

Lernstandsdiagnose im Rechtschreiben – mit und ohne Test ¹

Die Frage, ob Leistungsdiagnosen mit oder ohne Tests durchgeführt werden sollten, ist keine Entweder-oder-Frage. Lehrereinschätzung, informelle Beobachtungsverfahren und standardisierte Tests können sinnvoll miteinander integriert eingesetzt werden – wenn sie bestimmte Kriterien erfüllen.

1 Für und wider Tests

Ende der achtziger Jahre, als wir mit der Entwicklung der Konzeption der Hamburger Schreibprobe begonnen, stand die Mehrzahl der Lehrkräfte der Anwendung von Tests noch überwiegend skeptisch gegenüber. Tests wurden von vielen als einseitig ausgerichtete und fremdbestimmte Instrumente abgelehnt, die nach Ansicht vieler Lehrer einem ganzheitlich ausgerichteten Unterricht nicht entsprachen. Der frühere Herausgeber der Beltz-Test-Reihe vermutete damals bei den Lehrkräften generelle ideologische Gründe für ihre Ablehnung (siehe Ingenkamp 1989). Im Zuge der „empirischen Wende“ hat sich die Einstellung großer Teil der Lehrerschaft verändert. Dies liegt zum einen im wachsenden Einfluss eines pädagogischen Pragmatismus, der auch die empirische Überprüfung der Lernergebnisse einbezieht.²

Zum zweiten wurde die veränderte Einstellung zum Einsatz von Tests auch durch einen Paradigmenwechsel in der Testdiagnostik ermöglicht: Während die früheren Rechtschreibtests erfassten, ob die Kinder bestimmte Wörter, deren Auswahl durchaus fragwürdig erschien, richtig schreiben können, zielt die moderne Testkonzeption auf die Frage, welche Strategien und Zugriffsweisen die Kinder bereits realisieren, auch wenn die Schreibungen der einzelnen Wörter noch fehlerhaft sind. Damit bezieht sich die neue Generation von Rechtschreibtests nicht nur auf dieselben Grundannahmen eines entwicklungsorientierten Rechtschreibunterrichts, sondern die verbreitete Anwendung des strategiebezogenen Diagnosekonzepts beeinflusst ihrerseits auch die didaktische Diskussion. Denn mit der Feststellung eines Strategieprofils stellt sich die Frage, wie der Unterricht die Ausbildung relevanter Rechtschreibstrategien fördern kann.

Heute besteht für die meisten Lehrkräfte nicht die Alternative, ob sie grundsätzlich Tests für die Lernstandsdiagnose einsetzen sollen oder nicht, sondern sie entscheiden in Abhängigkeit von der Fragestellung, ob die Alltagsbeobachtungen und Klassenarbeiten durch Testdiagnosen

¹ Veröffentlicht in: Grundschule, 34. Jg., Heft 5/2002, S. 44 - 46.

² Die Bereitschaft zur Beteiligung an nationalen und internationalen Vergleichsstudien (TIMMS und PISA) wurde insbesondere durch entsprechende Vorläuferuntersuchungen im Bereich der Rechtschreibleistungen Anfang der neunziger Jahre (vgl. Brügelmann u.a. 1991; May 1995a und May 1994c) vorbereitet und durch offene Ergebnisdiskussion erhöht.

ergänzt werden sollen. Welche Gesichtspunkte bei der Wahl der Diagnosemethode beachtet werden sollten, dazu werden im folgenden thesenhaft einige Argumente vorgestellt (siehe auch den Kasten).

2 Maßstäbe für die Lernstandsdiagnose

Lernstandsdiagnosen dienen dem Ziel, den Unterricht bzw. die Förderung gezielt zu planen, den Lernerfolg einzuschätzen und Gesichtspunkte für die Entscheidungen über die weitere Förderung zu liefern. Lernentwicklungskontrollen werden darüber hinaus vorgenommen, um die Wirkung und den Erfolg von Unterrichts- und Fördermaßnahmen abzuschätzen.

Als Maßstäbe für die Beurteilung des Lernerfolgs kommen drei verschiedene Bezugsgrößen in Frage:

- Der *individuelle Vergleich* zielt auf die konkrete Lernentwicklung einzelner Lerner bzw. Lerngruppen: Welche Fortschritte hat der Einzelne oder die Lerngruppe bei der Aneignung des Lernstoffes erzielt? Ein deutlicher Lernzuwachs wäre demnach auch dann ein Erfolg, wenn andere Lerner vergleichsweise höhere Leistungen erzielen oder ein bestimmtes Lernziel noch nicht erreicht wurde.
- Der *sachliche Vergleich* zielt auf das fragliche Erreichen definierter Lernziele: Haben die Lerner ein bestimmtes Lernziel, das im Lehrplan vorgesehen ist oder das als Mindeststandard gilt, erreicht oder sich diesem angenähert? Die Bewertung erfolgt nur bezüglich dieser sachlichen Norm, unabhängig davon, wieviel Vorwissen die Lerner bereits hatten oder ob andere Lerner mehr oder weniger Lernzuwachs aufweisen.
- Der *soziale Vergleich* zielt auf den Vergleich verschiedener Lerner oder Lerngruppen: Ist der Lernzuwachs eines Lernenden oder der ganzen Klasse im Vergleich zu anderen Lernern höher oder niedriger? Bei dieser Bewertung wird die individuelle Lernentwicklung einer sozialen Bezugsgruppe (z.B. Klasse, Schulform, nationale Altersgruppe) zugeordnet.

Diese Maßstäbe können nicht nur für die Feststellung des individuellen Lernerfolgs und des erreichten Lernstandes in der Klasse bzw. Lerngruppe herangezogen werden, sondern sie dienen auch zur Beurteilung der Wirksamkeit des Unterrichts und der Fördermaßnahmen.

3 Lernentwicklungskontrolle

Um den Erfolg des Unterrichts zu kontrollieren, wird in regelmäßigen Abständen der Lernstand der Kinder bestimmt. Die Lernentwicklungskontrolle kann sich auf die gesamte Klasse, eine definierte Lerngruppe (z.B. Fördergruppe) oder auf einzelne Kinder beziehen. Aus der

Erhebung des erreichten Könnens in der Klasse bzw. Lerngruppe ergeben sich Hinweise für die weitere Planung und Gestaltung des Unterrichts, aus der Feststellung der individuellen Lernstände können Überlegungen zur gezielten Förderung einzelner Kinder abgeleitet werden. Lernentwicklungskontrollen dienen darüber hinaus dazu, die Wirkung und den Erfolg von Unterrichts- und Fördermaßnahmen abzuschätzen.

Bei der Kontrolle der *Lernentwicklung* müssen mindestens zwei (besser: mehrere) Zeitpunkte, an denen die Lernstände erfasst wurden (z.B. vor und nach einem Unterrichtsprojekt oder einer Fördermaßnahme), in Beziehung gesetzt werden. Legte man nämlich nur einen Zeitpunkt zugrunde, so würde nicht die tatsächliche Lernentwicklung, also das Ausmaß des Lernerfolgs, bewertet, sondern ausschließlich das erreichte Leistungsvermögen. Dieses wird jedoch wesentlich durch Faktoren (z.B. Intelligenz, bereits vorhandenes Vorwissen) mitbestimmt, die weitgehend unabhängig vom Lernverhalten des Kindes und vom Unterricht innerhalb des fraglichen Zeitraumes sind.

4 Lehrkräfte als Diagnostiker

Die Durchführung der Lernstandsdiagnose und Lernentwicklungskontrolle geschieht im Alltag meist durch die Lehrkräfte selbst. Sie sind Experten für die Einschätzung individueller Lernstände und müssen darüber regelmäßig schriftlich berichten (z.B. in Zeugnissen und Entwicklungsberichten). Lehrkräfte sind demnach professionelle Experten für die Einschätzung individueller Lernstände, um ihre Schüler im Unterricht zu fördern. Im Unterschied zur Anwendung standardisierter Diagnosemethoden haben Lehrkräfte dabei stets die gesamte Persönlichkeit des Lernenden im Blick und beurteilen die Ergebnisse auf dem Hintergrund länger andauernder Lernprozesse und unter Einbeziehung der institutionellen und personellen Rahmenbedingungen. Daher ist ihre Beurteilung der Lernfortschritte und des Unterrichtserfolgs unverzichtbar und durch kein anderes Diagnoseinstrument ersetzbar.

Allerdings ergeben sich bei der Beurteilung durch Lehrkräfte auch Einschränkungen, die es zu beachten gilt:

- Empirische Vergleiche zwischen verschiedenen Lehrkräften ergeben regelmäßig große individuelle Unterschiede hinsichtlich der Qualität ihrer Diagnosen.
- Einschätzungen von Lehrkräften unterliegen systematischen Wahrnehmungstereotypen (z.B. Halo- bzw. Ausstrahlungseffekt, klasseninterner Beurteilungsmaßstab, Pygmalion-Effekt), die ihre Vergleichbarkeit einschränken können.
- Darüber hinaus wirken soziale und emotionale Einflüsse auf die subjektiven Wahrnehmungen und Einschätzungen der Lehrkräfte ein. Insbesondere bei der Beurteilung von Lernprozessen, für die die Lehrkräfte selbst Verantwortung übernehmen – z.B. bei der

Einschätzung von Lernfortschritten geförderter Kinder oder bei klassenübergreifenden Leistungsvergleichen – ist auch das Selbstkonzept der Lehrkräfte bzw. die Selbsteinschätzung ihrer eigenen Wirksamkeit betroffen, und diese Tatsache begünstigt – je nach Persönlichkeit – eine Tendenz zur selbstwertdienlichen Überschätzung bzw. „Euphemisierung“ der Lernergebnisse.

Die Verbesserung der diagnostischen Qualität der Lehrereinschätzungen ist entsprechend eine wichtige Aufgabe für die Lehrerfortbildung, im Rahmen derer Lehrkräfte auch in der differenzierten und lernentwicklungsrelevanten Beurteilung von Schülerleistungen im förderdiagnostischen Zusammenhang trainiert werden sollten.

5 Vorteile von Schultests

Bei der klasseninternen Lernentwicklungskontrolle werden meist informelle Prüfaufgaben (z.B. spezielle Arbeitsblätter aus dem Unterricht) verwendet. Darüber hinaus werden mittlerweile in vielen Schulämtern auch klassen- und schulübergreifende Vergleichsarbeiten eingesetzt. Um die vorhandenen Förderressourcen angemessen steuern und den Erfolg von Bildungsmaßnahmen überprüfen zu können, reichen die Einschätzungen bzw. Selbsteinschätzungen der Lehrkräfte wegen der erwähnten Einschränkungen nicht immer aus. Daher ist es sinnvoll, von Zeit zu Zeit eine ergänzende Erfassung der schriftsprachlichen Leistungen in den Schulen durch objektive Erhebungsinstrumente (z.B. Lese- und Rechtschreibtests) vorzunehmen, mit denen die Lernstände objektiv und zuverlässig erfasst werden können..

Neben der Objektivität und Zuverlässigkeit ihrer Ergebnisermittlung liegt der Vorteil von Tests, die auf empirischer Grundlage konstruiert wurden, vor allem in der Standardisierung der Ergebnisse, wodurch die Möglichkeit zum Vergleich mit definierten Bezugsgruppen auch über die Lerngruppe hinaus eröffnet wird. Für die Auswahl der geeigneten Erfassungsinstrumente im Rahmen der Evaluation von Bildungsmaßnahmen wächst das Gewicht der Kriterien Objektivität und Zuverlässigkeit in dem Maße, wie der Umfang und die bildungspolitische Bedeutung der Evaluationsergebnisse steigen. Um die Notwendigkeit zusätzlicher Fördermaßnahmen zu begründen oder den Erfolg gezielter Bildungsmaßnahmen zu belegen, reicht die subjektive Beurteilung der Betroffenen im Allgemeinen nicht aus, um die Öffentlichkeit oder die Geldgeber zu überzeugen. Daher ist die Weiterentwicklung geeigneter objektiver Verfahren für die Leistungsdiagnose auch bildungspolitisch von hoher Bedeutung.

6 Anforderungen an Schultests

An die Qualität der Instrumente zur Erfassung schriftsprachlicher Leistungen sind dabei hohe Anforderungen zu stellen. Um einen Beitrag zur Verbesserung des schriftsprachlichen Unter-

richts leisten zu können, müssen Tests nicht nur objektiv und zuverlässig sein, sondern darüber hinaus sollten sie den Lehrkräften Hilfen für den Unterricht und für gezielte Fördermaßnahmen an die Hand geben. Dazu sind v.a. folgende Gesichtspunkte zu berücksichtigen.

- Leistungstests dürfen sich nicht nur auf die Ermittlung globaler Ergebnisse beschränken, die allenfalls ein allgemeines Leistungsranking erlauben, sondern müssen darüber hinaus gesicherte Hinweise auf den Lernstand geben. Dies gilt insbesondere für die differenzierte Beschreibung des erreichten Könnens der Lernenden aus dem schwachen Leistungsspektrum.
- Die Testergebnisse müssen vor allem gültig für den Leistungsbereich sein. Dazu gehört nicht nur, dass sie im Gesamtergebnis hoch mit anderen spezifischen Leistungsmessungen (z.B. Lehrereinschätzungen) korrelieren, sondern dazu gehört auch der Nachweis ihrer inhaltlichen Validität (Repräsentanz des zu erfassenden Leistungsmerkmals) sowie der theoretisch und empirisch überprüfbaren Konstruktvalidität (überprüfbare Annahmen über die Entwicklung der Fähigkeiten und Möglichkeiten ihrer Beeinflussung). Das bedeutet, dass die den Testverfahren zugrunde liegenden Annahmen und Konstruktionsmerkmale den neueren Ergebnissen der linguistischen, entwicklungspsychologischen und pädagogischen Forschung entsprechen müssen.
- Darüber hinaus sollen sie unbedingt alltagstauglich sein, das heißt, sie sollen den praktischen Erfordernissen des Unterrichtsalltags entsprechen und in das didaktische Konzept der Lehrkräfte integrierbar sein. Nur unter dieser Voraussetzung dienen die Leistungstests nicht nur einer übergeordneten Leistungskontrolle, sondern können direkt den laufenden Unterricht unterstützen.

7 Erfassung grundlegender Rechtschreibstrategien

Wie die verschiedenen Ansprüche an Schulleistungstests integriert und dennoch anwendungsökonomisch realisiert werden können, lässt sich am Beispiel der Hamburger Schreibprobe (HSP 1-9) zeigen. Mit dieser Testserie ist es möglich, die Leistungen im Rechtschreiben mit alltagsnahen und motivierenden Aufgaben objektiv und zuverlässig zu erfassen und auf der Basis der differenzierten Einschätzung des Lernstands Fördermaßnahmen zu planen und den Unterrichtserfolg zu überprüfen. Die verschiedenen Jahrgangsversionen folgen einem einheitlichen Konzept, so dass Lernentwicklungen auch über Jahre hinweg erfasst werden können. Mit der HSP werden die Schreibungen ausgewählter Einzelwörter und Sätze, die das Regelsystem der deutschen Orthographie repräsentieren, nicht nur nach der Anzahl richtig geschriebener Wörter sowie der sog. Graphemtreffer ausgewertet, sondern darüber hinaus nach den grundlegenden Rechtschreibstrategien (alphabetische, orthographische, morphematische sowie wortübergreifende Zugriffsweisen). Darüber hinaus erfasst die HSP

überflüssige orthographische Elemente, die auf noch bestehende Unsicherheiten beim Anwenden der orthographischen Strategie hindeuten, sowie Oberzeichenfehler, die auf den Grad der Sorgfalt und auf die Kontrolle beim Schreiben hinweisen. Auf diese Weise bietet der Rechtschreibtest den Lehrkräften nicht nur eine empirisch überprüfte Vergleichsmöglichkeit für die Schülerleistungen, sondern darüber hinaus ein theoretisch ausgewiesenes Konzept, das auch in der Alltagsdiagnose angewendet werden kann.

8 Anwendung der HSP-Kategorien zur Analyse freier Texte

Auch wenn es nicht sinnvoll oder möglich ist, einen standardisierten Rechtschreibtest einzusetzen (z.B. außerhalb der festgelegten Zeiträume, für die Normen vorliegen), kann das diagnostische Konzept der HSP für die qualifizierte Feststellung des Lernstand der Kinder genutzt werden. So ist es durchaus möglich, die Graphemtreffermethode und die strategiebezogenen Auswertungskategorien der Hamburger Schreibprobe auf andere Schreibwortschätze, wie sie in Übungsdiktaten, freien Texten usw. vorkommen, anzuwenden. Dabei ist jedoch zu beachten, dass die Aussagefähigkeit solcher informeller Diagnosen von der orthographischen Struktur des zugrunde liegenden Wortmaterials abhängt. Daher sollte sich die informelle Analyse der Rechtschreibung mit Hilfe der strategiebezogenen Auswertung stets auf eine Auswahl von Wörtern beziehen, die hinreichend repräsentativ für die deutsche Orthographie ist. Um die Schreibungen verschiedener Kinder vergleichen zu können, sollten neben freien Texten auch vorgegebene Wörter und Sätze ausgewertet werden, die von allen Kindern zu schreiben sind.

In gewissen Abständen (mindestens einmal pro Schuljahr) sollten die informellen Rechtschreibanalysen durch einen formellen Rechtschreibtest ergänzt werden, der die Schreibleistung mit einer repräsentativen Wortauswahl und mit Bezug auf eine altersentsprechende Vergleichsgruppe bestimmt, damit Auffälligkeiten in der Lernentwicklung einzelner Kinder oder der ganzen Klasse hinreichend sicher erfasst werden können.

Vorzüge von Expertenbeurteilung und Testdiagnose

Merkmale der Expertenbeurteilung

- spontan-flexible Planung und Durchführung der Diagnose
- jederzeit schnell und ohne Zusatzkosten durchführbar
- optimale Einpassung in den laufenden Unterricht und den Lehrgang, da verwendetes Material für die Diagnose dem Lehrgang bzw. Klassenunterricht entstammt
- ganzheitliche Betrachtung der Gesamtpersönlichkeit des Kindes
- Bezug auf Leistungsniveau in der Klasse bzw. Schule
- Diagnosegüte wird (in der Regel) nicht empirisch überprüft
- Grundlage der Interpretation der Diagnose ist das know-how der einzelnen Lehrkraft
- geringe Vergleichbarkeit der Ergebnisse mit anderen Klassen, anderen Leistungsbereichen usw.
- Beurteilungsgesichtspunkte häufig subjektiv, Beurteilungsfehler
- Diagnoseergebnisse können häufig nicht einfach kommuniziert werden, da verallgemeinerbare Kriterien fehlen

Merkmale der Testdiagnose

- standardisierte Durchführung zur Sicherung der Objektivität
- Durchführung muss vorausgeplant werden (Normzeitpunkte), Tests müssen bestellt und gesondert bezahlt werden
- Material für die Diagnose wird aufgrund systematischer Analyse und empirischer Überprüfung der Relevanz ausgewählt
- ausschließlich analytisch-sachliche Betrachtung unter Ausblendung der individuellen Lernerpersönlichkeit
- Bezugsrahmen ist die allgemeine Lernentwicklung einer repräsentativen Auswahl
- Diagnosegüte wird empirisch überprüft
- Grundlage der Interpretation der Diagnose ist das know-how der Testentwickler und ihrer wissenschaftlichen Überprüfungen
- hohe Vergleichbarkeit der Ergebnisse durch einheitliche Begrifflichkeit und standardisierte Normen
- Beurteilungsgesichtspunkte theoretisch analysiert und empirisch überprüft
- Testergebnisse können relativ leicht vermittelt und ausgetauscht werden

Literatur:

Brügelmann, H., Lange, I. & Spitta, G. (1991): Leistungs-Patt in der Rechtschreibung? Ein Vergleich in ausgewählten Regionen in Deutschland Ost und West. In: Grundschulzeitschrift, 5, Heft 43/1991, S. 32-34.

Ingenkamp, K.(1989): Diagnostik in der Schule. Beiträge zu Schlüsselfragen der Schülerbeurteilung. Weinheim, Basel: Beltz.

May, P. (1994c): Rechtschreibfähigkeit und Unterricht. Rechtschreibleistungen Hamburger Schüler im vierten Schuljahr im Zusammenhang mit Merkmalen schriftsprachlichen Unterrichts. Ergebnisse der Voruntersuchung zum Projekt Lesen und Schreiben für alle

(PLUS). Bericht der Wissenschaftlichen Begleitung Nr. 94/01. Hamburg: Behörde für Schule, Jugend und Berufsbildung.

May, P. (1995a): Rechtschreiblernen in West und Ost: Warum konnten DDR-Schüler besser schreiben? In: Niemeyer, W. (Hg.): Kommunikation und Lese- Rechtschreibschwäche. Bochum: Winkler, S. 169 – 178.

May, P.(1998a): Diagnose orthographischer Kompetenz. Zur Erfassung der grundlegenden Rechtschreibstrategien mit der Hamburger Schreib-Probe. Unter Mitarbeit von Ulrich Vieluf und Volkmar Malitzky. Hamburg: Verlag für pädagogische Medien. 4. Auflage.

May, P.(2002): PLUS – Wege zum schriftsprachlichen Lernerfolg. In: Zeitschrift Grundschule 4/2002.