

Peter May

Nach welchen Maßstäben lassen sich Schulleistungen vergleichen?¹

Das Thema „Leistungsmaßstäbe“ ist in den letzten Jahren zunehmend in den Blickpunkt der pädagogischen Öffentlichkeit gerückt, nachdem es über lange Zeit so schien, als sei der Leistungsbegriff aus der Diskussion weit gehend verbannt worden. Nicht zuletzt die nationalen und internationalen (z.B. PISA, TIMSS) sowie die regionalen (z.B. PLUS, LAU) Vergleichsstudien haben den Fingerzeig auf die lange verdrängte Frage gelegt, wie bestimmt werden kann und wer bestimmen soll, welche Leistungen in der Schule erbracht werden müssen bzw. wie die erbrachten Leistungen bewertet werden sollen.

Der diesem Thema gewidmete Schwerpunkt umfasst v.a. Beispiele aus der Schulpraxis, die vielfältige Facetten von Schulleistungen sichtbar machen, verschiedenartige Verfahren zur Erfassung von Leistungen darstellen und Belege dafür liefern, dass für die Beurteilung ganz unterschiedliche Maßstäbe herangezogen werden können und müssen. Es werden Leistungen beschrieben, die als Ausdruck besonderer Begabungen zu verstehen sind; daneben gibt es Leistungen, die sich aufgrund besonderer Anstrengungen der Lernenden ergeben, und es gibt Leistungen, die aufgrund gezielter Förderung in der Schule oder im Elternhaus erbracht werden können.

Was ist Schulleistung?

In der pädagogischen Praxis bleibt häufig unklar, was mit der Aussage eigentlich gemeint ist, die Leistung eines Schülers sei hoch oder niedrig, angemessen oder mangelhaft? Ist eine individuelle Leistung als gut zu bewerten, obwohl sie geradezu leicht und spielerisch erzielt wird, während andere Schüler sich enorm anstrengen müssen oder massive Hilfe benötigen? Und ist eine Leistung dennoch als gering anzusehen, wenn sich ein Schüler mit Lernschwierigkeiten um ein bestmögliches Ergebnis bemüht?

In der pädagogisch-psychologischen Lernforschung ist es üblich, die Schulleistung innerhalb eines Wirkungsgefüges zu betrachten, das neben der Persönlichkeit des einzelnen Schülers die Bedingungsfaktoren auf Seiten der Familie (z.B. Bildungsniveau, sprachlicher Hintergrund, elterliche Unterstützung) und auf Seiten der Schule (z.B. Unterrichtsqualität, individuelle Förderung) einbezieht. Und diese Faktoren müssen nicht nur in ihrer aktuellen Ausprägung, sondern auch in ihren längerfristigen Bedingungen betrachtet werden, da sich die aktuelle Leistungsfähigkeit im Laufe einer längeren Lerngeschichte entfaltet hat (siehe Abbildung 1).

Wenn wir also die aktuellen Leistungsergebnisse betrachten, müssen wir auch deren Entstehungsgeschichte und damit auch unsere eigenen Anteile (als unterstützende Eltern oder Lehrkräfte) einbeziehen. Die gezeigte Leistung ist demnach als das Ergebnis eines (meist längerfristigen) Lernprozesses zu verstehen, der von der Persönlichkeit des Lernenden und den Bedingungen in Elternhaus und Schule geprägt ist.

¹ veröffentlicht in: HAMBURG MACHT SCHULE, Heft 1/2003, S. 4-7

Die Beurteilung der Leistung bezieht sich indirekt auch auf die den Lernprozess stützenden oder behindernden Faktoren. Und dies macht es insbesondere für Lehrkräfte schwierig, die Leistungen der Schüler objektiv und fair zu bewerten, da sie ja mittelbar auch die Wirkungen des eigenen Unterrichts zu bewerten haben (siehe dazu ausführlicher: Weinert 2001).

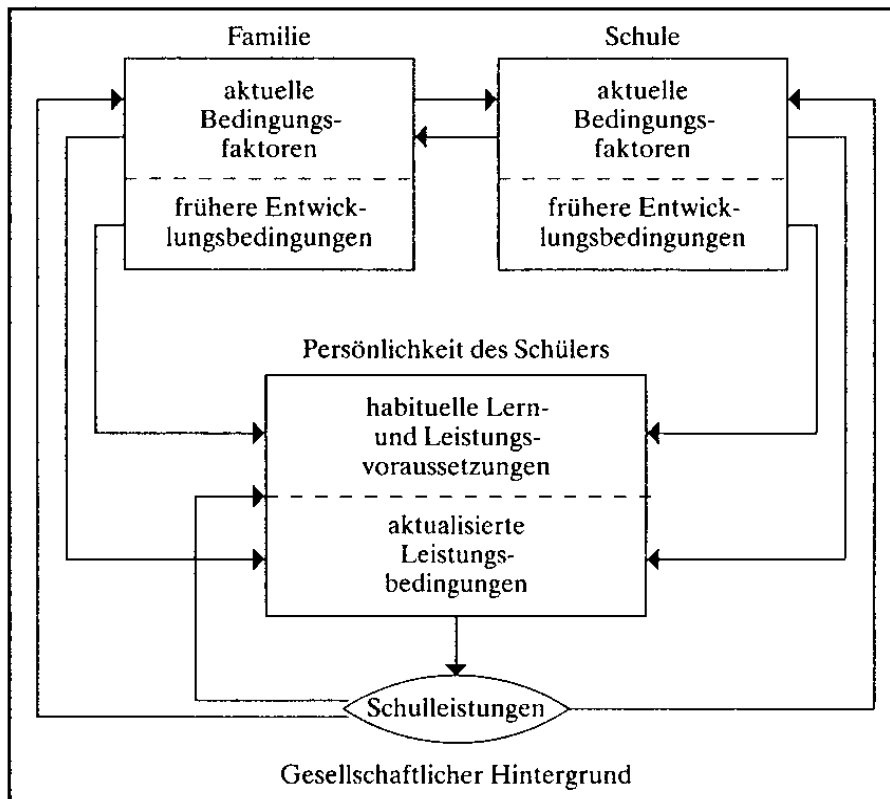


Abbildung 1: Bedingungsfaktoren der Schulleistung (Krapp 1986, S. 623)

Allgemeine Maßstäbe für die Leistungsbeurteilung

Als Maßstäbe für die Leistungsbeurteilung kommen vor allem drei Bezugsgrößen in Frage, nach denen die Bewertung derselben Leistung unterschiedlich ausfallen kann.

- Der *individuelle Vergleichsmaßstab* 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 wenn ein angestrebtes Lernziel noch nicht erreicht wurde. Nach diesem Maßstab kann jeder Lerner und jede Lerngruppe erfolgreich sein, wenn die Leistung erkennbar gesteigert wird. Beim Vergleich von Lernfortschritten hätte nach diesem Maßstab eine Klasse in Wilhelmsburg die gleiche Chance auf Lernerfolge wie eine Klasse im Alstertal. Umgekehrt könnten nach diesem Maßstab individuelle Leistungen, die zwar im Vergleich zu anderen relativ hoch sind, aber keine Verbesserung gegenüber der ursprünglichen Leistung darstellen, als unangemessen eingeschätzt werden.

- Der *sachliche Vergleichsmaßstab* zielt auf das Erreichen vorgegebener 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, wie viel Vorwissen die Lerner bereits hatten oder ob andere Lerner mehr oder weniger Lernzuwachs aufweisen. Die Erstellung solcher Mindeststandards (sog. Benchmarks), die von allen Schülern bzw. allen Klassen erreicht sein sollen, ist allerdings noch eine Aufgabe, die spätestens seit der Diskussion der LAU-Ergebnisse in Hamburg und der PISA-Ergebnisse in allen Bundesländern als dringlich angesehen und von der KMK vereinbart wurde. Die systematische Auswertung von Vergleichsarbeiten kann hierfür eine wichtige Grundlage liefern.
- Der *soziale Maßstab* zielt auf den Vergleich verschiedener Lerner oder Lerngruppen: Ist der Lernzuwachs eines Lernenden oder der Lerngruppe im Vergleich zu anderen Lernern höher oder niedriger? Bei dieser Abschätzung wird die individuelle Lernentwicklung einer sozialen Bezugsgruppe (z.B. Klasse, Schulform, nationale Altersgruppe) zugeordnet. Nach diesem Maßstab gibt es stets „Gewinner“ und „Verlierer“, denn es gibt stets einzelne Lerner oder Schulklassen, die bei diesem Vergleich besser bzw. schlechter abschneiden. Lehrkräfte beurteilen Ihre Schüler meist nach einem Klassen- oder Schulmaßstab, der sich auf ihr jeweiliges Umfeld bezieht. Dabei ist zu berücksichtigen, dass sich diese Lernmilieus sehr deutlich voneinander unterscheiden können. So erzielte z.B. bei der Hamburger PLUS-Untersuchung der Rechtschreibleistungen der leistungsbeste Schüler einer Klasse nicht das Niveau des schwächsten Schreibers in einer anderen Klasse.

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

Welche Instrumente sind in den Schulen vorhanden?

Üblicherweise sollen sich die Leistungsbewertungen in der Schulklasse an den Richtlinien und Lehrplänen für die einzelnen Fächer in der jeweiligen Klassenstufe und Schulform sowie an den Bestimmungen der Zeugnis- und Versetzungsordnung orientieren. Darüber hinaus wurden in Hamburg in den letzten Jahren einige Orientierungshilfen für die Gewinnung von Beurteilungsmaßstäben entwickelt (z.B. Handreichungen zu Vergleichsarbeiten und zur schulischen Evaluation) und Vorschläge zur Differenzierung der Leistungsbewertung erarbeitet (siehe Beutel, Lütgert, Tillmann & Vollstädt 1999). Da es in der Vergangenheit an überregionalen Vergleichsmaßstäben mangelte, wurden in Hamburg flächendeckende bzw. repräsentative Leistungserhebungen durchgeführt (z.B. für die Grundschule: PLUS und IGLU, für die Sek. I und II: LAU und ULME), die einen Vergleich der individuellen Leistungen und der Leistungsverteilung in den einzelnen Klassen und Schulen mit denen aller Hamburger Schüler ermöglichen. Allerdings steht die Systematisierung der einzelnen Erhebungsinstrumente, die Vervollständigung der einbezogenen Lernbereiche und die unterrichtspraktische Aufbereitung der verfügbaren Referenzdaten noch aus. Dies wird eine wesentliche Aufgabe des zukünftigen Landesinstituts für Bildung sein.

Lehrkräfte als Diagnostiker

Lehrkräfte sind professionelle Experten für die Einschätzung individueller Lernstände. Sie benötigen die Beurteilung des Lernstands, um Schüler im Unterricht gezielt zu fördern und über die Lernleistungen regelmäßig schriftlich berichten zu können (z.B. in Zeugnissen und Entwicklungsberichten). Im Unterschied zur Anwendung standardisierter Diagnoseverfahren haben Lehrkräfte dabei stets die gesamte Persönlichkeit des Lernenden im Blick und beurteilen die Ergebnisse auf dem Hintergrund längerfristige 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 (siehe: Ingenkamp 1997).
- Einschätzungen von Lehrkräften unterliegen systematischen Wahrnehmungsstereotypen (z.B. Halo- bzw. Ausstrahlungseffekt, klasseninterner Beurteilungsmaßstab, Pygmalion-Effekt), die ihre Vergleichbarkeit einschränken können (siehe: Schrader & Helmke 2001).
- Darüber hinaus wirken soziale und emotionale Faktoren 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 immer auch das Selbstkonzept der Lehrkräfte bzw. die Einschätzung ihrer Wirksamkeit betroffen, und diese Tatsache begünstigt – je nach Persönlichkeit – eine Tendenz zur selbstwertdienlichen Überschätzung bzw. „Euphemisierung“ der Lernergebnisse (May 2001).

Die Verbesserung der diagnostischen Qualität der Lehrereinschätzungen ist vor diesem Hintergrund eine wichtige Aufgabe für die Lehrerfortbildung, im deren Rahmen Lehrkräfte auch in der differenzierten und lernentwicklungsrelevanten Beurteilung von Schülerleistungen im förderdiagnostischen Zusammenhang geschult werden sollten.

Literaturhinweise

Beutel, S.-I., Lütgert, W., Tillmann, K.-J. & Vollstädt, W. (1999): Ermittlung und Bewertung schulischer Leistungen. Expertisen zum Entwicklungs- und Forschungsstand. Hamburg: Behörde für Schule, Jugend und Berufsbildung.

Ingenkamp, K. (1997): Lehrbuch der Pädagogischen Diagnostik. Weinheim: Beltz.

Krapp, A. (1986): Diagnose und Prognose. In: Weidenmann, B., Krapp, A., Hofer, M., Huber, G.L. & Mandl, H. (Hg.) (1986): Pädagogische Psychologie (S. 565-630). München: PsychologieVerlagsUnion.

May, P. (2001): Lernförderlicher Unterricht. Teil 1: Untersuchung zur Wirksamkeit von Unterricht und Förderunterricht für den schriftsprachlichen Lernerfolg. Er-

gebnisse der Evaluation des Projekts „Lesen und Schreiben für alle“ (PLUS).
Frankfurt a.M.: Peter Lang.

Schrader, F.-W. & Helmke, A. (2001): Alltägliche Leistungsbeurteilung durch Lehrer.
In: Weinert, F.E. (Hg.): Leistungsmessungen in Schulen (S. 45 – 58). Weinheim und
Basel: Beltz.

Weinert, F.E. (2001): Schulleistungen – Leistungen der Schule oder der Schüler? In:
ders. (Hg.): Leistungsmessungen in Schulen (S. 73 – 86). Weinheim und Basel:
Beltz.