

Hans Brügelmann

Entwicklung von Leseleistungen in Klasse 2 bis 4

Der Querschnitt LUST-1 revisited in einem Kurzlängsschnitt nach einem halben Jahr (LUST-1a)

Unsere Projekt LUST-1 („Lese-Untersuchung mit dem Stolperwörter-Test 1“) ist eine Querschnittstudie¹. Verglichen werden die Leistungen verschiedener Klassen/ SchülerInnen der Jahrgänge 2 bis 4. Auch wenn diese Klassen aus denselben Schulen stammen, lassen sich die Daten nur mit Vorbehalt als „Entwicklung“ interpretieren. Für eine solche Deutung sind Längsschnitte erforderlich, in denen *dieselben Kinder* erneut getestet werden. Diese Bedingung erfüllen erst die geplanten Nachfolgestudien LUST-2 und LUST-3.

Allerdings liegen auch für eine Teilstichprobe LUST-1a inzwischen Längsschnittdaten vor, weil einige LehrerInnen den Stolperwörter-Lesetest im Sommer, also ein halbes Jahr nach unserer Ersterhebung² erneut eingesetzt haben.

Wie Tab. 1 zeigt, entspricht unser Teillängsschnitt im Jahrgang 2 der Gesamtstichprobe, die wir in LUST-1 (vgl. die Werte in Klammern) ausgewertet haben; im dritten Jahrgang liegen seine Leistungen geringfügig unter, im vierten Jahrgang deutlich über dem Durchschnitt.

Tab. 1 ³ : Vergleich des Längsschnitts LUST-1a mit der Grundstichprobe LUST-1				
Jahrgang	N =	richtige Sätze/ Minute		
		Mittelwert (aM)	Streuung (SD)	Zuwachs ⁴
2	202 (5.849)	4.1 (4.2)	2.1 (2.2)	+ 0.9
3	169 (6.047)	5.7 (6.0)	2.3 (2.3)	+ 0.8
4	69 (5.708)	9.1 (8.1)	2.5 (2.7)	+ 1.1

¹ Vgl. Brügelmann, H. (2003k): Lese-Untersuchung mit dem Stolperwörter-Test. Abschlussbericht des Projekts LUST-1 → www.uni-siegen.de/~agprim/lust

² Im Durchschnitt beträgt der Abstand 6-7 Monate, mit den Extremen 5.0 – 9.0 . In den zweiten und dritten Klassen liegt das arithmetische Mittel bei 7 Monaten, in den vierten Klassen bei 6 Monaten; s. im einzelnen Tab. 0a im statistischen Anhang.

³ In dieser und den folgenden Tabellen sind die Werte um der besseren Lesbarkeit willen gerundet; vgl. für die genauen Daten die Grundtabellen im statistischen Anhang.

⁴ Der Zuwachs *rmin_dif* erfasst die Differenz der Rohwerte („richtige Sätze pro Minute“) zwischen dem zweiten und dem ersten Termin.

Im Durchschnitt steigen die Leistungen in der zweiten Klasse um 0.9 richtige Sätze/ Minute, in der dritten Klasse um 0.8 und in der vierten Klasse um 1.1 richtige Sätze/ Minute.

In unserem Querschnitt LUST-1 lagen die Jahrgangsdifferenzen zwischen 1.8 und 2.1 richtigen Sätzen pro Minute. Rechnet man die Halbjahreswerte des Teillängsschnitts LUST-1a auf ein ganzes Jahr um, so kommt man auf 1.6 – 2.2 Sätze pro Minute als Differenz zwischen den Jahrgängen. Damit werden die Vergleiche aus dem Querschnitt nachträglich als recht genaue Annäherung bestätigt.

Der Leistungszuwachs der schwächeren SchülerInnen ist generell größer als der der leistungsstarken SchülerInnen⁵. Diese Daten sprechen für einen Aufholeffekt über die Grundschulzeit hinweg und gegen eine Öffnung der Leistungsschere⁶. Tab. 2 belegt dies durch einen prozentualen Bezug der Zuwächse auf die Ausgangslage.

Tab. 2: Prozentuale Zuwächse bezogen auf das jeweilige Ausgangsniveau		
richtige Sätze/ min. (1. Termin)	N =	durchschnittlicher Zuwachs ⁷
12+	9	- 6 %
6-12	163	7 %
4-6	111	21 %
3-4	70	35 %
2-3	48	65 %
0-2	36	117 %
Gesamt	437	30 %

An den extremen veranschaulicht: In der Gruppe, die ihre Leistung gegenüber dem ersten Termin um mindestens ein Zehntel steigert, finden sich in den obersten Gruppen nur 11-41%, in den untersten dagegen über 90%.

Nun ist zu bedenken, dass eine geringe *prozentuale* Verbesserung der Teilgruppen mit höherem Ausgangsniveau in *absoluten* Zahlen durchaus einen vergleichbaren Zuwachs bedeuten *kann*, so dass der prozentuale Vergleich möglicherweise nicht fair ist. Wie Tab. 3 zeigt, verändert sich das Bild aber nur wenig, wenn man absolute Veränderungen vergleicht: Jeweils rund ein Drittel der verschiedenen Leistungsgruppen findet sich im mittleren Bereich nur geringfügiger Veränderungen (höchstens +/- ein Satz Zu- oder Abnahme der Leistung vom ersten zum zweiten Termin). In der Teilgruppe, die ihr Ergebnis um mehr als einen Satz pro Minute steigert, finden sich aus den beiden leistungsstärksten Gruppen aber nur 22-39%, in den beiden leistungsschwächsten sind es dagegen 59-69%.

⁵ Vgl. die Korrelationen von .34** bis .39** in Grundtabelle 1b im Anhang.

⁶ Vgl. ausführlicher zu konkurrierenden Befunden und Erklärungsmodellen: 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 → www.uni-siegen.de/~agprim/lust

⁷ RMIN-DIF by NIVEAU

Tab. 3: Veränderung vom 1. Termin zum 2. Termin in <i>absoluten</i> Zuwächsen				
richtige Sätze/ min. 1. Termin	N =	minus > 1 Sätze	minus 1 bis plus 1 Satz	plus > 1 Sätze
12+	9	44 %	33 %	22 %
6-12	163	22 %	39 %	39 %
4-6	111	5 %	45 %	50 %
3-4	70	3 %	39 %	59 %
2-3	48	-	31 %	69 %
0-2	36	-	36 %	64 %
Gesamt	437			

Man könnte nun evtl. einwenden, dass die oberen Leistungsgruppen bereits ein so hohes Niveau erreicht haben, dass nur noch geringfügige Verbesserungen möglich sind (Deckeneffekt). Dagegen sprechen unsere Befunde bei Erwachsenen, die noch einmal einen deutliche Zuwachs gegenüber den leistungsstarken GrundschülerInnen aufweisen⁸. Gegen einen Deckeneffekt spricht auch, dass der Leistungszuwachs schon unterhalb der Spitzengruppe absinkt und diese Abnahme der Zuwachsrate über das ganze Spektrum der Verteilung zu beobachten ist (vgl. Grundtabelle 3b im Anhang): im Durchschnitt verändert sie sich graduell von +1.2/1.5 Wörtern pro Minute auf - 0.8 Wörter pro Minute.

Dennoch bleibt zu prüfen, ob der Leistungszuwachs als echte Verbesserung der Leseleistung oder als bloßer Test-Retest-Effekt zu deuten ist, ob er also nicht schon die größere Vertrautheit mit dem Aufgabenformat bei dem zweiten Termin verursacht worden sein kann. Dieser Vertrautheitseffekt könnte in den verschiedenen Leistungsgruppen unterschiedlich stark ausgeprägt sein und damit die Entwicklungsdifferenzen erklären. In diesem Fall würde die stärkere Verbesserung in den unteren Gruppen gar keinen echten Leistungszuwachs abbilden.

Zur Aufklärung dieser Frage können unsere Befunde aus der Kurzzeitwiederholung in einer anderen Stichprobe beitragen.

Tab. 4: Kurzzeitvergleich des Leistungszuwachses binnen einer Woche (N = 158)	
Leistungsgruppe 1. Termin	durchschnittlicher Zuwachs binnen einer Woche
12+	0.3
6-12	1.9
4-6	1.3
3-4	1.6
2-3	.9
0-2	1.0
Durchschnitt	1.5

⁸ Brügelmann, H. (2004k): Leseleistungen von HandwerkerInnen im Stolperwörter-Lesetest. Erste Befunde und ihre Deutung. Vervielf. Ms. Projekt LUST/ FB 2 der Universität: Siegen; Brügelmann, H. (2004l): Leseleistungen von LehrerInnen und Lehramtsstudierenden im Stolperwörter-Lesetest. Erste Befunde und ihre Deutung. Vervielf. Ms. Projekt LUST/ FB 2 der Universität: Siegen. → www.uni-siegen.de/~agprim/lust

In dieser Zusatzstudie haben 158 SchülerInnen aus 3. und 4. Klassen den Stolperwörter-Lesetest binnen einer Woche wiederholt. Schlüsselt man die Veränderungen wiederum nach Leistungsgruppen auf, so ergibt sich folgendes Bild (vgl. Tab. 4 im Gegensatz zu den Werten in der Grundtabelle Tab. 3b):

Tab. 4 zeigt Zweierlei: Zum einen gibt es in der Tat einen erheblichen Test-Retest-Effekt, denn die Leistungen nehmen nach einer Woche erheblich zu – und zwar im Durchschnitt mehr als nach einem halben Jahr. Ob dieser Effekt auf die Vertrautheit mit dem Aufgabenformat oder mit der Erinnerung an den Inhalt der einzelnen Sätze zu tun hat, muss offen bleiben. Für unsere Frage interessant ist aber, dass der Vertrautheitseffekt über die Leistungsgruppen hinweg anders streut als der Leistungszuwachs nach einem halben Jahr. Dessen unterschiedliche Ausprägung ist insofern nicht durch den Vertrautheitseffekt zu erklären, sondern muss als echte Differenz in der Lernentwicklung interpretiert werden.

Analog kann für die Gruppe der MigrantInnen festgehalten werden, dass ihr Zuwachs dem der deutschsprachigen Kinder entspricht: plus 1.3 vs. 1.4 richtige Sätze pro Minute. Auch eine Aufgliederung nach den Leistungsgruppen der Ausgangslage führt zu keinem anderen Ergebnis (vgl. Grundtabelle 5a): Zweimal ist der Fortschritt der deutschsprachigen Kinder größer, zweimal der Lernzuwachs der Migrantenkinder, und in der untersten Gruppe ist der Fortschritt gleich groß:

Leistungsgruppe 1. Termin richtige Sätze/ Minute	Zuwachs (RMIN_DIF) von Termin-1 zu Termin-2	
	Deutschsprachig (N =)	Migrationshintergrund (N =)
12+	1.4 (2)	-
6-12	1.7 (53)	1.0 (7)
4-6	1.1 (49)	1.4 (22)
3-4	1.4 (24)	1.0 (25)
2-3	1.6 (16)	1.8 (17)
0-2	1.1 (9)	1.1 (16)

Fazit: Unsere Interpretation der Jahrgangsvergleiche aus dem Querschnitt LUST-1 als „Karawanen-Effekt“ wird insofern bestätigt, als auch der Teillängsschnitt zeigt, dass die leistungsschwachen SchülerInnen und die Migrantenkinder erfolgreich lernen – von ihrem jeweiligen Ausgangsniveau aus. Gegenüber unserer vorsichtigen Interpretation, dass die schwachen SchülerInnen nicht zurückfallen, zeigt der Längsschnitt LUST-1 sogar, dass sie in der gleichen Zeit größere Fortschritte machen, dass sich die Leistungsschere also etwas schließt. Der Abstand der Migrantenkinder gegenüber den deutschsprachigen Kindern nimmt zwar nicht ab, aber er wird wegen der vergleichbaren Leistungszuwächse auch nicht größer: sie entwickeln sich normal, erreichen dasselbe Niveau nur von ihrem niedrigerem Ausgangspunkt aus jeweils etwas später.

Diese für die Unterrichtspraxis – insbesondere für die Leistungsbewertung – sehr bedeutsamen Befunde werden in LUST-2 zur Zeit an einer größeren Stichprobe überprüft.