

Leistungsmessung und Schülerbeurteilung

Josef Leisen und Dietmar Höttecke

Leistungsmessung und Schülerbeurteilung gehören zum Kerngeschäft der Lehrertätigkeit. Das Thema ruft aber seit jeher viele Fragen, heftige Diskussionen und starke Kritik hervor. Ein immer wieder vorgetragener Verdacht behauptet, dass sich die Leistungsmessung negativ auf das Lernen auswirke. Demgegenüber betont der Lernforscher F. E. Weinert (2001), dass Leistungen und Leistungsbeurteilungen für Schüler notwendige Erfahrungen sind, damit sie eine Leistungsbereitschaft entwickeln und ein möglichst positiv getöntes realistisches Selbstbewusstsein eigener Tüchtigkeit und Leistungsfähigkeit aufbauen. Lernsituationen und Leistungssituationen sollten jedoch auch aus der Schülerperspektive unterschieden und voneinander getrennt werden, da sie unterschiedlichen psychologischen Bedingungen unterliegen. Im Bewusstsein der Schüler müssen sie so separiert werden, dass eine produktive Lernkultur entsteht, die eine Leistungskultur mitdenkt. Bei der Leistungsmessung steht die Leistungsdiagnostik mit Verfahren der Leistungsfeststellung und deren Funktion im Vordergrund. Bei der Schülerbeurteilung sind es die Beurteilungsverfahren mit deren Bezugsnormen und Qualitätsmerkmalen.

Gütekriterien und Merkmale der Leistungsdiagnostik

Aus der psychologischen Testtheorie heraus werden seit Jahrzehnten drei Gütekriterien auf die pädagogische Diagnostik übertragen, nämlich Objektivität, Reliabilität (Zuverlässigkeit) und Validität (Gültigkeit). Untersuchungen zeigen, dass Schulnoten weder objektiv, reliabel noch valide sind, es nicht sein können und auch nicht sein müssen. Statt dieser testtheoretisch begründeten Gütekriterien braucht die Leistungsmessung in der Schule pädagogisch begründete und für alle Akteure transparente Gütekriterien. Die Denkweise der „pädagogisch günstigen Voreingenommenheit“ (Helmke, 2003, 89-90) sollte sich zunehmend durchsetzen. Unter pädagogisch günstiger Voreingenommenheit als Gütekriterium ist zu verstehen:

- Die Leistungsunterschiede in der Klasse werden nur mäßig unterschätzt.
- Die individuellen Lernmöglichkeiten werden leicht überschätzt.
- Erfolge werden subjektiv mit der Begabung des Schülers begründet, Misserfolge eher durch mangelnde Anstrengung oder eigene Fehler im Unterricht erklärt.

Pädagogisch günstige Voreingenommenheit ist auf die Förderung der individuellen Schülerleistung und den förderlichen Ausbau des Selbstkonzeptes ausgerichtet.

Die Qualität der Schülerbeurteilung zeigt sich in eine Reihe von theoretisch abgesicherten und in der Praxis bewährten Merkmalen (s. Abb. xxx).



Abb. xxx: Qualitätsmerkmale der Leistungsbeurteilung.

Was bedeutet das für den Unterricht?

Lernsituationen und Leistungsüberprüfungssituationen von einander trennen

Den Lernenden muss klar sein, ob sie sich in einer Lernsituation oder einer expliziten oder impliziten Leistungs- und Beurteilungssituation befinden (Qualitätsmerkmal Entflechtung). Verbale und nonverbale Signale sind hier tauglich. Eine abfragende Gesprächsführung, wie sie im fragend-entwickelnden Unterrichtsgespräch im Physikunterricht häufig anzutreffen ist (s. Kap. 10 Unterrichtskonzeptionenxxx), erzeugt eine Atmosphäre der Überprüfung. Eine diskursive Gesprächsführung mit offenen Fragestellungen, andere Sozialformen, der Einsatz von Methoden-Werkzeugen (s. Kap. 11bxxx) und Lernaufgaben (s. Kap. 15xxx) schaffen eine adäquate Lernsituation.

Was beurteilt wird muss vorher auch gelernt und geübt werden

Lernende haben ein außerordentlich gutes Gespür für Fairness in der Beurteilung. Aus den Qualitätsmerkmalen ergibt sich zwingend die Verpflichtung, das zu Überprüfende in angemessener Intensität vorher zu üben, damit der Lernende zeigen kann, „was er kann“ und nicht das Gefühl hat, dass ihm bloß gezeigt wird „was er alles nicht kann“.

Das ganze Spektrum der Kompetenzen der Bildungsstandards überprüfen

Gemäß den Bildungsstandards für das Fach Physik (≈ Kap. 3 xxx) gehören alle vier Kompetenzbereiche (Fachwissen, Erkenntnisgewinnung, Kommunikation, Bewertung) in die Schülerbeurteilung. Dazu müssen Aufgabenformate erstellt werden, die diese Kompetenzen sowohl in der Breite (Kompetenzbereiche) wie in der Tiefe (Anforderungsbereiche) überprüfen. Es empfiehlt sich, dazu Aufgabenteile und Formen mündlicher und praktischer Prüfung zu entwickeln, die in die Kompetenzmatrix der Bildungsstandards eingeordnet werden können (Erweiterte Beurteilung).

Leistungsmessung im Physikunterricht

Im Physikunterricht bieten sich folgende Formen der Leistungsfeststellung und -Beurteilung an:

- schriftliche Leistungsfeststellungen wie Klassenarbeiten, Tests, Vergleichsarbeiten, Protokolle, schriftliche Hausaufgaben, Facharbeiten, schriftliche Ausführungen, Portfolios, ...
- mündliche Leistungsfeststellungen wie mündliches Abfragen, mündlicher Vortrag, Referat, mündliche Beiträge, Erfahrungsberichte, mündliche Hausaufgaben, ...
- praktische Leistungsfeststellungen wie dokumentierte Schüler- und Hausexperimente, Projektergebnisse, Plakate, Präsentationen, ...

Für die Lehrkraft stellen sich im Physikunterricht einige besondere Fragen:

Wie stellt man eine Klassenarbeit zusammen?

Mit der Erstellung der Klassenarbeit sollte man Wochen vor der Arbeit am besten begleitend zum Unterricht beginnen. Jede Minute, die in die Erstellung investiert wird, erspart Korrekturzeit und Ärger. Die Klassenarbeit ist das Spiegelbild des Unterrichts. Sie muss die verschiedensten Unterrichtstätigkeiten und Lernvorgänge angemessen, ausgewogen aber umfassend widerspiegeln und alle Arten und Qualitäten von Schülerleistungen überprüfen (z.B. Experimente entwickeln und beschreiben, Schaltskizzen entwerfen, Basiswissen kennen, Hypothesen formulieren, logische Schlüsse ziehen, Rechnungen durchführen, Formeln herleiten, Diagramme zeichnen, interpretieren, Formeln diskutieren, etc.). Leistungsaufgaben vermeiden Verunsicherung, die dagegen bei Lernaufgaben manchmal produktiv sein kann.

Zur Überprüfung, ob die Klassenarbeit alle Kompetenzbereiche erfasst, bietet sich die Kompetenzmatrix an, indem die einzelnen Aufgabenteile in die Felder eingeordnet werden.

		Kompetenzbereiche			
		Fachwissen	Erkenntnisgewinnung/ Fachmethoden	Kommunikation	Reflexion/Bewertung
Anforderungsbereiche	I	Wiedergeben von Sachverhalten	Beschreiben und Einsetzen von Fachmethoden	Darstellen von Sachverhalte in vorgegebenen Formen	Angeben von Bezügen
	II	Anwenden von Sachverhalten eines abgegrenzten Gebietes	Anwenden von Fachmethoden	situationsgerechtes Anwenden von Kommunikationsformen	Herstellen einfacher Bezüge und Wiedergeben von Bewertungsansätzen
	III	problembezogenes Erarbeiten, Einordnen, Nutzen und Werten von Wissen	problembezogenes Auswählen und Anwenden von Fachmethoden	situationsgerechtes Auswählen und Einsetzen von Kommunikationsformen	Herstellen von Bezügen und Bewerten von Sachverhalten

Tab. 8.1: Kompetenzmatrix der Bildungsstandards Physik

Kommentar: Kann man hier vielleicht auf das Kapitel zu Bildungsstandards verweisen?

Folgende Empfehlungen sind zu berücksichtigen:

Die Arbeit ist im Aufbau klar gegliedert.

- Die Aufgabenabschnitte weisen das Thema, Prüfungsschwerpunkt und die Prüfungsqualität aus (sinnvolle Überschriften).
- Die Aufgabenformulierung enthält Signalwörter als Operatoren (z.B. zeichne, berechne, erkläre, nenne, diskutiere, beschrifte, beschreibe, skizziere, diskutiere, kommentiere, ...), deren Bedeutung den Schülern bekannt ist.
- Die Aufgaben erfassen und formulieren für die Schüler erkennbar alle Lernzielstufen (Reproduktion, Reorganisation, Transfer und problemlösendes Denken) in angemessener Weise.
- Der Aufgabentext ist sprachlich einfach und verständlich. Die Aufgabenlänge entspricht im Textumfang in etwa der Gewichtung der Teilaufgabe.
- Aus prüfungspsychologischen Gründen überschreitet der Aufgabentext zwei Seiten nicht.
- Der Anteil einzelner Aufgaben und Teilaufgaben an der Gesamtnote ist transparent (z.B. durch Angabe von Punkten).
- Arbeitsunterlagen (z.B. Versuchsskizzen, Diagramme, Tabellen, Texte, ...) werden entweder in eine Anlage übersichtlich und nummeriert untergebracht oder optisch übersichtlich in die Aufgabenstellung integriert.

Wie führt man eine mündliche Prüfung durch?

Die mündliche Prüfung ist nicht eine schriftliche Prüfung, die lediglich mündlich vorgetragen wird. Die mündliche Prüfung ist eine Leistung anderer Art, die in erster

Linie die Kompetenzbereiche Kommunikation und Bewertung/ Reflexion überprüft. Dazu bedarf es entsprechender Aufgabenformate, die im Unterricht gepflegt und geübt werden müssen. Bei der Aufgabenstellung sind zu vermeiden:

- reines Abfragen gedächtnismäßig fixierten Wissens
- umfangreiche, trickreiche Rechnungen, Ableitungen, etc.
- verkürzte Klassenarbeitsaufgaben
- trickreiche Fragen, die ein allzu langes Nachdenken erfordern
- Suggestivfragen und Fragen nach dem Motto: „Denk dir mal, was ich mir jetzt denke.“

Gute Aufgabenstellungen beinhalten Aufgaben, Materialien, Experimente, fachliche Probleme, Situationen, Fallbeispiele, Quellen, etc.

- die Anlass geben, in Fachkommunikation zu treten
- vielfältige fachliche Methoden tangieren
- Bezüge zwischen verschiedenen Gebieten herstellen
- dem Schüler Raum für verschiedenartige Antworten geben
- das Basiswissen kurz und auskunftssicher anstechen
- eine Fachkommunikation ermöglichen.

Beim Führen des Prüfungsgesprächs ist zu beachten:

- den Prüfling erst einige Minuten ohne Eingreifen sprechen lassen
- ermutigendes Zuhören, damit sich der Schüler einreden kann und seine Linie findet
- ggf. wichtiges für den weiteren Prüfungsverlauf (z.B. Formeln, Begriffe, Skizzen, ...) durch einen nonverbalen Wink an der Tafel fixieren lassen
- bei kleinen Fehlern nicht unterbrechen, es sei denn sie führen in eine Sackgasse
- auf Fehler nur hinweisen, damit evtl. eine Eigenkorrektur möglich ist und erläuternde Fehlerkorrekturen im Sinne einer Nachhilfe vermeiden
- Fehler als Rückmeldung an den Schüler zwecks Selbsteinschätzung kommentieren, z.B. mit 'falsch', 'das ist nicht ganz vollständig', im Kern richtig, aber unpräzise', nonverbale Hinweise bevorzugen.

Das Prüfungsgespräch muss den Charakter eines diskursiven Fachgesprächs haben, das Leistungspotenziale und –defizite feststellt.

Wie bewertet man Versuchsprotokolle?

Ein Protokoll ist die Niederschrift über den Verlauf und/ oder die Ergebnisse einer Veranstaltung und kann deshalb als Ergebnisprotokoll oder als Verlaufsprotokoll verfasst werden. Das Versuchsprotokoll ist eine Mischform. Alle Protokolle zeichnen sich durch formale Vorgaben aus, deren Berücksichtigung in die Bewertung eingeht.

Alle Kriterien, die bei der Bewertung des Versuchsprotokolls herangezogen werden, müssen den Schülern vorab bekannt sein:

- Vollständigkeit und Form
- Fertigkeiten beim Umgang mit Experimentiergeräten und beim Aufbau der Versuchsanordnung
- Bewältigen der Beobachtungs- und Messaufgabe, Festhalten der Messwerte oder Formulierung der Beobachtungen
- Ableiten von Schlussfolgerungen aus den gewonnenen Messwerten und den Beobachtungen,
- Erklären des Ergebnisses mit Hilfe vorhandener Kenntnisse,
- Herstellen des Bezugs zur Aufgabenstellung oder zur anfangs formulierten Hypothese,
- Nennen möglicher Fehlerursachen.

Falls das Protokoll im Sinne eines Lernprotokolls bewertet werden soll, muss es als solches angelegt sein. Dann werden Lern- und Erkenntniswege, Entwicklungen und Reflexionen bewertet und weniger die Berücksichtigung formaler Normen.

Wie bewertet man Tätigkeiten bei der Durchführung von Experimenten?

Bei der Durchführung von Experimenten können in besonderem Maße die Methoden der Erkenntnisgewinnung, die Arbeitsmethoden der Physik und auch die Kompetenzen aus dem Kompetenzbereich „Kommunikation“ bewertet werden. Schülerexperimente sind so ein- und durchzuführen, dass die gewünschten Kompetenzen von den Schülern gefordert und auch bewertet werden können. Die Bewertungskriterien werden am besten vorab mit den Lernenden erarbeitet:

- aktive Beteiligung an der Planung in der Gruppe
- Beschaffung von Materialien und Informationen (auch aus dem Internet)
- richtiger Aufbau der Versuche und Sorgfalt beim Experimentieren
- sachgerechter und pfleglicher Umgang mit den Arbeitsmaterialien
- Beachtung der Sicherheitsregeln beim Experimentieren und der Ordnung im Fachraum
- Offenheit für alternative Interpretationen und Variationen des Experiments
- Eigenständigkeit bei der Erarbeitung von Zusammenhängen
- Eigenständigkeit bei der Auswertung, sachgerechte Ergebnisdarstellung
- selbstständige Arbeitsorganisation (u.a. termingerechte Abgabe).

Die Kriterien lassen sich mit entsprechenden Indikatoren in Beobachtungsbögen erfassen.

Wie kann man die Methodenkompetenz von Schülern bewerten?

Die Methodenkompetenz der Schüler umfasst allgemeine Arbeitsmethoden (Methoden der Informationsbeschaffung und -verarbeitung, Darstellung und Präsentation von Sachverhalten, Kommunikation und Kooperation mit anderen), Arbeitsmethoden der Physik (intuitiv-spekulatives Entdecken, hypothetisches Formulieren, induktives Finden, deduktives Ableiten, analoges Übertragen, schöpferisches Modellbilden, kreatives Theoretisieren, systematisches Explorieren...) und die Lernmethoden des Schülers. Alle Methoden können und sollen im Physikunterricht bewertet werden. Die Arbeitsmethoden der Physik zu bewerten ist den Bildungsstandards und den Einheitlichen Prüfungsanforderungen entsprechend verpflichtend.

Die Methodenkompetenzen können auf folgende Arten bewertet werden:

- Bewertung durch den Lehrer anhand von entsprechenden Kriterienlisten.
- Bewertung in Form eines Gesprächs mit Selbsteinschätzung des Schülers.
- Metareflexion über die angewandten Methoden und Bewertung durch den Lehrer.

Wie kann ein Portfolio zur Leistungsmessung genutzt werden?

Portfolios sind vorrangig als Förderinstrument und nachrangig als Beurteilungsinstrument anzusehen. Ambivalente Wirkungen des Portfolio-Einsatzes gibt es, wenn sie doppelgleisig eingesetzt werden. Mit Portfolios dokumentieren Schüler nach Vereinbarung mit der Lehrkraft ihre individuellen Lernfortschritte. Sie sammeln in eigens angelegten Mappen ihre besten physikalischen Leistungen, gelungene Tests, originelle Hausaufgaben, dokumentierte Hausexperimente, bearbeitete Aufgaben, Rechercheergebnisse zu bestimmten Themen, originelle Fragen, Beobachtungen im Alltag, Zusammenfassungen, Sammlung von physikalischen Alltagsphänomenen, etc. Die Leistungsbewertung im Portfolio

- bewertet auch Verständnis, Durch- und Ausführungen
- reflektiert auch die Prozesse des Lernens
- bezieht eine Vielzahl von Methoden zur Dokumentation von Ergebnissen ein
- ermöglicht bereits während eines Lernprozesses eine Selbstevaluation
- dient unter Beschreibung von Stärken und noch auszugleichenden Defiziten insbesondere der weiteren Förderung des Lerners.

Portfolios lassen nicht so einfach wie bei Standardtests hinreichend eindeutig miteinander vergleichen und daher im Rangvergleich bewerten. Deshalb findet die Leistungsfeststellung im Rahmen eines Gesprächs statt, das am Ende mit einer Note abschließt. Das Gespräch über Produkte und Prozesse gibt Rückmeldungen zu Leistungsentwicklungen und bezieht die Ressourcen und individuellen Lösungen des Lerners mit in die Bewertung ein. Bei den Bewertungskriterien wird immer auch ein Wert für die formale Qualität der Arbeit gegeben, der nach äußerer Gestalt und

Handhabung, sachlich-formaler Richtigkeit und Stil aufgeteilt wird. Eine persönliche Note vermag das Portfolio von der Form her aufzuwerten. Formale Bewertungskriterien sollten nicht bloß isoliert erhoben werden, sondern Teil der sonstigen Bewertung sein und in einem konkreten Bezug zu anderen Kriterien stehen.

Wie setzt sich eine Physiknote im Zeugnis zusammen?

Die Festsetzung der Zeugnisnoten ist den jeweiligen landesspezifischen Schulordnungen geregelt, insbesondere die anteilige Berücksichtigung von schriftlichen, mündlichen und sonstigen, z. B. praktischen Leistungen.

- Die Zeugnisnote wird von dem zuständigen Fachlehrer festgesetzt, d.h. sie werden in der pädagogischen Verantwortung der Lehrkraft gesetzt und nicht berechnet.
- Die Gesamtnote soll durch eine hinreichende Zahl von Einzelnoten begründet sein, muss aber nicht dem rechnerischen Durchschnitt der Einzelnoten entsprechen. Einzelnoten können verschieden gewichtet werden, wenn dies durch den Schwierigkeitsgrad oder den Umfang der überprüften Leistungen begründet ist. Das Runden erfolgt nicht arithmetisch, denn Noten sind pädagogische Botschaften an die Lernenden.
- Es müssen nicht bei allen Schülern die gleiche Anzahl von Noten für sonstige Leistungen erteilt werden, sondern hinreichend viele.
- Nur die mitgeteilte Note ist eine erteilte Note!
- Die Notentransparenz hinsichtlich der Maßstäbe und der Gewichtung muss gewährleistet sein.

Die Physiknote im Zeugnis muss die Leistungen des Schülers in allen Aspekten und Facetten physikalischen Arbeitens über den Gesamtzeitraum des Schul(halb)jahres hinweg unter den gegebenen Lernbedingungen zutreffend erfassen.

Diagnostikfehler und Subjektive Fehlerquellen

Beurteilende Lehrkräfte wissen aus eigener Erfahrung um folgende Fehlerquellen:

- Da stellt man fest, dass man mit der Benotung der Klassenarbeiten nach hinten hin immer strenger (oder milder) wird.
- Da gibt man dem Schüler noch den Punkt, weil man sich nicht vorstellen kann, dass der so etwas Einfaches nicht weiß.
- Da ist man von dem ersten Teil der Bearbeitung so überrascht, dass der positive Eindruck die anderen Teile überstrahlt.
- Da gibt man dem Schüler eine Vier in der Epochalnote, weil er die doch immer hatte.

- Da werden schriftliche und mündliche Noten einander angepasst.
 - Da wird mit Tendenznoten (Plus/Minus) eine hohe Messgenauigkeit vorgegaukelt.
- Die Diagnostik von Lehrern fußt oftmals auf impliziten, hausbackenen, empirisch nicht belegbaren subjektiven Theorien, z.B. dass ein wacher, verständiger Gesichtsausdruck, eine ordentliche und saubere Arbeitsweise, ein geordnetes Elternhaus auf Intelligenz hindeuten und der Schüler besser benotet wird. Sympathie und Antipathie, eigenes Lernverhalten, Vorlieben und Persönlichkeitsprofile haben einen nachgewiesenen Einfluss auf die Leistungsdiagnostik. Nicht die Tatsache selbst ist erschreckend, sondern der unprofessionelle Umgang damit. Ein professioneller Umgang erfordert Wachsamkeit, Selbstevaluation und Reflexivität den Fehlerquellen gegenüber. Aus vielen empirischen Untersuchungen sind subjektive Fehlerquellen bekannt, von denen hier nur einige genannt werden:
1. *Effekt der Zusatzinformationen*: Negative oder positive Zusatzinformationen über einzelne Schüler beeinflussen die Korrektur und Benotung der Arbeit.
 2. *Sympathie-Effekt*: Etliche Lehrer bewerten Schüler, die ihnen sympathisch sind, zu günstig und diejenigen, die ihnen unsympathisch sind, zu ungünstig.
 3. *Effekt des ersten Eindruck*: Der erste positive oder negative Eindruck beeinflusst die nachfolgende Wahrnehmung des Schülers bis hin zur Bewertung.
 4. *Voreinstellungs-Effekt*: Beispiel: „Jungen sind in Physik begabter als Mädchen“.
 5. *Reihenfolgen-Effekt*: Die erste Note in einer Reihe von Prüfungen oder bei der Korrektur schriftlicher Arbeiten setzt den Maßstab für die nachfolgenden Noten.
 6. *Halo-Effekt*: Von einem hervorstechenden Merkmal oder Gesamteindruck wird auf andere, nicht direkt beobachtbare Merkmale geschlossen. Beispiel: "Wer in Mathematik gut ist, ist auch in Physik gut“.
 7. *Matthäus-Effekt*: „Wer hat, dem wird noch mehr gegeben“. Ein Schüler mit bereits guten Noten hat es einfacher, weitere gute Noten zu bekommen.

Literatur zur Vertiefung

- Franz E. Weinert (Hrsg.): Leistungsmessungen in Schulen. Weinheim und Basel: Beltz 2001
- Helmke, Andreas: Unterrichtsqualität erfassen, bewerten, verbessern. Seelze: Kallmeyer 2003.
- Becker, Georg: Unterricht auswerten und beurteilen. Weinheim: Beltz 1998
- Grunder, Hans-Ulrich u. Thorsten Bohl (Hrsg.): Neue Formen der Leistungsbeurteilung in den Sekundarstufen I und II. Hohengehren: Schneider 2001.