

Leistung gerecht bewerten

Leistungsmessung spielt eine immer größere Rolle im Schulalltag. Die Vergleichbarkeit schulischer Leistungen zu gewährleisten und Schulqualität zu sichern sind die wesentlichen Ziele. Die Leistungsbewertung muss dabei möglichst objektiv und transparent sein. Doch kein anderer Bereich schulpädagogischer Arbeit ist so umstritten und bietet so viel Konfliktstoff. Dieses Praxisbuch stellt für die Leistungsbewertung in offenen und geschlossenen Unterrichtsformen klare Kriterien auf und präsentiert praxisorientierte Hilfestellungen und Kopiervorlagen, die den Bewertungsprozess für alle Beteiligten nachvollziehbar und handhabbar machen.

Cornelsen

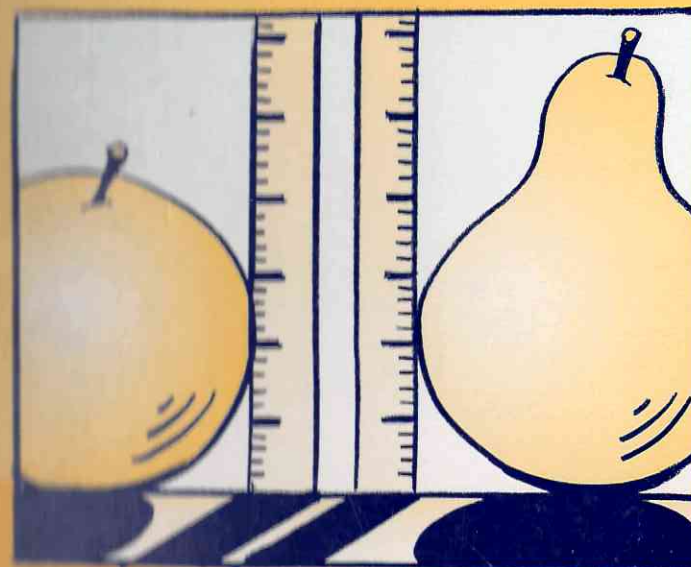
ISBN 978-3-589-22171-4



9 783589 221714

Hebtk

Liane Paradies
Franz Wester
Johannes Greving



f o r h .

Leistungsmessung und -bewertung

Cornelsen

SCRIPTOR

1.2 Dimensionen des Leistungsbegriffs

Unsere modernen (post-)industriellen Gesellschaften sind ohne das Streben nach und das Erbringen von Leistung schlechterdings nicht denkbar – das hat MAX WEBER mit seinem grundlegenden Werk zur protestantischen Ethik bereits vor hundert Jahren nachgewiesen, und der weltgeschichtliche (Globalisierungs-)Prozess zeigt, dass keineswegs kontemplative Sozietäten im Vormarsch sind, sondern die Leistungsgesellschaft überall in der Welt an Boden gewinnt.

Wie aber wird der Leistungsbegriff in unserer Gesellschaft gesehen? Beispiele für Leistungen gibt es jeden Tag in unserer Medienwelt: Leistungen in Sport, Politik, Technik, Kultur, im täglichen Leben, bei abstrusen Höchstleistungen für das Guinness-Buch etc. Man fragt sich oft vergeblich, wem diese Leistungen eigentlich nützen und was sie Tolles bewirken, aber dennoch werden sie durch die Bank positiv konnotiert – etwas leisten ist durch und durch gut und sinnvoll! (Niemand würde dagegen ernsthaft auf die Idee kommen, die Taten eines Mörders oder Bankräubers als Leistung zu bezeichnen!) Dieser Begriff ist also sehr unmittelbar mit unserem gesellschaftlichen Wertesystem verknüpft.

Wir alle werden tagtäglich mit Leistungserwartungen an uns konfrontiert. Doch wie unterscheiden sich die Leistungsanforderungen an bestimmte Gruppen oder auch an Einzelne in der Gesellschaft, wie werden sie definiert, wer definiert sie und welche Leistungen sind gefragt? Für Schüler aller Schulformen und Jahrgänge lässt sich das relativ einfach beantworten: Hier wird der Leistungsbegriff ausschließlich mit Schule verbunden und – noch eine Kategorie enger gefasst – unmittelbar an die schulischen Noten geknüpft.

1.2.1 Leistung als gesellschaftliches Phänomen

Auch außerhalb der Schule werden schulische Leistungen fast unmittelbar an die schulischen Noten geknüpft.

Für die Gesellschaft hat die Schule eine fünffache Funktion:

1. Die *Qualifikationsfunktion*: Noten sollen eindeutig, nachvollziehbar und objektivierbar den Leistungsstand des zu Beurteilenden unter speziellen, genau definierten Aspekten dokumentieren und dem zu Beurteilenden selbst eindeutige Kriterien liefern. Unter diesem Aspekt ist Leistungsbewertung also ein Teil der kritischen Selbsteinschätzung. In der Konse-

quenz für den zukünftigen Unterricht müsste sich ein Schüler also die Frage stellen: Wie komme ich zu einer realistischen Selbsteinschätzung?

2. Die *Selektionsfunktion*: Die Schule ist in unserer freien Leistungsgesellschaft der Ort, an dem für die jungen Menschen die Sozialchancen für das weitere Leben primär vergeben werden – diese manchem Pädagogen schmerzhaft anmutende Bestimmung von Schule darf und kann nicht ignoriert werden. Ohne ein komplexes System von Leistungsmessungen und -bewertungen könnte Schule diesem gesellschaftlichen Auftrag nicht nachkommen. Die Selektionsfunktion generiert in der Praxis die Frage: Wie komme ich zu guten Noten? Sie kann kontraproduktiv wirken im Hinblick auf die Qualität der Selbsteinschätzung.
3. Die *Legitimationsfunktion*: Die beiden oben skizzierten Funktionen kann Schule nur dann erfüllen, wenn sowohl die betroffenen Schüler als auch die „abnehmende“ Gesellschaft die Kriterien und die konkrete Durchführung der individuellen Leistungsbewertung als „gerecht“ empfinden. Nur dann, wenn – ungeachtet aller, vielleicht „ungerechter“ Details – alle die Maßstäbe für die Verteilung der Lebenschancen als angemessen und als fair begreifen, kann Schule ihren gesellschaftlichen Auftrag erfüllen.
4. Die *Informationsfunktion*: Leistungsbeurteilungen geben den auszubildenden Betrieben, den weiterführenden Schulen und den Hochschulen Informationen über Ausbildungs- und Leistungsstände der sich Bewerbenden. Noten haben allerdings einen geringen Informationswert über vorhandene Kompetenzen, vielmehr signalisieren sie den Platz in einer Rangfolge. Das ist ein besonderes Problem für weiterführende Schulen, wenn sie lern- oder leistungsschwache Schüler aufnehmen müssen (Berufsvorbereitungsklassen oder Berufsgrundbildungsklassen). Diese Schüler haben in der Regel Probleme, ihre Kompetenzen zu präsentieren, nicht zuletzt deshalb, weil sie ein diffuses Selbstbild mitbringen.
5. Die *Sozialisierungsfunktion*: Schule vermittelt die Existenz und Gültigkeit von Leistungsnormen. Diese können sich von den gültigen Normen in der Familie, im Freundeskreis etc. erheblich unterscheiden. Andererseits können gleiche Leistungen in unterschiedlichen Bereichen unterschiedlich bewertet werden: Hilfsbereitschaft zählt im Freundeskreis viel, in der Schule geht es (bisher) mehr um den Nachweis individueller Leistungsfähigkeit. Anstrengungsbereitschaft zählt in der Schule mehr als emotionale Zuwendung; gute Leistungen in Mathematik sagen nichts über den Status in der Freundesgruppe aus.

Am ehesten lassen sich Leistungsnormen in der Schule mit denen der Arbeitswelt verbinden. Insofern sozialisiert die Schule nicht nur mit Blick auf die Leistungen allgemein, sondern auch mit Blick auf die Anforderungen in der Arbeitswelt. Das gelingt umso eher, als auch die sogenannten Sekundärtugenden in unterschiedlichen Formen zum Tragen kommen. In Projekten und anderen kooperativen Arrangements lassen sich Sinn und Wirksamkeit erfahren. Deshalb ist es so wichtig, dass z. B. Projekte nicht von der Leistungsorientierung freigemacht bzw. von der Leistungsidee abgekoppelt werden.

1.2.2 Leistung als psychologisches Phänomen

Leistung wird in der Psychologie so definiert: Leistung ist der Vollzug und das Ergebnis von Tätigkeiten in unterschiedlichen Handlungsfeldern.

„Leistung, Gegenstand von Arbeitsprozessen, Eignungsuntersuchungen. ... Innere und äußere Bedingungen der Leistungsbereitschaft werden besonders von Arbeitsprozessen und -Physiologie untersucht. ... Sozialpsychologisch gilt die Leistung als unterscheidendes Merkmal der Einzelnen im Aufbau der Industriegesellschaft (Leistungsgesellschaft). Die Diagnose der individuellen Leistungsfähigkeit gibt zugleich den Ort der Person im sozialen Gefüge an. Die Erziehung und Ausbildung zielen auf Leistungssteigerung in den Grenzen der Veranlagung. Doch ist die Leistung in hohem Grade auch von dem Anspruchsniveau abhängig.“
(HEHLMANN, WILHELM: Wörterbuch der Psychologie. Stuttgart 1965)

Leistung als *psychologisches Phänomen* hat drei unterschiedliche Dimensionen, die individuelle, die soziale und die zielorientierte Dimension:

- Die Selbsteinschätzung der eigenen geistigen Fähigkeiten und Kapazitäten setzt die je persönliche individuelle Norm.
- Die Eingebundenheit in das soziale Miteinander des privaten und familiären Bereichs und die damit verbundenen Standards konstituieren die soziale Norm.
- Die eigene Lebensplanung schließlich bezieht sich auf die zielorientierte Dimension – Berufs- und Karrierewünsche und -ziele, die angestrebte Stellung in der Gesellschaft, die innere Bereitschaft, hierfür etwas zu tun etc.

Auch persönliche Eigenarten und Eigenschaften oder Fähigkeiten beeinflussen die Beurteilungen eines anderen. So konnten wir in den Schulen häufig beobachten, dass Schüler bei schlechten Leistungen die Gründe für

ihr Verhalten eher in den äußeren Umständen begründet sahen, die zu dieser Situation geführt hatten, und nicht in ihrem eigenen Unvermögen. Natürlich gibt es, wie die Forschung gezeigt hat, unterschiedliche Strategien zwischen verschiedenen leistungsstarken Schülern, die eng an das jeweilige Selbstkonzept gebunden sind: Je leistungsschwächer ein Schüler ist, desto größer seine (statistische) Neigung zu destruktiver Selbstkritik. Einige der biografischen Kurzporträts in den letzten Shell-Jugendstudien belegen dies eindrucksvoll (z. B. FISCHER, Opladen 2000, Bd. 2, S. 169ff.).

Der Lehrer dagegen macht das schlechte Abschneiden in erster Linie an den individuellen, persönlichen Merkmalen des Schülers fest, an seinen fehlenden Kenntnissen und am mangelnden Üben. Einige der genannten Kriterien waren: Anstrengungsbereitschaft, Fähigkeiten, Selbstständigkeit, Konzentrationsfähigkeit, Ausgeglichenheit. Die Qualität des eigenen Unterrichts spielte eine ebenso untergeordnete Rolle wie die individuellen fachspezifischen Kenntnisse der Schüler.

Vorkenntnisse und Vorinformationen beeinflussen unsere Sicht über die menschlichen Eigenschaften eines Schülers. Sie bilden praktisch die Basis, mit der wir alle weiteren Informationen verknüpfen und die wir durch aktuelle Erlebnisse erweitern. Wir ordnen die neuen Informationen in Bezug auf die schon bekannten ein und interpretieren sie. So entsteht ein Schülerbild, das wir uns „zurechtgeschustert“ haben, indem wir Informationsdefizite durch eigene plausibel erscheinende Rückschlüsse ergänzt haben (implizite Persönlichkeitstheorien).

Schon 1946 hat ASCH diese Zusammenhänge in Experimenten nachgewiesen. Gesunder Menschenverstand und Menschenkenntnis bilden hier die Grundlage für Schülerbeurteilungen.

In unserer Schulzeit war es üblich, dass der Beruf des Vaters in das Klassenbuch eingetragen wurde. Führen diese Informationen zu einer besseren Bewertung eines Schülers, wenn wir wissen, dass der Vater an der Universität Mathematik unterrichtet oder Arzt ist? Das muss der Fall gewesen sein, denn die sogenannten „familiären Verhältnisse“ sind mittlerweile schon lange aus den Klassenbüchern verschwunden.

Subjektive Theorien (GROEBEN u. a., 1988) über Schüler wie „Fleißige, ordentliche Schülerinnen haben auch Ordnung in ihren Heften und Gedanken“ oder „Schülerinnen sind sprachbegabter als Schüler“ haben ebenfalls Auswirkungen auf die Beurteilungen. Die individuelle Sichtweise des Lehrers (seine subjektive Alltagstheorie) bildet dabei den Maßstab für konkrete Entscheidungen.

1.2.3 Leistung als pädagogisches Phänomen

Im 19. Jahrhundert bildete sich ein Leistungsbegriff in den Schulen aus, der sich an überprüfbarstem Wissen und Können orientierte, damit sollten die Voraussetzungen für die weiterführenden Schulen gesichert sein. Die unterschiedlichen Fächer und ihre fachspezifischen Anforderungen prägen diesen Leistungsbegriff. Mithilfe von Tests oder Klassenarbeiten werden die individuellen Leistungen gemessen und bewertet (zensiert). Die nachweisbaren Ergebnisse – die Produkte – stellen die schulische Leistung dar.

„Die Leistungsbereitschaft und die Leistungsfähigkeit ist entwicklungsabhängig und variiert je nach Individuum und Zuständigkeit. ... Der erzieherische Wert der Leistung liegt einmal in der Willensschulung durch die richtig eingesetzte Kraftentfaltung im Arbeitsprozeß und endlich in dem Erfolgs- und Bündniserlebnis nach Vollendung der Leistung. Die Erziehung zur Leistung gehört zu den Aufgaben der Schul- und Berufserziehung. Die Schule verlangt von jedem Schüler ein Mindestmaß von Lern- und Arbeitsleistung. Sie tritt damit als Leistungsschule in Gegensatz zu allen denjenigen Bestrebungen, die die Schularbeit durch die zufälligen Erlebnisse und das jeweilige Ausdrucksbedürfnis der Schüler bestimmen lassen wollen.“
(HEHLMANN, WILHELM: Wörterbuch der Pädagogik. Stuttgart 1964)

Dagegen wird in der Reformpädagogik Anfang des 20. Jahrhunderts ein Leistungsbegriff in den Vordergrund gerückt, der sich an der individuellen Lern- und Entwicklungsmöglichkeit der Schüler orientiert: Ausgehend von der Chancenungleichheit der Schüler und den fehlenden Kriterien und Maßstäben zur Leistungserfassung und -messung entwickeln z. B. Maria Montessori, Celestin Freinet und Peter Petersen eine ganzheitliche, am jeweiligen Individuum orientierte Sichtweise, die sich mit der Leistungsmessung durch schlichte Benotung nicht verträgt.

Heute versuchen wir, diese beiden Leistungsbegriffe miteinander zu verbinden. Dabei haben sich in der neueren Diskussion über die pädagogischen Dimensionen der Leistungsbewertung einige Eckpunkte herauskristallisiert:

- Leistung ist nicht „frei“ feststellbar; der die Leistung Beurteilende setzt immer Normen und Schwerpunkte.
- Leistung bedarf nicht nur der Fremdbeurteilung (durch den Lehrer), sondern auch der Selbstbeurteilung – nur so können die je eigene Leistungsfähigkeit realistisch eingeschätzt, neue Lernsituationen angemessen antizipiert und damit letztlich selbstständig und selbstbewusst bewältigt werden.

- Leistung ist subjektbezogen, d. h., alle objektivierbaren, testtheoretisch orientierten Verfahrensweisen messen die je spezifisch individuellen Leistungen immer weniger, je mehr sie sich entindividualisierten, objektivierten Kriterien nähern. Ein allgemeiner Beobachtungsbogen, der für alle Schüler der BRD entwickelt würde, müsste notwendigerweise sehr abstrakt sein, wäre zwar scheinbar extrem objektiv, könnte aber die konkrete Individualität jedes einzelnen Schülers nur mit Blick auf die Ausfüllung der Kriterien erfassen. Wollte er die Individualität insgesamt erfassen, müsste er verändert, angepasst, spezifiziert werden, und genau dann ginge die ganze „Objektivität“ wieder verloren.
- Leistung ist wesentlich auch prozessorientiert. Alle traditionellen Formen der Leistungsbewertung haben ausschließlich das Produkt im Auge; die Beurteilung des (Lern-)Prozesses, der zu diesem Produkt geführt hat, bleibt außerhalb des Wahrnehmungshorizontes. Dies ist mittlerweile auch in den Fokus der Bildungspolitiker geraten. Im neuen niedersächsischen Organisationserlass für die Sekundarstufe I aus dem Jahr 2004 (www.nibis.de) findet sich die folgende Formulierung: „Die Leistungsbewertung darf sich nicht in der punktuellen Leistungsmessung erschöpfen, sondern muss den Ablauf des Lernprozesses einbeziehen.“ Die Konsequenzen sind nicht immer klar und eindeutig. Wird das gleiche Ergebnis unterschiedlich oder werden Teilergebnisse getrennt bewertet?

Die *pädagogische Funktion der Leistungsbewertung* unterscheidet sich daher zum Teil erheblich von ihrer gesellschaftlichen Aufgabe und psychologischen Funktion:

Die *Qualifikationsfunktion* erfüllt auch unter dem pädagogischen Aspekt die gleiche Funktion wie oben skizziert. Zeugnisse sind für die Schüler in erster Linie Rückmeldungen über den erreichten Leistungsstand.

Zwei ganz wesentlich komplexere Funktionen, die der Selektionsfunktion in der Praxis oft diametral widersprechen, hat die Leistungsmessung aber unter dem pädagogischen Aspekt: die der *Lern diagnose* und der *Lernberatung*. Dahinter steht das Konzept der individuellen Förderung des Schülers, also die Vorstellung, dass man nicht alle „über einen Kamm scheeren“ kann.

Die *Berichtfunktion* gilt für Schüler, Eltern und Lehrer: Der Schüler erhält Informationen über seine Lernentwicklung, Lernfortschritte, Lernerfolge und Leistungsstände, er wird aber gleichzeitig auch mit seinen Mitschülern verglichen. Der Lehrer erhält Informationen zur Wirksamkeit

seines Unterrichts und über die Lernstände seiner Schüler. Die Eltern werden über Lernstände informiert.

Die *Motivations- und Förderungsfunktion* ist ambivalent: Schüler werden durch Leistungsbeurteilung nicht nur motiviert, sich mit bestimmten Lerninhalten zu beschäftigen, sondern sie können so auch individuell gefördert werden. Andererseits muss eine sehr wichtige Einschränkung gemacht werden: Da Leistungsbeurteilung immer auch eine selektive Wirkung hat, können Schüler dann, wenn sie nicht die gewünschte Leistung erbringen, sehr leicht demotiviert und blockiert werden! Wir sind daher der Meinung, dass es im Lernprozess immer auch beurteilungsfreie (Monitoring-)Phasen geben muss, und werden auf diesen Aspekt ausführlich später zu sprechen kommen.

Als Fazit dieser kurzen Sichtung der unterschiedlichen Dimensionen des Leistungsbegriffes lassen sich zwei Ansprüche formulieren, die ab Seite 40 wieder aufgegriffen, präzisiert und reflektiert werden sollen:

- Der eng gefasste, auf das Endergebnis bezogene und mit testtheoretischen (Schein-)Vorgaben operierende Leistungsbegriff muss um wesentliche Dimensionen erweitert werden.
- Der Prozess der Leistungsbewertung muss getragen werden von klaren Vereinbarungen zwischen Lehrer und Schülern sowie permanenter Reflexion der Leistungsnormen und der Kommunikation darüber.

1.2.4 Konsequenzen für die Praxis der schulischen Leistungsbewertung

Welche Fragen müssen Pädagogen beantworten, wenn sie als Lehrer die Lernfortschritte, Lernergebnisse und Lernfähigkeiten ihrer Schüler messen?

	Förderfunktion Anreize geben, zur Persönlichkeitsstärkung beitragen, zur Selbstdisziplin befähigen	Kontrollfunktion Arbeitsverhalten und Arbeitsergebnisse kontrollieren	Berichtsfunktion Rückmeldungen an Schüler und Eltern	Selektionsfunktion Platzierung, Schullaufbahneempfehlungen	Prognosefunktion Einschätzung inner- und außerschulischen Verhaltens
Beschreibung	Welche Fortschritte gibt es? Welche Kenntnisse, Fertigkeiten und Fähigkeiten werden sicher beherrscht?	Welche Unterrichtsziele sind wie weit erreicht?	Welche Lernentwicklung ist zu beobachten, woran muss der Schüler verstärkt arbeiten?	Über welche, für die Schulform besonders wichtige Fähigkeiten verfügt der Schüler?	Welche Einstellungen zur Gemeinschaft sind feststellbar, welche Aktivitäten seitens des Schülers gibt es?
Bewertung	Welche Anerkennung, welche Kritik für Verhalten und Leistung ist angebracht?	Gehen unbefriedigende Lernergebnisse auf Fehler des Lehrers, unzureichende Arbeitsleistung oder ungenügende Lernvoraussetzungen zurück?	Inwieweit sind die verlangten Wissens- und Könnenselemente fest verankert und jederzeit verfügbar?	Worauf gründet sich die Erwartung einer erfolgreichen Mitarbeit in der nächsten Klasse, Stufe oder Schulform?	Welcher Grad an Selbstständigkeit, welches Maß an Interesse und Lernenergie waren feststellbar?
Einstufung	In welchen Punkten ist die erwartete Leistung übertroffen, voll erreicht, teilweise erreicht, nicht erreicht?	Wie können die Kenntnisse, Fähigkeiten und Fertigkeiten eingestuft werden im Hinblick auf die Lerngruppe?	Welches (Zwischen-)Zeugnis kann ausgestellt werden, ist der Eintritt in die nächste Klasse oder Schulform zu empfehlen?	Ist das Klassenziel in allen Fällen erreicht oder müssen Ausgleichsleistungen herangezogen werden?	Wo sind besonders aussichtsreiche Merkmale der Persönlichkeitsbildung zu berücksichtigen?
Auswertung	Wie und wodurch kann dem Schüler Anerkennung vermittelt, ein Ansporn gegeben werden?	Wo können oder müssen spezielle Fördermaßnahmen ergriffen werden?	Wo ist der Schüler in Gefahr, den Anschluss zu verlieren?	Welche Wahlangebote sollte der Schüler wahrnehmen, zu welcher Wahl ist anzuraten?	Welches Begaubungs- und Fähigkeitsprofil deutet sich an?

1.3 Bezugsnormen der Leistungsbeurteilung

Quer zu der in den letzten drei Unterkapiteln entwickelten Systematik der verschiedenen Inhaltsebenen liegt die Frage nach den Bezugsnormen jeder Leistungsbeurteilung, die vor einem resümierenden Fazit noch zu klären ist:

Leistung per se gibt es nicht. Leistungen müssen definiert werden und lassen sich je nach Situation immer nur in Bezug auf eine Norm bestimmen. Unterschiedliche Maßstäbe prägen das Urteil, ob etwas schlecht oder gut ist. Folgende drei Bezugsnormen bilden die möglichen normativen Grundlagen für die Bewertung von Schülerleistungen:

1.3.1 Individualnorm – individuelle Bezugsnorm

Für und über einen bestimmten Zeitraum wird der individuelle Lernzuwachs eines Schülers erfasst. Individuelle Leistungen werden bewertet und geben dem Schüler Rückmeldung über seinen persönlichen Lernstand. So kann der Schüler die Entwicklung seines eigenen Lernfortschritts erfassen. In der Regel erfolgen diese Beurteilungen in Form von mündlichen Rückmeldungen oder durch Lernentwicklungsberichte (Grundschule, IGS, Waldorfschule ...). Der Schüler wird nicht mit anderen verglichen und kann optimal in seinem Lernverhalten bestätigt oder individuell gefördert werden. Fehler bieten eine Chance zum Lernen und zur Entwicklung neuer Lernstrategien, allerdings nur unter der Voraussetzung, dass Prozesse dokumentiert und reflektiert werden. Das Portfolio und das Lerntagebuch sind dazu geeignete Instrumente.

1.3.2 Sachnorm – kriteriumsorientierte Bezugsnorm

Die Beurteilungen der Lernleistungen nach Fach- und Sachaspekten sind gültiges Schulrecht. Die Lernziele der Unterrichtsfächer bilden die Grundlage der Zensierung. Entsprechend den Notenstufendefinitionen hat ein Schüler die Anforderungen erfüllt oder auch nicht. Die Noten beziehen sich (entsprechend den RRL und Schulcurricula) auf das Erreichen der fachspezifischen Lernziele. Soziale Lernziele werden bei der Zensierung nicht berücksichtigt.

Die Erweiterung der Lern- oder Kompetenzfelder hat den Begriff „Sachnorm“ veralten lassen. Kriterien und Normen können auch aus so-

zialen oder methodischen Lernfeldern definiert werden. Für die Messung und Bewertung gibt es noch wenig Routine. Probleme bereiten insbesondere die klare Beschreibung der Leistung und deren Erfassung. Hier ist Entwicklungsbedarf und Unterstützung nötig.

Die Zensierung sollte unabhängig von Gruppenleistungen erfolgen. Sie gibt Rückmeldung darüber, ob ein Schüler seine bisherigen Lernanstrengungen noch verbessern sollte oder nicht. Hier stellt sich die Frage, ob das wirklich mithilfe einer Notenskala angemessen gelingt. Vermutlich ist der Lernbericht die bessere Variante. Er generiert für die alltägliche Unterrichtspraxis die folgenden Fragen:

- Was soll geleistet werden?
- Warum soll es geleistet werden?
- Wie soll es geleistet werden?
- Wer muss was leisten?
- Welche Bezugsnorm gilt für die Bewertung?

Nur die erste und die letzte Frage besitzen Relevanz für die Notengebung – anders als beim Lernbericht. So wird deutlich, dass die Form der Beurteilung auf den Unterricht zurückwirkt.

Bisher galt als Stärke der Lehrer die Orientierung an der Sachnorm. PISA hat gezeigt, dass gleiche Leistungen sehr unterschiedlich bewertet werden, selbst wenn man sie in der gleichen Schule misst. Die Fragen sind: Was führt dazu? Gehen andere Normen in die Bewertung ein oder sind die Messungen ungenau?

1.3.3 Sozialnorm – kollektive Bezugsnorm

VIERLINGER (1999, S. 80) hat den Vorschlag gemacht, die soziale Norm als kollektive Norm zu bezeichnen, damit der positiv besetzte Ausdruck „sozial“ von vornherein falsche Assoziationen vermeidet. Die schulischen Lernleistungen eines Einzelnen orientieren sich an den Leistungen einer Referenzgruppe. Sie werden innerhalb dieser Gruppe unter- und miteinander verglichen und bewertet. Im Regelfall bildet die eigene Klasse die Vergleichsgruppe bei der Notengebung. Die Schüler werden entsprechend ihrer erreichten Punktzahlen bei der Überprüfung in einer Rangreihe geordnet. Die Zuordnung von Noten zu einzelnen Punkten erfolgt dann in Anlehnung an die Gauß'sche Normalverteilung.

Viele Schüler, Eltern und Lehrer empfinden diese Art der Beurteilung als gerecht – die meisten Schüler befinden sich im Mittelbereich –, auch wenn

z. B. der Lernfortschritt schwächerer Schüler so nicht angemessen bewertet werden kann. Vorgegebene Standards und einheitliche Anforderungen werden nur teilweise berücksichtigt, da kein Vergleich mit anderen Lerngruppen erfolgt. Die Form der komparativen Leistungsbewertung steht allerdings zumindest teilweise im Widerspruch zu den schulischen Normen und Gesetzen – so verbieten es z. B. die niedersächsischen Durchführungsbestimmungen zur Abiturprüfung explizit, Leistungen von Schülern miteinander zu vergleichen (auch wenn zwei oder drei Schüler direkt nacheinander eine mündliche Prüfung zum selben Thema ableisten). Es gibt aber noch einen schwerwiegenden Grund gegen diese Form der Bewertung: Eine Schulklasse stellt nichts weiter als eine kleine, zufällig zusammengewürfelte Testgruppe dar, deren Zusammensetzung nicht für standardisierte Tests geeignet ist. Erst bei extremen Abweichungen von der Normalverteilung – zu viele Gute oder zu viele Schlechte – wird diese hinterfragt.

Alle Bezugsnormen haben ihre Berechtigung. Was liegt näher, als sie alle zu berücksichtigen? Aber was so einfach klingt, lässt sich in der Praxis nur schwer umsetzen. Wir plädieren dafür, zwischen Bewertungen, die sich an der Sozialnorm, und Bewertungen, die sich an der Individualnorm orientieren, deutlicher zu trennen – sowohl was die Messung als auch was die Form der Beurteilung angeht.

1.4 Fehlerquellen bei der Leistungsbeurteilung

Jede Art von Wahrnehmung, von Bewusstwerdung und von Beurteilung hängt von vielen subjektiven Faktoren ab bzw. wird von den unbewussten Voreinstellungen und Erwartungen des Beobachtenden beeinflusst. Im privaten Bereich sind diese Verhaltensweisen sicherlich für viele Gesprächssituationen eher förderlich, verbessern sie doch die Spontaneität der Kommunikation untereinander. Bei der Leistungsbeurteilung sind gerade diese Merkmale eher von Nachteil. Sie haben wenig mit Objektivität oder Transparenz zu tun, werden der Leistungsbeurteilung nur in Teilbereichen gerecht und behindern die Vergleichbarkeit der Beurteilungen. GISLINDE BOVET und VOLKER HUWENDIEK haben in ihrem „Leitfaden Schulpraxis“ (1998) aufgezeigt, welche subjektiven Fehlerquellen bei der Leistungsbeurteilung in der Schule auftreten können.

1.4.1 Der Einfluss von Vor- und Zusatzinformationen

Positive oder negative Zusatzinformationen – auch über außerschulische Bedingungen – beeinflussen die Beurteilung von schriftlichen und mündlichen Leistungen (WEISS 1971 b), was Auswirkungen auf die alltägliche Benotungspraxis hat. Ein unauffälliger, stiller Schüler, über den der Lehrer sonst wenig weiß, hat bei der Korrektur einer Arbeit wahrscheinlich deutlich „schlechtere Karten“ als ein sprachgewandter, in seiner Freizeit im Orchester spielender Schüler, dessen Schwester gerade einen Sportwettbewerb gewonnen hat, wie in der Zeitung zu lesen war.

1.4.2 Der Einfluss von Sympathie und Geschlecht

Objektiv gleiche Leistungen von Mädchen werden günstiger benotet als die von Jungen – und zwar von Lehrern wie von Lehrerinnen! In Befragungen geben beide Gruppen an, Mädchen im Vergleich zu Jungen als fleißiger, angepasster, ordentlicher usw. wahrzunehmen. Untersuchungen haben gezeigt, dass zahlreiche Lehrer diejenigen Schüler günstiger beurteilen, die ihnen sympathisch sind (HADLEY 1971), aber es gab auch viele Lehrer, die in dieser Hinsicht völlig immun waren.

1.4.3 Der Einfluss von subjektiven Theorien

Die subjektiven, berufsbezogenen Theorien eines Lehrers leiten in hohem Maße sein Handeln; ausgeprägte Grundüberzeugungen beeinflussen die Wahrnehmung und Einschätzung von Schülerleistungen. Denn die allgemeine Tatsache, dass man vornehmlich das wahrnimmt, was man wahrzunehmen erwartet, führt häufig zu Beobachtungsverzerrungen und -einseitigkeiten.

1.4.4 Halo-Effekt und logischer Fehler

Allgemeineindrücke oder hervorstechende Merkmale können die Wahrnehmung einzelner, nicht direkt beobachtbarer Merkmale bestimmen. Höfliches Auftreten, ordentliche Kleidung, saubere Heftführung und angemessenes Sprachverhalten führen zu einem gesamtordentlichen Eindruck eines Schülers. Es besteht aber die Gefahr, dass er bessere Leistungsbeurteilungen erhält, als er es eigentlich verdient hätte. Ähnlich wirkt mitunter auch der Ruf, der einem Schüler vorausleitet.

Voreilige Schlussfolgerungen werden von einem Leistungsmerkmal auf ein anderes geschlossen, das als quasi logisch mit dem beobachteten verbunden angenommen wird (logischer Fehler). Erbringt z. B. ein Schüler vorzügliche Leistungen in Mathematik, wird leicht angenommen, dass seine Leistungen auch im Fach Physik sehr gut sein müssen. Ähnlich wird nicht selten von einer vorzüglichen Gedächtnisleistung auf ein entsprechend hohes Maß an Verständnis geschlossen.

1.4.5 Stabile Urteilstendenzen

- a) Gute und sehr gute Noten werden selten benutzt. Es besteht die Neigung, kleinere Mängel relativ stark zu gewichten und vorwiegend negative Urteile abzugeben und schlechte Noten zu erteilen. (Strengfehler)
- b) Es werden hauptsächlich gute Noten und günstige Beurteilungen vergeben. (Mildefehler)
- c) Die Scheu vor extremen Urteilen führt zur Häufung von mittleren Urteilen und durchschnittlichen Noten. (Tendenz zur Mitte)
- d) Es kommt relativ selten zu mittleren Urteilen und durchschnittlichen Noten, sondern häufig zu überzogenen extremen Urteilen. (Tendenz zu Extremurteilen)

1.4.6 Reihenfolgeeffekte

Reihenfolgen- und Positionsfehler ergeben sich aufgrund vorangehender Urteile oder wenn mehrere Beurteilungen nacheinander durchgeführt werden. Eine durchschnittliche Leistung wird beispielsweise oft besser beurteilt, wenn direkt vorher eine sehr mäßige zu bewerten war; eine Leistung wird leicht schlechter beurteilt, wenn ihr eine besonders gute voranging.

1.5 Gerechtigkeit und Chancengleichheit

1.5.1 Das Rasenmäherprinzip

Wenn Chancengleichheit so aufgefasst wird wie in der Karikatur, führt sie sich selbst ad absurdum. Aber: Kann es und darf es sein, dass Lehrer unterschiedliche Bewertungsmaßstäbe anlegen? Dürfen Lehrer sozial



benachteiligte Schüler, die z. B. kein eigenes Zimmer haben oder jeden Morgen Zeitungen austragen müssen, über die normalen pädagogischen Fördermaßnahmen hinaus (notfalls bis zum kostenlos erteilten Nachhilfeunterricht) kompensatorisch dergestalt unter die Arme greifen, dass sie für gleiche Leistungen verschiedene Noten geben? Auf den ersten Blick ist sicher jeder geneigt, diese Frage mit „nein“ zu beantworten, denn so ein Vorgehen erscheint gerade deswegen als ungerecht und auch Unrecht, weil es verschiedene Maßstäbe anlegt, mit zweierlei Maß misst. Dürfen Lehrer also niemals die individuellen Umstände einer Einzelsituation wertend ins Auge fassen? Fragen wir die Experten von der „Ständigen Konferenz der Kultusminister“ (KMK):

„Die Kultusministerkonferenz sieht es als zentrale Aufgabe an, die Qualität schulischer Bildung, die Vergleichbarkeit schulischer Abschlüsse sowie die Durchlässigkeit des Bildungssystems zu sichern. Bildungsstandards sind hierbei von besonderer Bedeutung. Sie sind Bestandteile eines umfassenden Systems der Qualitätssicherung ... Bildungsstandards beschreiben erwartete Lernergebnisse ..., die Schülerinnen und Schüler bis zu einer bestimmten Jahrgangsstufe an zentralen Inhalten erworben haben sollen. Bildungsstandards formulieren fachliche und fachübergreifende Basisqualifikationen, die für die weitere schulische und berufliche Ausbildung von Bedeutung sind und die anschlussfähiges Lernen ermöglichen. Die Standards stehen im Einklang mit dem Auftrag der schulischen Bildung. Sie zielt auf Persönlichkeitsentwicklung und Weltorientierung, die sich aus der Begegnung mit zentralen Gegenständen unserer Kultur ergeben ... (Bildungsstandards) beschreiben die fachbezogenen Kompetenzen einschließlich zugrunde liegender Wissensbestände, die Schülerinnen und Schüler bis zu einem bestimmten Zeitpunkt ihres Bildungsganges erreicht haben sollen. ...

Die Zuordnung zu Anforderungsbereichen orientiert sich an den Einheitlichen Prüfungsanforderungen in der Abiturprüfung (EPA). ...“

Quelle: Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss (Jahrgangsstufe 10) (Beschluss der Kultusministerkonferenz vom 04.12.2003)

So weit, so gut – immerhin nennt die KMK Ziele wie „Persönlichkeitsentwicklung und Weltorientierung“, und da kann doch vom Rasenmäherprinzip nicht ernsthaft gesprochen werden!

Wenn man allerdings genauer hinsieht, wird klar, dass die eigentliche Intention dieses Beschlusses über Bildungsstandards in der Implementierung von möglichst bundesweit (oder zumindest landesweit) einheitlichen und zentralen Leistungsüberprüfungen für den mittleren Bildungsabschluss besteht – es geht also keineswegs nur um die Aussortierung einer (gymnasialen) Elite! Mit diesem Beschluss, der dem allgemeinen Post-PISA-Trend im vom Bildungskatzenjammer befallenen Deutschland folgt, wird die vom Grundgesetz und den Länderschulgesetzen geforderte und geförderte (?) Chancengleichheit reduziert auf den Abgleich punktueller kognitiver Leistungen. Der Kern des KMK-Beschlusses ist – ungeachtet aller verbalen Euphemismen – die Prüfung, und die funktioniert nur dann „gerecht“, wenn alle mit dem gleichen Kamm gekämmt, also formal gleich behandelt werden – genau das versteht man unter dem Rasenmäherprinzip.

In anerkannter Offenheit wird dieses Ziel in den Anfang 2004 erschienenen „curricularen Vorgaben für die Schuljahrgänge 5/6“ des Niedersächsischen Kultusministeriums formuliert:

„Die verbindlichen Vorgaben zielen auf die Möglichkeit landesweit einheitlicher Überprüfung. Entscheidend ist nicht die prozessuale Abfolge vorgegebener Lernziele und Inhalte ... , sondern der tatsächliche Kompetenzerwerb ...“

Quelle: Nds. Kultusministerium, Curriculare Vorgaben für das Gymnasium Schuljahrgänge 5/6, Hannover 2004, S. 7

1.5.2 Das Prinzip der individuellen Förderung

Gegen dieses Prinzip der formal gerechten Gleichbehandlung gibt es eine Reihe von gewichtigen Argumenten:

Die Objektivität von Prüfungen stellt notwendigerweise eine Reduktion und Verzerrung dar, weil sowohl in den Prozess der Leistungserbringung als auch in den der Beurteilung viele psychologisch-subjektive Faktoren hineinspielen, die ausgeblendet werden müssen, um Ergebnisse nicht zu verfälschen.

Die Vorstellung, dass dann, wenn Kollege A und Kollege B den gleichen Stoff im Unterricht behandeln, der Lernprozess und der Lernerfolg bei allen Schülern auch tatsächlich gleich sei, ist naiv: Zu unterschiedlich sind die Lehr- und Lernpersonen, zu unterschiedlich ist das soziale und das Lernklima, zu unterschiedlich sind letztlich auch die äußeren Bedingungen.

Anstrengungsbereitschaft und Mobilisierung unserer Lernenergien wie unserer Interessen sind von den positiven Rückmeldungen derer abhängig, auf deren Urteil wir etwas geben. (Selbst in Leistungskursen der gymnasialen Oberstufe haben wir die Erfahrung gemacht, dass Schüler „für den Lehrer“ lernen!)

Das Lernpotenzial der Bezugsgruppe und ihr Verhaltensrepertoire entscheiden mit über die Lernmöglichkeiten und das Leistungsniveau jedes Individuums.

Die Vorstellung, Lehrer hätten wie Richter zu sein, um so objektiv, unparteilich und gerecht zu entscheiden, ist nicht haltbar.

„Wie unterscheidet sich die Gerechtigkeit des Pädagogen von der Gerechtigkeit des Richters?“

Die Klassenarbeit und das Zeugnis sind primär Instrumente der Förderung, denn wir sind Pädagogen, nicht Richter. Der Richter sieht mehr in die Vergangenheit und stellt fest, was gewesen ist; als Pädagoge sehe ich mehr in die Zukunft und bedenke, was werden kann. Er muss ein Verhalten als einen Fall z. B. von Fahrlässigkeit oder von Vorsatz erkennen, einordnen, abschließend bewerten und ein entsprechendes Strafmaß festsetzen. Dabei muss er auf Gleichbehandlung aller vor dem Gesetz bedacht sein. Als Pädagoge kann ich die Einmaligkeit und die Kostbarkeit, natürlich auch die Schwächen der mir anvertrauten Schülerin oder des mir anvertrauten Schülers viel mehr in das Zentrum meiner Aufmerksamkeit rücken; und meine erste Aufgabe – ohne Gesetzbuch im Hintergrund – ist, die Entwicklung dieses Individuums zu fördern, ihre oder seine Möglichkeiten zu entfalten. Die Gerechtigkeit des Pädagogen besteht also nicht so sehr in der Zuteilung genau der gleichen Anforderungen an alle wie in der unvoreingenommenen, genauen Beobachtung jedes einzelnen, ihrer oder seiner individuellen Voraussetzungen und Möglichkeiten, in der gleichen Zuwendung und im Finden der ihr oder ihm zuträglichen Lernaufgaben. Wenn schon die Bewertung von Schülerleistungen immer im Spannungsfeld steht zwischen den beiden Maximen ‚allen das gleiche‘ und ‚jedem das Seine‘, dann muss uns ... im Zweifel das Prinzip ‚jedem das Seine‘ wichtiger sein.“

(DIETRICH ALBRECHT 1989)

1.6 Notwendigkeit eines weiter gefassten Lern- und Leistungsbegriffs

Was kann eine neue Kultur der Leistungsbeurteilung nicht leisten?

Zwei wesentliche Aspekte sollten zumindest genannt werden:

1. Die Widersprüchlichkeit zwischen den Polen Selektion und Förderung, die die Schule strukturell kennzeichnet, lässt sich über die Leistungsbeurteilung nicht aufheben.
2. Intersubjektivität garantiert nicht absolute Gerechtigkeit.

Die Entwicklung muss auf die Interaktion und deren qualitative Umsetzung setzen. In den schülerorientierten Formen der Leistungsbeurteilung sind Konflikte nicht zu vermeiden. Sollten sie auch nicht, denn sie zeigen

Unterschiede an, die sonst oft nicht wahrgenommen würden. So erhöht sich die Komplexität einer nicht immer bis ins Letzte planbaren Situation, in der sich Lehrer wie Schüler bewähren können, für die aber größtenteils noch Routinen geschaffen werden müssen, um unnötige Aufregungen vermeiden zu können. Diese Routinen können sich mit Blick auf Vorbereitung, Durchführung und Reflexion der Interaktion herausbilden.

Die im Folgenden skizzierten Überlegungen haben große Nähe zu Evaluationsverfahren, die ein wesentliches Element von Schulentwicklung sind. Selbstevaluation hat darin einen zentralen Stellenwert. Die Erfahrungen damit auf die Leistungsbeurteilung zu übertragen, kann diese voranbringen. Die bisherige Bewertungspraxis in der Schule zeigt eindeutig, dass in erster Linie die Konzentration auf dem inhaltlich-fachlichen Lernbereich liegt – und das ist kein Zufall, denn dieser kann mit den traditionellen Bewertungsformen wie Tests, Klassenarbeiten und Klausuren, Referaten, Beteiligungsnoten etc. beurteilt werden. Die anderen Lernbereiche aber sind mit diesen Bewertungsmodalitäten nicht oder nur unzureichend erfassbar. Daher benötigen Schulen neue Formen der Leistungsbewertung, die über den sachlich-inhaltlichen Lernbereich hinausgehen und sowohl den methodischen als auch den sozialen und den persönlichen Lernbereich erfassen.

Hinter dieser Forderung steht die inzwischen auch auf der Ebene der Kultusbürokratien nicht mehr angezweifelte Erkenntnis, dass Schule in Zukunft mit einem weiter gefassten Lern- und damit auch Leistungsbegriff operieren muss, dessen Dimensionen in der folgenden Übersicht skizziert werden:

1.6.1 Die vier Dimensionen des Lern- und Leistungsbegriffes

Inhaltlich-fachlicher Lernbereich	Methodisch-strategischer Lernbereich	Sozial-kommunikativer Lernbereich	Persönlicher Lernbereich
wissen, kennen, beherrschen, anwenden können	aus Materialien Informationen entnehmen, exzerpieren, strukturieren, ordnen	zuhören, argumentieren, fragen, kooperieren	ein realistisches Eigenbild entwickeln und Selbstvertrauen gewinnen
verstehen, übertragen, erschließen, sich selbstständig auseinandersetzen, ordnen, übertragen, transferieren	Lern- und Arbeitsprozesse planen, organisieren, gestalten, Arbeitsdisziplin wahren, Ordnung halten	sich in andere einfühlen, Signale wahrnehmen, integrieren, Konflikte lösen	die Fähigkeit zum Engagement entwickeln, (Selbst-)Kritikfähigkeit aufbauen
urteilen, begründen, reflektieren, problematisieren, erörtern	Entscheidungen treffen	Ergebnisse oder Prozesse präsentieren, Diskussionen und Gespräche leiten	Werthaltungen entwickeln

Diese vier Dimensionen des Leistungsbegriffes sind zunächst einmal als analytische Kategorien zu verstehen, d. h., sie lassen sich nicht einfach während des Unterrichts beobachten und zu einem Gesamtbild addieren, sondern verstehen sich in der Praxis als organische Einheit, die sinnvoll nur an den Kompetenzbegriff zu binden ist. Leistungen lassen sich am ehesten durch die Aufgaben voneinander unterscheiden, die ein Schüler zu lösen imstande ist.

Wenn Leistungsbewertung sich an einem erweiterten Leistungsbegriff orientieren will, müssen methodisch-strategische Kompetenzen in den Aufgaben explizit gefordert werden und in der Beschreibung der Bewertungskategorien enthalten sein. Explizit fordern bedeutet, dass das Vorgehen der Schüler in der Bearbeitung einer Aufgabe als Ergebnis einer Entscheidung erkennbar gemacht und mit Blick auf die Effektivität vom Schüler bewert-

tet wird. Das Abwägen von Alternativen in der Planung oder Reflexion beweist ein hohes methodisches Kompetenzniveau. Sie voneinander abzugrenzen führt zu einer differenzierteren Bewertung bzw. Rückmeldung an den Schüler und zu mehr Transparenz. HILBERT MEYER (KIPER u. a. 2003) hat dafür ein Modell entwickelt, das fünf Kompetenzstufen nach dem Grad der Selbstständigkeit und der Reflektiertheit unterscheidet.

1.7 Notwendigkeit neu gefasster Kriterien für Beurteilungen

1.7.1 Grundlegende Ansprüche an die Leistungsbewertung

Objektivität (viele Wissenschaftler sprechen heute lieber von „Intersubjektivität“) bedeutet, dass verschiedene Bewerter unabhängig voneinander zum exakt gleichen Ergebnis kommen. Was für wissenschaftliche Experimente aber wünschenswert und notwendig sein mag, ist für schulische Bewertung nicht so ohne weiteres postulierbar!

„Objektivität (dies hieße: Alle Lehrkräfte stimmen bei der Beurteilung einer Arbeit oder eines Schülermerkmals perfekt miteinander überein) ist nicht automatisch gleichbedeutend mit Korrektheit und sachlicher Richtigkeit; es gibt auch die Möglichkeit kollektiven Irrtums, z. B. wenn bei der Beurteilung der Intelligenz eines Schülers das Merkmal ‚Intelligenz‘ mit ‚Kreativität‘ verwechselt wird.“ (HELMKE 2003, S. 87) Ziel schulischer Leistungsbewertung kann nur sein, die unkontrollierte Subjektivität zugunsten einer kontrollierten und transparenten Leistungsbewertung möglichst weit zurückzudrängen. Dies kann geschehen durch eine möglichst präzise Formulierung von Kriterien, durch eindeutige Binnengewichtungen und rechtzeitig vorher diskutierte Musterlösungen oder Erwartungshorizonte – auch und gerade bei offenen Aufgabenstellungen wie bei Deutschaufsätzen.

Reliabilität bedeutet Zuverlässigkeit. Die angewendeten Beurteilungskriterien und -verfahren müssen die zu beurteilenden Merkmale auch wirklich korrekt messen – und nicht etwas ganz anderes (vgl. den Abschnitt über Fehlerquellen, siehe S. 34 ff.). Die Reliabilitätskontrolle, die die Wissenschaft vorschlägt, ist leider im schulischen Alltag recht aufwändig bzw. kaum zu realisieren: „Reliabel ist ein Urteil dann, wenn es sich – vorausgesetzt, das zu beurteilende Schülermerkmal bzw. die Schülerleistung ist

im Zeitverlauf stabil geblieben – bei wiederholten Beurteilungen nicht ändert. Probe aufs Exempel: wiederholte Korrektur und Bewertung der gleichen Serie von Arbeiten (Aufsätze, Klassenarbeiten, Hausarbeiten) einige Monate später.“ (HELMKE 2003, S. 87)

Dennoch muss als ernsthafter Anspruch an jeden einzelnen Leistungsmessungsprozess die Frage gestellt werden, ob die gewählten Verfahren sachangemessen sind und wirklich das messen, was gemessen werden soll. Dabei gilt der Grundsatz, dass die Anzahl der unabhängig voneinander gestellten Einzelaufgaben zu einem bestimmten Lernziel oder -bereich die Reliabilität erhöht. HANS-GERT WENGERT gibt ein gutes Beispiel: „Wenn ein Deutschlehrer überprüfen will, ob seine Schüler verlässlich zwischen ‚das‘ und ‚dass‘ unterscheiden können, ist ein herkömmliches Diktat, in dem diese Unterscheidung nur drei- oder viermal getroffen werden muss, weniger reliabel als ein Lückentext, in dem die Schüler an 15–20 Stellen ihre Entscheidung eintragen müssen.“ (WENGERT, in: BOVET/HUWENDIEK 1998, S. 282)

Validität bedeutet Gültigkeit: Ein Untersuchungsverfahren ist dann gültig, wenn es den zu messenden Gegenstand exakt misst und nichts anderes und sich das Urteil auch tatsächlich auf die Leistung bezieht, die gemessen werden sollte. Wenn ein Lehrer z. B. seine Klassenarbeiten so mit Stoff überlädt, dass ein guter Teil der Lerngruppe es einfach nicht schafft, alle Aufgaben in der vorgegebenen Zeit zu bearbeiten, wird ein wesentlicher Bestandteil der zu erbringenden Leistung völlig fachfremd sein – es geht dann nämlich nicht mehr nur um z. B. das Lösen von quadratischen Gleichungen, sondern genauso um Schnelligkeit! Das eigentlich beabsichtigte Testziel – die Überprüfung des Lernstandes der Schüler – wird so verfälscht. In den folgenden Abschnitten werden wir die Möglichkeiten erläutern, Validität durch gemeinsame Absprachen und „Arbeitsbündnisse“ innerhalb der Schule herzustellen bzw. zu verbessern.

1.7.2 Notwendigkeit neu gefasster Gütekriterien für Beurteilungen

Unterrichtsverfahren, -methoden und -arrangements, die prozessorientiert die drei „neuen“ Aspekte (methodisch-strategisch, sozial-kommunikativ, persönlich) der Leistungsbeurteilung integrieren wollen, sind per se (ergebnis-)offener und flexibler als traditionelle Lehr-/Lern-Arrangements. Dies bedeutet aber auch zwingend:

- Die Normen und Standards der Leistungsbewertung dürfen nicht a priori festgesetzt werden, sondern können (bzw. müssen) sich mit dem Lern- und Arbeitsprozess ändern, wenn sie sich im Laufe des Beurteilungsverfahrens als nicht brauchbar erweisen. Je offener die gewählte Unterrichtsform, desto gewichtiger diese Flexibilität, z. B. im projektorientierten Unterricht.
- Diese Normen und Standards müssen selber immer kritisch beleuchtet und reflektiert werden, also selbst Gegenstand des Unterrichts und – last, but not least – schulischer Absprachen, Konferenzbeschlüsse etc. werden. Nicht nur der Lehrer bzw. die Schule gibt die Maßstäbe des Beurteilungsverfahrens vor, sondern diese werden in einem gemeinsamen, dialogischen Prozess immer wieder überarbeitet, *kommunikativ validiert*. Widerspruch und Kritik der Schüler sind also nicht als Störquelle, sondern im Gegenteil als Möglichkeit und Chance zu verstehen, zu gemeinsamen Kriterien zu gelangen. Das kostet Zeit, denn die konkrete unterrichtliche Sach- oder Handlungsebene muss immer wieder verlassen werden, um darüber reflektieren zu können. Insbesondere dann, wenn aus der Reflexion erkennbar Konsequenzen für die folgenden Handlungsphasen gezogen werden, wird dieser Zeitverlust mehr als wieder aufgewogen.

Es gibt einen Gewinn an *Transparenz* der Schüler ihren eigenen Leistungen gegenüber, aber auch der Lehrer im Kontakt zu Eltern, Kollegen und der Aufsichtsbehörde.

1.7.3 Kommunikative Validierung in der Schule

Beide Schlussfolgerungen, die wir eben skizziert haben, können auch von einer einzelnen Lehrkraft gezogen und in die Praxis umgesetzt werden, wozu wir an dieser Stelle alle Kollegen ausdrücklich ermuntern möchten! Aber einfacher geht es doch immer im Team – und das muss ja nicht gleich ein Gesamtkonferenzbeschluss der Schule sein! Absprachen und Aktivitäten können auf unterschiedlichen Ebenen stattfinden: gemeinsam im Fach, in einer Jahrgangsstufe, zwei oder mehr Parallelklassen in einem oder mehreren Fächern etc. Je mehr eine Schule sich dem je schulspezifischen Schulentwicklungsprogramm verschreibt, desto breiteren Raum werden diese Formen von kommunikativer Validierung auch der Leistungsbeurteilung einnehmen (müssen), und desto entlasteter (auch gegenüber Eltern und Schulaufsicht) kann der einzelne Kollege diese neuen Formen der Bewertung praktizieren.

1.7.4 Kommunikative Validierung zwischen Schülern und Lehrer

Die in den oben angesprochenen Metaphasen zu treffenden Vereinbarungen sind selbstredend nicht das Produkt von völlig gleichberechtigten Partnern. Den letztendlichen Zwang zur Notengebung und damit auch zur Zuweisung von Sozialchancen kann auch die kommunikative Validierung von Leistungsnormen nicht abschaffen. Diese Metaphasen sind also nicht durch die symmetrische Kommunikation gleichberechtigter Partner gekennzeichnet. Dennoch gibt es eindeutig benennbare *Gütekriterien für die Reflexionsphasen*. HANS-ULRICH GRUNDER und THORSTEN BOHL (2001, S. 46) nennen fünf:

1. Information und Diskussion des konkreten Beurteilungsverfahrens

In der Regel werden Lehrer ihre Überlegungen hinsichtlich Verfahren und Kriterien der Beurteilung vorstellen. Diese Überlegungen, die nun vor einer konkreten Anwendung stehen, werden detailliert erläutert und gemeinsam diskutiert.

2. Offenheit für Änderungen

Änderungsvorschläge und Verbesserungen, die aufgrund dieser Diskussion entstehen, sind soweit wie möglich einzuarbeiten.

3. Beteiligung

Bei zunehmender Erfahrung und mit zunehmendem Alter können Schülerinnen und Schüler selbst bei der Erstellung der Kriterien mitwirken und Beurteilungen durchführen, z. B. mittels Beobachtungen, als Schülermitbeurteilung.

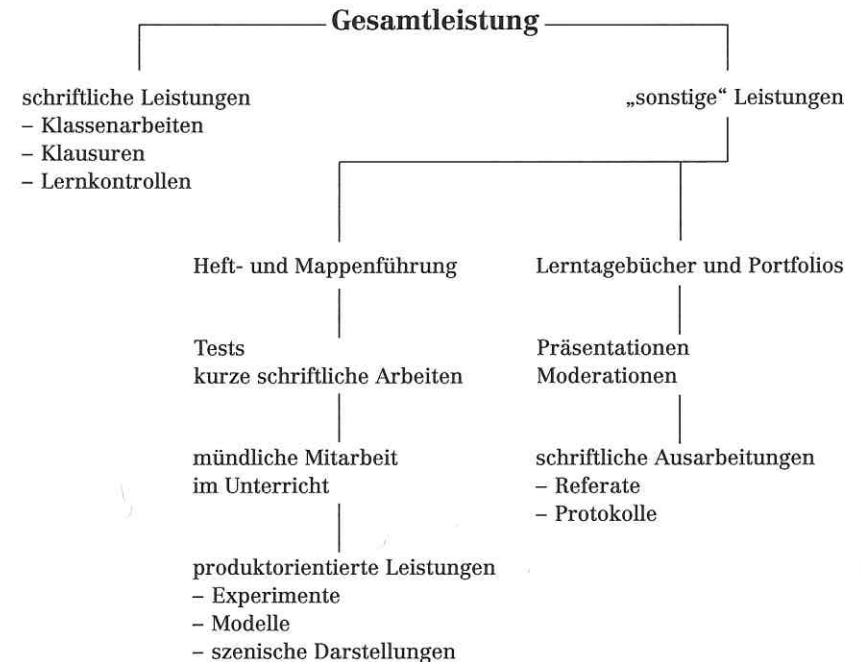
4. Detailverständnis

Schülerinnen und Schüler müssen im Detail verstehen, welche Leistung sie erbringen müssen, um die einzelnen Beurteilungskriterien zu erfüllen. Dies setzt detaillierte Überlegungen der Lehrkraft voraus, wie die jeweiligen Kriterien feststellbar sind, wie sie bewertet werden und ob sie für Schülerinnen und Schüler verständlich (formuliert) sind, also ihren sprachlichen Möglichkeiten entsprechen.

5. Reflexion

Während des Unterrichts- und Beurteilungsprozesses werden immer wieder Reflexionsphasen eingefügt, sodass regelmäßig und gemeinsam über diesen Prozess nachgedacht und reflektiert werden kann. Unklarheiten können dann beseitigt, notwendige Veränderungen eingebracht und Erkenntnisfortschritte ermöglicht werden.

2. Anforderungen an eine transparente und gerechte Leistungsbeurteilung



Leistungspuzzle: die Einzelleistungen, aus denen sich die Gesamtleistung eines Schülers zusammensetzt