Klasse ist. Lediglich in der 3. Klassenstufe lässt sich tendenziell eine leichte lineare Verschlechterung ausmachen.

Ein größerer Anteil von Kindern mit anderer Muttersprache bedeutet danach in unserer Stichprobe weder für die Kinder mit anderer Muttersprache noch für die deutschsprachigen Kinder eine Gefährdung ihrer Leseentwicklung²⁵.

Der Eindruck einer Benachteiligung in solchen Klassen entsteht fälschlich dadurch, dass die Kinder mit anderer Muttersprache *generell* schwächere Leseleistungen erbringen. Wenn ihr Anteil in einer Klasse ansteigt, müssen also auch die Leistungen der Gruppe insgesamt sinken, selbst wenn die beiden Teilgruppen (deutsch- und fremdsprachige Kinder) gleiche Leistungen wie sonst auch erbringen.

Die für die Anteilsgruppen 60+ vorsichtig angedeuteten Negativ-Tendenzen sind zusätzlich dadurch zu relativieren, dass Klassen mit hohem Migrantenanteil mehrheitlich in sozio-ökonomisch schwächeren Stadtteilen liegen, so dass selbst in den Extremgruppen eine schlechtere Durchschnittsleistung nicht ohne Weiteres – und jedenfalls nicht allein – dem Migrantenanteil zugerechnet werden kann.

Wie stark die Kategorie „andere Muttersprache“ die sehr unterschiedlichen Voraussetzungen der gemeinten Schüler und Schülerinnen vereinfacht, wird deutlich, wenn man bedenkt, dass in 15-20% der Klassen (mit durchaus unterschiedlich hohen Anteilen von Kindern anderer Muttersprache) die deutschsprachigen Kinder sogar schlechtere Leseleistungen erbringen als ihre KlassenkameradInnen mit Migrationshintergrund.

Fazit zu 4 :

²⁵ So auch für die Schweiz Rüesch (1999, 43) und für eine bundesweite Auswertung von Erhebungen mit dem Stolperwörter-Test Metzke (2003) – anders dagegen die ersten Daten aus dem Berliner Längsschnitt von Schröder-Lenzen (2003) und die Befunde zum Übergang in die Sekundarstufe 1 bei Kristen (2000). Gegen das gängige Vorurteil spricht auch der folgende Befund aus einer Studie in 14 hessischen Grundschulklassen, auf die mich mein Siegener Kollege *Jürgen Zinnecker* aufmerksam gemacht hat: Je mehr eingewanderte Kinder in einer Schulklasse waren, um so aufmerksamer waren die Schüler im Unterricht, denn: „...befinden sich viele oder auch sehr viele Kinder aus Migrantenfamilien in einer Klasse, so scheinen alle Kinder der Klasse, also deutsche wie nicht-deutsche, besonders zur Mitarbeit und damit auch tendenziell zu besseren Lernleistungen angeregt zu werden.“ (Walter 2001, 118) Im Hamburger BLK-Modellversuch „Elementare Schriftkultur“ fielen die Leistungen erst bei einem Anteil von mehr als 80% Kindern anderer Muttersprache ab – dann aber sehr deutlich (pers. Mitteilung von Mechthild Dehn per e-mail v. 4.4.03).

Kinder, deren Eltern eine andere Muttersprache sprechen (rund 20%), schaffen zu allen Zeitpunkten weniger Sätze als die deutschsprachigen SchülerInnen. Aber auch in dieser Gruppe sind die Leistungen der Viertklässler doppelt so gut wie die der Zweitklässler.

Da Migrantenkinder deutlich häufiger nicht versetzt oder auf eine Sonderschule überwiesen werden als deutschsprachige SchülerInnen, schönt der Querschnittsvergleich allerdings auch hier das Bild. Rechnet man die Abgänge von etwa 10% zu den unteren Leistungsgruppen hinzu, so bleiben rund 20% Kinder mit nur geringen Fortschritten – gegenüber rund 10% unter den deutschsprachigen Kindern. In Klassen mit einem hohen Anteil an Migrantenkindern sind entgegen dem gängigen Vorurteil weder deren Leistungen noch die der deutschsprachigen Kinder schlechter.

Beachtung verdient schließlich der Befund, dass Migrantenkinder bei gleichem Testergebnis durchgängig schlechtere Noten erhalten als deutschsprachig aufgewachsene Kinder – außer im untersten Leistungsbereich, wo sie möglicherweise einen Mitleidsbonus erhalten.

5. Resümee

Am Ende der Grundschulzeit können 90% der Kinder unbekannte Sätze einigermaßen zügig (mindestens fünf Sätze/ Min.) und ohne größere Schwierigkeiten (weniger als 10% Fehler) erlesen und auf ihre Stimmigkeit prüfen. Allerdings streuen am Ende der Grundschulzeit die Leistungen in der Breite über mehrere Schuljahre hinweg. Diese Ergebnisse von LUST bestätigen damit die über einen anderen methodischen Zugang gewonnenen zentralen Befunde von IGLU ²⁶ .

Die schwächsten 5% können die Aufgabe entweder nur mit erheblichen Schwierigkeiten bewältigen (mehr als 15 sek pro Satz und im Durchschnitt mehr als 15% Fehler) oder gar nicht richtig lesen (ca. 0.5%). Weitere 5% der Viertklässler (Prozentrang 5-10) lesen zwar selbstständig, aber recht langsam (12-15 sek pro Satz).

In Bezug auf die große Mehrheit der Kinder, auch die mit unterdurchschnittlichen Leistungen, ist andererseits zu berücksichtigen, was wir als „Karawaneneffekt“ bezeichnen. Die unteren 10%, 20% oder 30% einer Gruppe sind *definitionsgemäß* immer schlechter als der Durchschnitt. Mit der Fixierung des Blicks auf ihren Platz in der Bezugsgruppe wird leicht übersehen, dass *alle* SchülerInnen von Jahr zu Jahr Fortschritte machen – bezogen auf ihre jeweiligen Voraussetzungen und Lernbedingungen. Auch wenn unser Befund aus dem

²⁶ s. Bos u. a. (2003) .

Querschnittsvergleich noch in einem echten Längsschnitt abzusichern ist: Pädagogisch gesehen sind die *Fortschritte* jeder *Teilgruppe* bedeutsamer als die *Abstände* innerhalb der *Gesamtgruppe*. Das gilt für leistungsschwache Schülerinnen oder Schüler generell und es gilt im Besonderen auch für Migrantenkinder.

Diese Einsicht muss auch Konsequenzen für die Leistungsbewertung haben. Erstaunlich ist, dass die unteren 10-20% trotz ständig negativer Rückmeldung überhaupt noch Fortschritte machen. Aber wie unsere Ergebnisse für die *untersten* 5% zeigen, bleiben diese immer mehr zurück, während sonst alle Teilgruppen vergleichbare Fortschritte machen.

Wie bereits gesagt: Bei einer Karawane verwundert es niemanden, wenn der, der zuletzt auf die Reise gegangen ist, auch als letzter ankommt. Bedeutsamer ist der Weg, den jeder Einzelne und die Karawane als *ganze* geschafft haben. In dieser Hinsicht sind unsere Befunde zu den Fortschritten der meisten Kinder im Lesen während der Grundschulzeit ermutigend. Allerdings: Diese Förderung muss in der Sekundarstufe fortgesetzt werden, um die Lesentwicklung derjenigen zu stabilisieren, die als letzte gestartet sind. Denn der entscheidende Faktor scheint für die meisten Kinder „mehr Zeit zum Lernen“ zu sein: Dieselben Ziele werden von fast allen Kindern erreicht – nur von den später gestarteten zu einem späteren Zeitpunkt.

„Bildungsstandards“ in Form verbindlicher Niveaustufen für alle zu demselben Termin sind vor diesem Hintergrund nicht nur sinnlos, sondern auch gefährlich. Zu befürchten ist trotz aller anderslautenden Beteuerungen, dass die Standards in Deutschland vor allem als ein weiteres Selektionsinstrument missverstanden werden. So legen z. B. die drei Kompetenzstufen der Lernstandserhebungen VERA, die sieben Bundesländer in den vierten Klassen durchführen, eine Zuordnung zu den drei Schulformen der Sekundarstufe I nahe. Ohne Berücksichtigung der individuellen Lernfortschritte werden sich die Leistungsvergleiche für das propagierte Ziel, dass alle Kinder und Jugendlichen zuverlässig besser lesen lernen sollen, eher kontraproduktiv auswirken (vgl. Brügelmann 2005, Kap. 48-50).

Literaturnachweise

Anderson, R.C., et al. (1988a): Growth in reading and how children spend their time outside school. In: Reading Research Quarterly, Vol. 23, No. 3, 285-303.

Arbeitsgruppe Leseförderung (1978): Taktiken des Lesens. In: Grundschule, 10. Jg., H. 7, 299-303.

Auernheimer, G., u. a. (Hrsg.) (2001): Interkulturalität im Arbeitsfeld Schule. Interkulturelle Studien Bd.8. Leske + Budrich: Opladen.

Backhaus, A., u.a. (2004): Leseuntersuchung mit dem Stolperwörtertest. Erste Ergebnisse aus einer Längsschnittuntersuchung im Satzlesen und Textverstehen (Vorstellung des Projekts LUST). Ms. für: Sasse, A./ Hoffmann, B. (i.V.): Bericht über die DGLS-Tagung „Erste Erfahrungen mit Schriftsprache – Konzepte für Vorschule und Schuleingangsphase“ (erscheint 2005).

Balhorn, H./ Brügelmann, H. (Hrsg.) (1995): Rätsel des Schriftspracherwerbs. Neue Sichtweisen der Forschung. "Auswahlband Theorie" der DGLS-Jahrbücher 1-5. Libelle: CH-Lengwil.

Bartnitzky, H., u. a. (1999): Zur Qualität der Leistung. 5 Thesen zur Evaluation und Rechenschaft der Grundschularbeit. Grundschulverband – Arbeitskreis Grundschule e. V.: Frankfurt. (auch in: Schmitt 1999, 164 ff.)

- Baumert, J., u. a. (Hrsg.) (2001): PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Leske + Budrich: Opladen.
- Baumert, J., u. a. (Hrsg.) (2002): PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich. Leske + Budrich: Opladen.
- Baumert, J., u. a. (2003a): PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Zusammenfassung zentraler Befunde. Max-Planck-Institut für Bildungsforschung: Berlin.
- Baumert, J., u. a. (2003b): PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Leske+Budrich: Opladen.
- BMB+F (Hrsg.) (2001): Grund- und Strukturdaten 1997/98. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie: 53170 Bonn.
- Bos, W., u. a. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Waxmann: Münster u. a.
- Brinkmann, E. (1997): Rechtschreibgeschichten – Zur Entwicklung einzelner Wörter und orthographischer Muster über die Grundschulzeit hinweg. Bericht No. 35 des Projekts OASE, FB 2 der Universität, 57078 Siegen.
- Brügelmann, H. (2000): IGLU: Glasklare Information über den aktuellen Leistungsstand – oder droht den Grundschulen wieder eine pädagogische Eiszeit? In: Grundschulverband aktuell, Nr. 72, 3-8.
- Brügelmann, H. (2003a): IGLU-2001: Zu Risiken und Nebenwirkungen... In: Grundschulverband aktuell. Nr. 82 (April 2003, 12-16).
- Brügelmann, H. (2003b): Der Karawanen-Effekt beim Rechtschreiblernen. Kinder anderer Muttersprache lernen genauso schnell, aber von einem anderen Startpunkt aus. Bericht aus der Reanalyse der freien Texte von 4.- bis 10-Klässlern in der NRW-Kids Studie 2001. Vervielf. Ms. FB 2 der Universität: Siegen (i. V.).
- Brügelmann, H. (2004s): Zur Bedeutung ausgewählter Rahmenbedingungen des Unterrichts auf die Leseleistung in der Grundschule. Vervielf. Ms. Projekt LUST/ FB2 der Universität: Siegen. → www.uni-siegen.de/~agprim/lust
- Brügelmann, H. (2005): Schule verstehen – Forschungsbefunde zu Kontroversen über Erziehung und Unterricht. Libelle: CH-Lengwil (i.D.).
- Brügelmann, H./ Backhaus A. (2003): Entwicklung der Leseleistung in verschiedenen Leistungsgruppen. Lese-Untersuchung mit dem Stolperwörter-Test (LUST-3). Antrag an die Deutsche Forschungsgemeinschaft (Entwurf: Stand 30.9.). FB 2 der Universität: Siegen.
- ddp (2003): Drei Prozent „Sitzenbleiber“. In: Süddeutsche Zeitung v. 12.3.03, Nr. 59, 50.
- Geißler, R. (2002): Die Sozialstruktur Deutschlands. Die gesellschaftliche Entwicklung vor und nach der Vereinigung, Hrsgg. von der Bundeszentrale für politische Bildung. Westdeutscher Verlag: Wiesbaden.
- Heinzel, F. / Prengel, A. (Hrsg.): Heterogenität, Integration und Differenzierung in der Primarstufe. Jahrbuch Grundschulforschung Bd. 6. Leske+Budrich: Opladen.
- Köller, O. (2002): Des Schülers Leid, des Lehrers Freud. Schulnoten sind nötig und besser als ihr Ruf. In: Schule – Wissen – Bildung. Klett ThemenDienst Nr. 16: Dezember 2002, 7-10.
- Kristen, C. (2000): Ethnic differences in educational placement: The transition from primary to secondary schooling. Arbeitspapier des Zentrums für Europäische Sozialforschung: Mannheim.
- May, P. (1995): Kinder lernen rechtschreiben: Gemeinsamkeiten und Unterschiede guter und schwacher Lerner. In: Balhorn, H. / Brügelmann, H. (1995, 220-229). [Nachdruck aus 1990]
- May, P. (2000): Lernförderlichkeit im schriftsprachlichen Unterricht. Effekte des Klassen- und Förderunterrichts in der Grundschule auf den Lernerfolg. Ergebnisse der Evaluation des Projekts „Lesen und Schreiben für alle“ (PLUS). Juni 2000. Behörde für Schule, Jugend und Berufsbildung: Hamburg.
- May, P. (2002b): Hamburger Schreibprobe 1+. Hinweise zur Durchführung und Auswertung. Verlag für Pädagogische Medien: Hamburg.

- Metze, W. (2003): Stolperwörter-Lesetest. Ergebnisse der Stichprobenerhebung (Stand 3.8.2003). Vervielf. Ms. Auch zugänglich über: www.lesetest1-4.de .
- Panagiotopoulou, A. (2002): Lernbiografien von Schulanfängern im schriftkulturellen Kontext. In: Heinzel, F. / Prengel, A. (2002, 235-241).
- Peschel, F. (2002): Offener Unterricht -- Idee - Realität - Perspektive und ein praxiserprobtes Konzept zur Diskussion. Teil I: Allgemeindidaktische Überlegungen. Schneider Verlag Hohengehren: Baltmannsweiler.
- Peschel, F. (2002): Offener Unterricht - Idee, Realität, Perspektive und ein praxiserprobtes Konzept in der Evaluation. Dissertation. FB 2 der Universität: Siegen. Schneider Verlag Hohengehren: Baltmannsweiler.
- Ratzka, N. (2003): Mathematische Fähigkeiten und Fertigkeiten am Ende der Grundschulzeit – Empirische Studien im Anschluss an TIMSS. Dissertation im FB 2 der Universität: Siegen.
- Richter, S. (1996): Unterschiede in den Schulleistungen von Jungen und Mädchen. Geschlechtsspezifische Aspekte des Schriftsprachenerwerbs und ihre Berücksichtigung im Unterricht. S. Roderer: Regensburg.
- Rolff, H.-G. (1999): Schulentwicklung in der Auseinandersetzung. In: Pädagogik, 51. Jg., H. 4, 37-40.
- Rüesch, P. (1999): Gute Schulen im multikulturellen Umfeld. Hrsgg. von der Bildungsdirektion des Kantons Zürich. Orell Füssli: Zürich.
- Schmitt, R. (Hrsg.) (1999): An der Schwelle zum dritten Jahrtausend. BundesGrundschulKongress 1999. Grundschulverband – Arbeitskreis Grundschule: Frankfurt.
- Schründer-Lenzen, A. (2003): Lesekompetenzentwicklung von Migrantenkindern. Erste Ergebnisse einer Langzeitstudie zur Schulleistungsentwicklung von Kindern mit Migrationshintergrund im Verlauf der Grundschule. Vervielf. Ms. zum Vortrag auf der Jahrestagung „Grundschulforschung“ der DGfE am 1.1.2003. Universität: Bremen.
- Walter, P. (2001): Pädagogische Kompetenz und Erfahrung in kulturell heterogenen Grundschulen: In: Auernheimer u.a. (2001).