

Welche Formen der Leistungsbewertung können eine Lehrerbeurteilung objektivieren?

Antje Siebert

„Die Schwierigkeit, zu einem gerechten Urteil zu gelangen, belastet gerade die besten Lehrer. Angesichts der Bedeutung einer gerechten Beurteilung ist es erforderlich, sich intensiv um Möglichkeiten einer Objektivierung des Lehrerurteils zu bemühen.“ (Deutscher Bildungsrat)

In der Literatur finden sich zahlreiche Formen der Leistungsbewertung. Um für den jeweiligen Lernprozess bzw. das jeweilige Lernergebnis die angemessene Form der Bewertung auswählen zu können, müssen u.a. die folgenden Fragen geklärt werden:

a) Welche Ziele sollen mit der Leistungsbewertung erreicht werden?

- Selektion
- Prognose und Steuerung
- Festhalten des Leistungsstands

b) Welche Art der Leistung soll bewertet werden?

- Fähigkeiten, Fertigkeiten
- musische, sprachliche, mathematisch-naturwissenschaftliche, sozial-kundliche
- kognitive, affektive, soziale

c) Anhand welcher Tests bzw. Aufgaben sollen die Leistungen erfasst werden?

- standardisiert, nicht standardisiert
- mündlich, schriftlich

d) Welche Merkmale werden dem Bewertungsprozess zugrunde gelegt?

- Fehler
- korrekte Antworten
- lernzielorientierte Kriterien

e) Nach welcher Bezugsnorm erfolgt die Bewertung?

- sachlich = kriteriumsorientiert
- individuell = lernprozess-orientiert
- Sozial = normorientiert

f) Wie wird das Bewertungsergebnis festgehalten?

- Noten
- Bericht

Zur Beantwortung der Fragen sollen folgende Modelle dienen:

- **Richtlinien-Modell vs. Maßstab-Modell**
- **Standardisierte vs. Nicht-Standardisierte Verfahren**
- **Verbalzeugnisse vs. Notengebung**

Richtlinien-Modell vs. Maßstab-Modell

Die erste zu treffende Unterscheidung ist die in **Richtlinien-** und **Maßstab-Modell**. Beide werden vorwiegend bei der Korrektur von Klassenarbeiten verwendet, welche nach wie vor die Hauptform der Überprüfung schriftlicher Schülerleistung ist.

Richtlinien-Modelle

Richtlinien-Modelle werden im Bereich solcher Schülerleistungen genutzt, die von den Lehrern als schwer quantifizierbar eingeschätzt werden. Das sind vorwiegend Leistungen sprachlicher Fächer. Die Ermittlung der Abweichung von den Erwartungen bezeichnet man durch subtraktive und additive Formen. Richtlinien-Modelle führen meist unmittelbar zu Noten. Diese sind als Versuch der Quantifizierung des Bewertungseindrucks zu verstehen. Hierbei ist allerdings zu beachten, dass in seltenen Fällen nur eine Erwartung existiert, meist ist es ein ganzes Erwartungsbündel.

Subtraktives Richtlinien-Modell

Die **subtraktive Form** ist an fehlerhaften Sachverhalten oder an Mängeln orientiert. Bei ihr werden durch den Lehrer Verstöße gegenüber einer Norm oder Erwartung registriert. Lassen sich keine solche Verstöße feststellen, stellt die Arbeit eine sehr gute Leistung dar. Mit wachsender Anzahl der Verstöße sinkt die Bewertung der Leistung.

Eine Lehreraussage:

„Ich arbeite eigentlich eher nach dem Abstrichverfahren. Ich sehe also, was ein Schüler alles hätte machen können, und dann streiche ich ab, wieviel von seiner Arbeit dem eigentlich noch entspricht. Dann bleibt eben mehr oder minder viel übrig. Ja, das könnte man so beschreiben, dass man von einem Idealbild ausgeht.“

Additives Richtlinien-Modell

Die **additive Form** ist an richtigen Sachverhalten orientiert. Es wird also registriert, welche Sachverhalte der Schüler im Hinblick auf eine Erwartung oder Norm richtig dargestellt hat. Durch additive Zusammenfassung dieser richtigen Sachverhalte gelangt man daraufhin bezüglich der Erwartung zu einer Beurteilung.

Eine Lehreraussage:

„Da könnte man ein Raster anlegen, und sagen, welche Aussagen die Schüler bei der Beschreibung hätten liefern müssen.“

Maßstab-Modelle

Den **Maßstab-Modellen**, das heißt, den anderen Typen der Beurteilungsmodelle, liegt im Regelfall eine Bewertung nach Punkten, die ein Schüler in einer Klassenarbeit erreicht hat, oder nach Fehlern, die der Schüler gemacht hat, zugrunde. So ergeben sich zwei Grundformen:

Punkte-Modelle und **Fehler-Modelle**.

Sie werden im Bereich von leichter quantifizierbaren Schülerleistungen verwendet, also in Mathematik und Naturwissenschaften.

Beide Modelle führen im Gegensatz zu Richtlinien-Modellen nicht unmittelbar zu Noten. Es muss entschieden werden, welchen Punkt- bzw. Fehlerintervallen welche Notenstufen zugeordnet werden.

Dabei repräsentieren die beiden Modelle in dieser Hinsicht detailliertere Formen der erwarteten Schülerleistung.

Punkte-Modelle

Bei Punkte-Modellen gibt es zwei Ansätze, die sich als „**Additives Punkte-Modell**“ und „**Subtraktives Punkte-Modell**“ beschreiben lassen.

Additives Punkte-Modell

Dieses Modell besagt, dass der Schüler bei einer bestimmten Aufgabe eine – im Regelfall – vorher festgelegte Punktzahl erreichen muss. Die Summe der Aufgabenpunktzahlen ergibt die maximal mögliche Gesamtpunktzahl. Eine individuelle Schülerleistung lässt sich dabei dann im Hinblick auf diese Gesamtpunktzahl beurteilen. Additive Punkte-Modelle gehen im Schwerpunkt von richtigem Sachverhalten aus. Dabei korrespondieren sie mit der additiven Form der Richtlinien-Modelle. Dabei sind Unterscheidungen zu treffen, ob die Aufgabe beispielsweise insgesamt betrachtet wird, oder in Teilaufgaben zu zerlegen sei. Auch die Frage, ob falsche Sachverhalte zu Punktabzügen führen sollen, ist von Bedeutung.

Eine Lehreraussage:

„In den letzten Schuljahren habe ich in der Regel die Mathematikarbeiten der Mittelstufe mit Bewertungseinheiten, also Punkten, belegt. Diese Festlegung der Punkte geschehen in der Regel schon bei der Planung der Arbeit. Sie orientiert sich am Umfang dessen, was für eine einzelne Aufgabe zu leisten ist. Die Schwierigkeit der Aufgabe spielt dabei keine Rolle. Die schwerere Aufgabe bekommt nicht notwendigerweise mehr Punkte, sie bekommt in etwa gleich

Aufgaben kommen bei der Umsetzung von Punkten zu Notenstufen sowieso zum tragen.“

Subtraktives Punkte-Modell

Dies bedeutet, dass für einen bestimmten Aspekt zum Beispiel 10 Punkte vorgegeben werden. Verstöße gegenüber festgelegten Anforderungen, Normen oder Erwartungen führen dann zu einem Punktabzug. In dieser Hinsicht korrespondiert das subtraktive Punkte-Modell mit der subtraktiven Form der Richtlinien-Modelle.

Eine Lehreraussage:

„Der Bewertungsmaßstab wird meist im groben Rahmen bei der Planung festgelegt, bei der Auswertung können dann Korrekturen erfolgen. Das heißt, wir legen Punkte fest und versuchen, die Arbeit nach diesen Punkten auszuwerten.“

Fehler- Modelle

Die **Fehler-Modelle** sind dadurch gekennzeichnet, dass im Beurteilungsprozess Fehler ausgezählt werden. Es lassen sich hierbei 2 Formen unterscheiden: das **einfache Fehler-Modell** und das **Fehler-Index-Modell**. Bei beiden Formen ist zu entscheiden, ob alle Fehler gleich gewichtet werden, oder ob bestimmte Fehler höher (zum Beispiel lernzielrelevante) oder niedriger (zum Beispiel Interpunktionsfehler) gewichtet werden.

Eine Lehreraussage:

„Wir haben beschlossen, dass bestimmte Arten von Fehlern als 1,5 Fehler gewertet werden sollen.... Ebenso haben wir definiert, was als leichter Fehler gelten soll und was als Wiederholungsfehler gilt.“

Das einfache Fehler-Modell

Hierbei bilden die gezählten Fehler bereits die Basis der Notengebung.

Eine Lehreraussage:

„Wir haben es in der Mathematik einfach, jeder Aufgabe eine bestimmte Punktzahl zuzuordnen, und können daher dann der Punktzahl wieder die Note zuordnen. Wir gehen in der Regel so vor, dass wir bei der Hälfte der zu erreichenden Punktzahl die Note ausreichend erteilen, gehen aber auch nach oben und nach unten, wenn es erforderlich ist, z.B. wenn wir erkennen: die Arbeit war wirklich zu schwierig oder zu leicht, dass wir dann den Maßstab versetzen.“

Das Fehler-Index-Modell

Dabei wird die Fehlerzahl im Hinblick auf eine andere gezählte Größe zu einem Indexwert relativiert. Das heißt beispielsweise, dass der Fehlerquotient sich aus der Zahl der Wörter insgesamt, die durch die Zahl der Fehler dividiert wird, errechnet. Bei der Zuordnung von Punkten- bzw. Fehlerintervallen zu Noten erfolgt jedoch keine explizite Orientierung am Klassendurchschnitt. Das kann häufig zu Problemen bei der Notengebung führen, da sich beispielsweise an einer Intervallgrenze bestimmte Punkte- oder Fehlerzahlen häufen, so dass etwa gleichstarke Schüler unterschiedliche Noten erhalten. Leichte Modifizierungen würden zu gleichen Noten bei etwa gleichen Leistungen führen.

Eine Lehreraussage:

„Über die Punktvergabe mache ich mir schon bei der Planung Gedanken. Aber den Bewertungsmaßstab selbst, das mache ich unmittelbar vor der Korrektur. Da lege Ich einen Zettel an – Punkte, Gesamtzahl – und dann wird die Einteilung nach Noten getroffen.... Bevor ich das erste Heft in die Hand nehme, liegen die Noten fest, ohne Rücksicht darauf, wie die Arbeit ausfällt.“

Standardisierte vs. Nicht-Standardisierte Verfahren

Bei diesen Verfahren ist die Unterscheidung zu treffen, ob sich das Messen der Leistung an einem **normierten Verfahren** oder an einer vom Lehrer angestellten **Beobachtung** orientiert.

Standardisierte Verfahren

Häufig werden unzuverlässige Techniken der Beurteilung der psychomotorischen Fertigkeiten oder der Fähigkeit zum Problemlösen eines Schülers verwendet. Da dies häufig zu ungerechten Einschätzungen führen kann, sollten Verfahrensregeln aufgestellt werden, die allen untersuchten Kindern vergleichbare Chancen ermöglichen. Dies kann durch standardisiertes Testen geschehen.

Standardisierte Tests zeichnen sich dadurch aus, dass sie testtheoretische Gütekriterien erfüllen und somit **objektiv**, **reliabel** sowie **valide** sind. Deswegen werden sie mit Hilfe großer repräsentativer Stichproben geeicht.

Ein standardisierter Test ist ein Test, der einer großen und repräsentativen Stichprobe aus einer bestimmten Population vorgelegt wurde. So können die Werte in dem Test mit denjenigen der Person dieser Stichprobe verglichen werden.

Normbezogenes Testen

Normbezogenes Testen bedeutet, dass die Testleistungen verschiedener Personen durch die Wahl desselben Messens als Grundlage für die Interpretation der Ergebnisse verwendet werden.

Dadurch können alle Individuen miteinander verglichen werden. Leistungsschwächere Schüler werden von leistungsstärkeren unterschieden, was beispielsweise bei der Vergabe von Stipendien oder Positionen von Vorteil ist. Ein Nachteil kann allerdings darin gesehen werden, dass es zu verzerrten Ergebnissen kommen kann, wenn in der Normgruppe (Bezugsgröße für Beurteilung von Testwerten) zu wenige Mitglieder von Minderheitsgruppen enthalten sind.

Beispiele hierfür sind:

- PISA-Studie
- Intelligenztests
- Schuleignungsprüfungen
- Tests der allgemeinen Begabung

Kriteriumsbezogenes Testen

Kriteriumsorientierte Schulleistungstests sind Tests, bei denen das individuelle Ergebnis in Relation zu einem vorher gesetzten Kriterium mitgeteilt wird.

Dabei soll die Frage untersucht werden, ob und eventuell wie gut ein bestimmtes Lehrziel erreicht wurde, und was ein Schüler im Vergleich zur Peergruppe zu leisten vermag. Der kriteriumsbezogene Test wird also immer so konstruiert, dass er zu einer bestimmten Information führt. Die dabei verwendeten Testaufgaben sind nicht identisch mit dem Lehrziel, sondern repräsentieren es nur. Sie dienen dazu, den individuellen Fähigkeitsgrad eines Schülers mit einem gewünschten Fähigkeitsgrad zu vergleichen.

Für diesen Vergleich sind folgende Voraussetzungen erforderlich:

- Quantifizierung des Lehrziels
- Quantitative Erfassung der Schülerleistung
- Messmodell für zufallskritische Entscheidung über Erreichung des Lehrziels.

Diese Tests messen also die Fähigkeit oder Eignung eines Individuums im Hinblick auf ein bestimmtes Kriterium. Die Kriterien werden schon vor dem Test festgelegt. So bleiben sie von den Leistungen der Schüler während des Testens unbeeinflusst.

Da es bei der Messung dieser Tests um das Erreichen eines Lehrziels geht, wird die Methode häufig auch „lehrzielorientierte“ Leistungsmessung genannt. Ein Test soll hierbei inhaltlich gültig sein, das heißt, er soll genau die Inhalte prüfen, zu deren Prüfung er konstruiert wurde. Die dabei gewählten Testaufgaben sollten jedoch möglichst die Gesamtheit der möglichen Testaufgaben repräsentieren.

Beurteilungsergebnisse, die dabei entstehen, werden meist in die Abstufungen „bestanden und nicht bestanden“, oder die der Notenskala unterteilt.

Nicht-Standardisierte Verfahren

Diese Art des Testens wird notwendig, wenn der Lehrer den Schüler anhand der Lernziele beurteilen möchte, die er sich für den eigenen Unterricht, und für die Leistungen seiner Schüler gesetzt hat.

Diese Notwendigkeit ist gegeben, da jeder standardisierte Test etwas prüfen kann, das in der speziellen Klasse nicht gelehrt wurde, und daher unfair ist. Hierbei ist es nicht wichtig, nicht Unterschiede zwischen den Schülern hervorzuheben, sondern die Leistungen der Schüler anhand von Lernzielen zu bewerten. Das Ergebnis des Testes gibt dann darüber Aufschluss. Auf die verschiedenen Arten des vom Lehrer erstellten Tests, im Besonderen auf die informelle Beurteilung, soll im Folgenden eingegangen werden.

Informelle Beurteilung

Die informelle Beurteilung gibt dem Lehrer anhand nicht normbezogener Kriterien Einblick in das Leistungsverhalten seiner Schüler. Sie informiert ihn über den Leistungsstand der Schüler in Bezug auf seine Lernziele (Schülereinschätzung) und über die Effektivität seines Unterrichts (Unterrichtseinschätzung).

Schülereinschätzung

Diese Art der Beurteilung vermittelt praktische Einsichten in die kognitiven, affektiven und psychomotorischen Fertigkeiten des Schülers.

Diese Einschätzungen sind jedoch vorläufig, da sie immer von bestimmten Situationen abhängen. Der Lehrer sollte sich hierbei immer ein eigenes Urteil bilden, und nicht das anderer übernehmen. Es stellte sich heraus, dass Lehrer bei der Einschätzung ihrer Schüler meist ein gerechtes Urteil über die kognitiven Fähigkeiten fällten, und ein ungerechtes über emotionale Stabilität, das Selbstbild des Schülers und seine Interes-

sen. Häufig fließen auch kulturelle Stereotypen mit in das Bild ein, zum Beispiel Vorurteile gegenüber Mädchen etc..

Unterrichtseinschätzung

Die zweite Form der informellen Beurteilung ist die **Unterrichtseinschätzung**. Dabei soll der Lehrer die Schüler beobachten und in deren Verhalten den Erfolg seiner Arbeit sehen.

Da sich aber die Unterrichtsbedingungen ändern, kann es zu keinen verlässlichen Einschätzungen kommen. Auch die Tatsache, dass der Lehrer bestimmte Reize nur von machen Schüler aufnimmt, und andere Schüler vorurteilsgesteuert betrachte, kann ihm kein genaues Bild des Unterricht vermitteln. Der informelle Eindruck sollte durch formale Beweismittel wie beispielsweise Hausaufgaben ergänzt werden. Des Weiteren sollte eine ständige Kontrolle erfolgen, ob die Beurteilungsbasis breit genug gefächert ist, und ob alle Lernfortschritte objektiv betrachtet werden können.

Verbalzeugnisse versus Notengebung



Sowohl Verbalzeugnisse als auch Ziffernoten dienen als Anreiz, Warnung, Orientierungshilfe, Auslesekriterium, Berechtigungsnachweis, Disziplinierungsmaßnahme, Ermutigungs-, Benachrichtigungs-, Belohnungsinstrument.

Zum Thema Zeugnisgebung gibt es eine Fülle von Veröffentlichungen, die meist kritisch zu der bestehenden Praxis Stellung nehmen, ohne dass sich inzwischen an der Praxis viel verändert hätte. Über die Funktionen, die Zeugnisse üblicherweise zugeschrieben werden, ist man sich in der Regel einig.

Verbalzeugnisse

Verbalzeugnisse dienen des weiteren auch als Förderung einer stärkeren Individualisierung des Unterrichts in der Grundschule. Der Lernprozess des einzelnen Schülers und eine auf konkrete Lernziele bezogene Leistungserfassung sollen dabei im Mittel-

punkt stehen. Die Verbalgutachten beziehen sich neben dem Leistungsstand auch auf nichtkognitive Fähigkeiten in den Dimensionen „Sozialverhalten“ und „Arbeitsverhalten“. Da die Einführung der Verbalbeurteilung ohne eine gezielte Evaluation erfolgte, sind die Kenntnisse über Formen und deren Wirkung unsystematisch.

Ziele der Verbalbeurteilung

1. ermutigende Erziehung statt Leistungsdruck
2. Förderung der sozialen Kooperation statt Konkurrenzverhalten
3. Erhöhung der Chancengleichheit statt Leistungsabfall der Benachteiligten
4. individuelle Förderung statt frontal gesteuertem Lerngleichschritt

Typologie der Verbalzeugnisform

1. **Normatives Zeugnis:** Die Lehrperson nimmt die Beurteilung des Schülers nach den feststehenden Unterrichtszielen vor.
2. **Schönes Zeugnis:** Entspricht einer ermutigende Pädagogik, ist allerdings wenig informativ.
3. **Deskriptives Zeugnis:** Hier erfolgt eine möglichst objektive Beschreibung des Gewesenen, jedoch finden sich keine Hinweise zur Leistungsverbesserung.
4. **Entwicklungsbericht:** Dieser ist nur sehr selten anzutreffen, obwohl er dem Anspruch an ein Verbalzeugnis am ehesten entspricht.

Inhalte von Verbalurteilen

In **Verbalurteilen** werden Auskünfte über das **Verhalten gegenüber Mitschülern, Lehrern und gegenüber Sachen**, über das **Arbeitsverhalten** in der Gruppe und als Einzelner (Ausdauer, Aufmerksamkeit), sowie über das **Lernverhalten** (Sprachverständnis, Ausdrucksfähigkeit, Beherrschung der Motorik und der Sinne und kognitive Leistungen) erteilt.

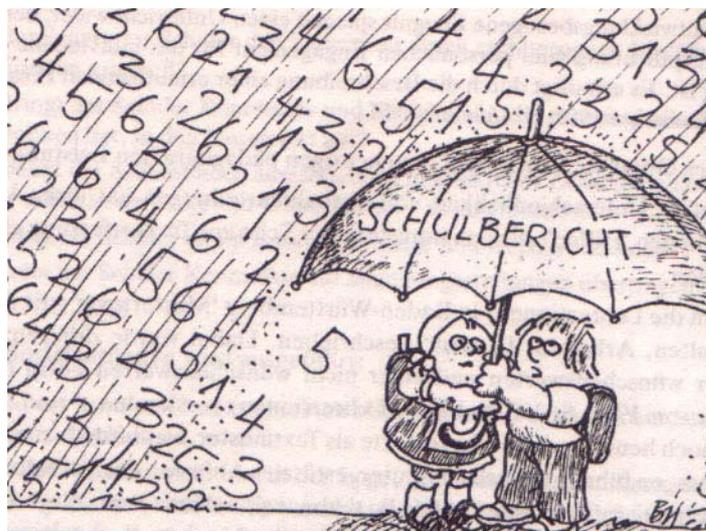
Diese Beurteilung soll nur beschreiben und nicht bewerten, wobei jedoch unklar bleibt, wie der Lehrer zu seinem Urteil kommt und welche Situationen ihm dabei wichtig erscheinen.

Ziffernzeugnisse

Bei der Erstellung von **Ziffernzeugnissen** misst der Lehrer das Schülerverhalten spontan an seinen Anforderungen und am Klassendurchschnitt. Schüler mit gleichen Noten haben damit automatisch gleiche Rangplätze. Hinzu kommt, dass Zensurenskalalen Schätzskaalen (Ordinalskalen) sind, die keine Aussagen über die Größe der Unterschiede machen. Dies führt dazu, dass eine Verstärkung der Grenz- und Auslesefunktion erfolgt. Die schwache Differenziertheit der Skala bedingt eine enge Anbindung an die Polarität „gut“ und „schlecht“. Da bei der Zensurengebung keine Zwischennoten erteilt werden, muss sich der Lehrer aufgrund seiner eigenen Einschätzung entweder für die bessere oder schlechtere Note entscheiden.

Ziele der Ziffernzeugnisse

1. Auskunft über den Leistungsstand
2. Prognose und Steuerung der schulischen Laufbahn
3. Erteilung von positiven oder negativen Sanktionen für Leistung



„Klasse Schirm – leider hält er bloß zwei Jahre“
(Zeichnung: Bühs)

Doch sind Ziffernzeugnisse überhaupt in der Lage, die genannten Funktionen angemessen zu erfüllen?

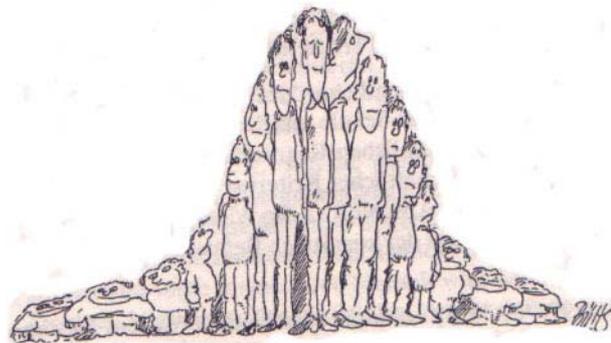
Charakteristik der Notengebung

Die Bedeutung der Noten ist oft unklar, weil in der Regel nicht zu ersehen ist, in welchem Maß tatsächlich erbrachte Leistung oder aber Fleiß oder auch gutes Betragen am Zustandekommen beteiligt waren.

Noten müssen **logisch eindeutig** sein, das heißt, auf gleiche Leistungen müssen immer gleiche Noten vergeben werden, wobei gleiche Noten durchaus unterschiedlichen Leistungen entsprechen können. Da dem Lehrer nur sechs Benotungsstufen zur Verfügung stehen, fassen Zensuren immer einen ganzen Leistungsbereich zusammen.

Die Note 3 steht ebenso für die 3, die schon fast eine 2 ist, wie für die 3, die an der Grenze zur 4 steht.

Zensuren können kaum dem **Gesetz der Normalverteilung** untergeordnet werden, das heißt also, dass man nicht von Abweichungen sprechen kann, die vom Mittelwert so streuen, dass geringe Abweichungen häufiger, größere Abweichungen seltener und extreme Abweichungen sehr selten sind; theoretisch müssten sich aber Schulleistungen „normal“ verteilen.



„Normalverteilung von Schülern“ (Zeichnung: Bühs)

Zensuren geben keinen Aufschluss über Begabung, Eignung, Fähigkeit, Neigung, Bildsamkeit und Verhalten eines Schülers. Es erfolgt keine differenzierte Information über den Schüler.

Heißt die Note 3, dass der Schüler auf allen Teilgebieten durchschnittliche Arbeit geleistet oder dass er in einigen Teilgebieten über-, in anderen unterdurchschnittliche Leistungen erbracht hat?

Sie verursachen Konkurrenzdenken, stehen kreativem Arbeiten im Wege und fördern das Lernen der Note wegen. Unklar bleibt auch, welche Maßstäbe der jeweiligen Note zugrunde liegen.

Die Beurteilung mit Hilfe von Zensuren ist weder objektiv noch reliabel, die Validität im Sinne der inhaltlichen Vorhersagegültigkeit ist nicht genügend hoch.

Praktiken der Notengebung

1. Es gibt Lehrer, die Klassenarbeiten mit einer Aufgabenzahl ausstatten, die ein Mehrfaches von 6 darstellt.
2. Andere verwenden eine „Proportional-Notenskala“, wobei Anzahlen von richtigen Lösungen bestimmt und Notenwerten zugeordnet werden. Wenn sich herausstellt, dass die Arbeit insgesamt sehr leicht oder schwer war, muss dann nachträglich korrigierend eingegriffen werden.
3. Die Methode der „natürlichen Einteilung“ versucht, bei der Betrachtung der Verteilungskurve „natürliche“ Einschnitte festzustellen.
4. Bei der Zeugnisgebung müssen natürlich verschiedene Daten und Eindrücke zu einer Zeugnisnote verarbeitet werden. Es müssen Gesichtspunkte wie Mitarbeit, Erledigung der Hausaufgaben, Kurzarbeiten und mündliche Überprüfungen berücksichtigt werden. Deshalb sollte der Lehrer die Gewichtung der unterschiedlichen Bewertungskomponenten auch mit seinen Schülern besprechen.

So zum Beispiel:

Standardisierter Leistungstest am Ende des Schuljahres	25%
Informelle Tests im Laufe des Schuljahres	30%
Mitarbeit und Diskussion in Gruppen	25%
Erledigung der Hausaufgaben	20%

Egal für welche Form der Leistungsbeurteilung man sich entscheidet, es sollte versucht werden, so **objektiv wie möglich** zu bewerten und **Vieldeutigkeiten** zu **vermeiden**.

Um der wichtigsten Funktion der Zeugnisse, nämlich der Informationsvermittlung gerecht zu werden, sollten Beurteilungs- und Zensierungssysteme gemeinsam mit Lehrern, Eltern und Schülern entwickelt werden.

Und wenn schon Ziffern zur Zeugnisgebung verwendet werden, sollten sie sich auf den aktuellen Leistungsstand des Schülers beziehen und möglichst nicht auch noch andere Faktoren beinhalten, wie Fleiß, Betragen, Mitarbeit, relativer Lernfortschritt. Nur mit Unterstützung durch zusätzliche Information kann das gegenwärtige Zensierungssystem seiner wichtigsten psychologischen und pädagogischen Funktion einigermaßen gerecht werden.

Literatur:

Gage, N.-L./ Berliner, D.-C. (1996) Pädagogische Psychologie. Weinheim: Psychologie Verlag- Union

Ingenkamp, K. (1988). Lehrbuch der pädagogischen Diagnostik. Weinheim: Beltz.

Ingenkamp, K. (1995) Die Fragwürdigkeit der Zensurenggebung. Weinheim: Beltz

Nuding, A. (1997). Beurteilen durch Beobachten: pädagogische Diagnostik im Schulalltag. Hohengehren: Schneider

Pache, D. (1978). Einstellung und Schulleistungen: Vergleichende Untersuchungen zwischen Sportunterricht und Hauptfächern. München: Ernst Reinhardt

Pfistner, H.-J. (1981). Lehrer beurteilen Schüler: Beiträge zur Beurteilungspraxis und zur Theorie der Leistung. Frankfurt/ Main: Peter Lang

Rapp, G. (1974). Messung und Evaluierung von Lernergebnissen in der Schule. Bad Heilbronn: Klinkhardt

Weinert, P.-E. (2001). Leistungsmessungen in Schulen. Basel: Beltz