

Überprüfung und Beurteilung von Lernleistungen



Zum Ziel einer gerechten
Auslese lautet die Aufgabe
für alle gleich: Klettert
auf den Baum!

Inhalt

- 1. Das Problem**
- 2. Grundproblematik aller Leistungsbeurteilung**
- 3. Funktionen und Formen der Leistungsüberprüfung und -beurteilung**
- 4. Gütekriterien der traditionellen Leistungsmessung**
- 5. Urteilsfehler**
- 6. Bezugsnormen der Leistungsbeurteilung**
- 7. Ausgestaltung von Prüfungssituationen**
- 8. Formen der Leistungserbringung und –beurteilung im Rahmen der Neuen Lernkultur**
- 9. Fazit**

1. Das Problem

Ein elfjähriges Bauernkind aus dem österreichischen Mühlviertel hat einen Aufsatz geschrieben, der über die Berechtigung zum Besuch des Gymnasiums entscheiden soll. Das Thema lautet:

Ein schöner Herbsttag

Vor ein paar Tagen sagte mein Vater zu mir: „Bua, morgen ziagst di besser an, wir fahren in die Stadt.“ So sind wir nach Urfahr gefahren. Von dort gingen wir zu Fuß über die Brücke nach Linz. Dort besuchten wir einen Optiker. Er sah mich an und sagte: „Bua, du schiagelst ja.“ Dann musste ich viele größere und kleinere Buchstaben lesen. Jedes Mal fragte er: „Ist es so besser oder so?“ Ich sagte es ihm. Dann bekam ich Brillen. Jetzt schiagle ich nicht mehr. Das war mein schönster Herbsttag!

Fragestellung:

Welche Note würden Sie dem Elfjährigen im Hinblick auf Inhalt und Sprachgestaltung geben? Das österreichische Notensystem sieht die Noten „Sehrgut“, „Gut“, „Befriedigend“, „Genügend“ und „Nichtgenügend“ vor.

Einschätzung von 200 Experten des oberösterreichischen Schulwesens (Landesschul-Inspektoren [Regierungsschulräte], und Bezirksschul-Inspektoren [Schulräte], Direktoren und Fachgruppenleiter aus allen Schultypen) anlässlich eines Vortrages von *Rupert Vierlinger* über die Objektivität der schulischen Leistungsbeurteilung:

- ca. 10 Personen: „Ungenügend“ (Thema verfehlt)
- ca. 10 Personen: „Sehr gut“ (Originalität, Thema voll getroffen)
- ca. 50 Personen: „Gut“
- ca. 90 Personen: „Mittlere Note“
- ca. 50 Personen: „Genügend“

(aus: Rupert Vierlinger, Das Konzept der „Direkten Leistungsvorlage“, in: Felix Winter/ Annemarie von der Groeben/ Klaus-Dieter Lenzen (Hrsg.), Leistung sehen, fördern, werten. Neue Wege für die Schule, Bad Heilbrunn [Klinkhardt] 2002, S. 28f.)

2. Grundproblematik der Leistungsbeurteilung:

- **Der/ die Bewertende:**
 - Das pädagogische Selbstkonzept oder „toughe“ Lehrkräfte und „Weicheier“
 - Subjektive Theorien und das Bild vom „guten“ oder „schlechten“ Schüler
 - geschlechts- und sozialschichtspezifische Bewertung
 - Beurteilungsfehler
- **Das bewertete Individuum und das Kollektiv:**

Problematik des klasseninternen Bezugsmaßstabes:
Mangelnde Vergleichbarkeit von Schülerleistungen über die eigene Lern-gruppe hinaus

2. Grundproblematik aller Leistungsbeurteilung

* Der/ die Bewertete oder das Problem “Performanz versus Kompetenz”

Mögliche **Diskrepanz** zwischen der nicht unmittelbar beobachtbaren **Kompetenz** als „*nicht wahrnehmbare Tiefenstruktur der Kognition*“ (Bonnet 2004) und der **Performanz** als „*Oberflächenstruktur der Handlungen und Aussagen eines Menschen*“ (Bonnet 2004)

Beispiele:

Konzentrationsmängel, eine nicht verstandene Aufgabenstellung, fehlende Zeit etc können verhindern, dass „*eine Person ihre Kompetenz in einer konkreten Situation auch zur Performanz bringt.*“ (Bonnet 2004)

2. Grundproblematik aller Leistungsbeurteilung

- **Das Bewertete:**

strenge Hierarchie von Hauptfächern über Nebenfächer zu den musisch-künstlerischen Fächern, die sich in der Zensierung widerspiegelt

3. Funktionen und Formen der Leistungsüberprüfung und – beurteilung

- Bürgerliches Leistungsprinzip als **gesellschaftliches Verteilungsprinzip** im Hinblick auf Berufs- und Lebenschancen
- **Konkurrierende Verteilungsprinzipien** wie Herkunft, Anciennität, Ideologie, Bekanntheit/Beliebtheit
- Leistungsprinzip ist **nicht ohne weiteres auf schulische Bildung übertragbar** (Erziehungsauftrag versus Zwang zur Anpassung an Leistungsorientierung, soziale Herkunft verzerrt Wettbewerb um Bildungschancen).
- **Mangelnder Konsens** in der pluralistischen Gesellschaft darüber, was gute Leistung ist
- Primat der **Qualifizierungs- und Selektionsfunktion** über die pädagogische Funktion (Rückmeldefunktion, diagnostische bzw. prognostische Funktion, Ansatzpunkt zur individuellen Förderung) als gesellschaftliche Realität

Formen der Leistungsbeurteilung:

Formative Beurteilung:

- prozessbegleitend
- *Rückmeldungs- und Verstärkungsfunktion*
- informelle Leistungsbeurteilung (z.B. Unterrichtsbeobachtungen, Protokolle)

Summative Beurteilung:

- ergebnisorientiert im Hinblick auf die Vergabe einer Qualifikation
- *Qualifizierungs- und Selektionsfunktion*
- Gültigkeit von Gütekriterien

4. Gütekriterien der traditionellen Leistungsmessung

Objektivität = Unabhängigkeit der Messergebnisse von der Person des Prüfers

Reliabilität = Zuverlässigkeit einer Messung, indem das Messergebnis möglichst frei von Messfehlern ist

Validität = Gültigkeit einer Messung. Es wird tatsächlich das gemessen wird, was man vorgibt zu messen

Hinweise:

- **Objektivität:**

Formulierung von **Kriterien** und **Gewichtungen** samt detaillierter **Punkteverteilung** („Erwartungshorizonte“) vor allem bei freieren Aufgabenstellungen

- **Reliabilität:**

Formulierung mehrerer **voneinander unabhängiger Einzelaufgaben** (Problematik der Komplexität von Oberstufenaufgaben)

- **Validität:**

Gewährleistung „**unterrichtlicher Validität**“ (geprüft wird nur, was im Unterricht auch thematisiert wurde) unter Berücksichtigung der unterschiedlichen Taxonomiestufen (Wissen, Verständnis, Anwerndung, Analyse, Synthese, Evaluation)

5. Urteilsfehler

A Systembedingte Schwachstellen

- **Ungenauigkeit der Berechnung von Mittelwerten** (Vergleichsbeispiel: Ermittlung des Mittelwerts bei Buchstabennoten A, D und E ?)
- **Erhöhung der Ungenauigkeiten bei der Vergabe von Viertel- oder Halbnoten** und deren Addition
daher: **Pädagogische Verantwortung** anstelle eines „Dezimalfetischismus“
- **Problematik der Vergleichbarkeit von Schulnoten** über Klassen, Schulen oder Bundesländer hinweg
daher: direkte Vergleichsarbeiten mit Parallelklassen und Erfahrungsaustausch mit Kolleginnen/Kollegen

B Subjektive Fehlerquellen

- **Pygmalion-Effekt** (Wechselwirkung zwischen Lehrkraft und Schüler/in im Hinblick auf die Erwartungshaltung des Lehrenden gegenüber dem Lernenden)
- Einfluss von **Vor- und Zusatzinformationen**
- **Sympathie** und **Geschlecht**
- Beeinflussung durch handlungsleitende „**subjektive Theorien**“ über das Zustandekommen bestimmter Schülerleistungen (Beispiel: „Lateinklassen sind besser“)
- **Halo-Effekt** und **logische Fehler** (Gesamteindruck wird aus einem besonders hervorstechenden, beurteilungsfremden Einzelmerkmal hergeleitet: [z.B. Eloquenz = logisches Denken])
- **Stabile Urteilstendenzen** (Milde- und Strengeeffekt)
- **Reihenfolgen-Effekte** (z.B. zunehmende Milde bei der Korrektur oder eine erste Prüfung, die Maßstäbe setzt)

6. Bezugsnormen der Leistungsbeurteilung

- eine Übung (nach F. Rheinberg [2002]) -

Erreichte Punkte				Beurteilung des letzten Test- ergebnisses (+ + bis - -)
1. Test	2. Test	3. (letzter) Test		
1. 60	55	50		
2. 25	25	25		
3. 85	80	75		
4. 50	50	50		
5. 65	70	75		
6. 15	20	25		
7. 40	45	50		
8. 75	75	75		
9. 35	30	25		

Problematik der Bezugsnormenorientierung

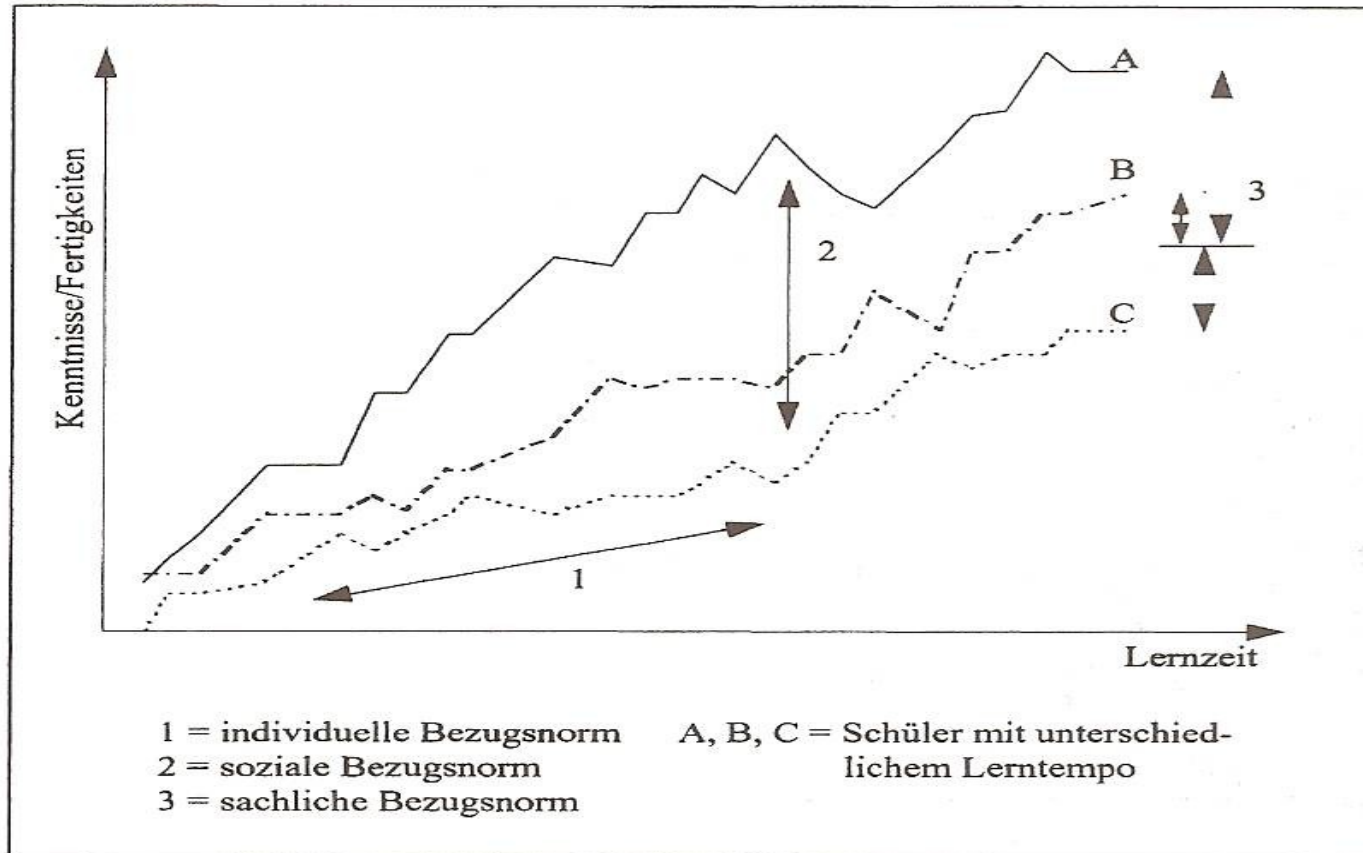
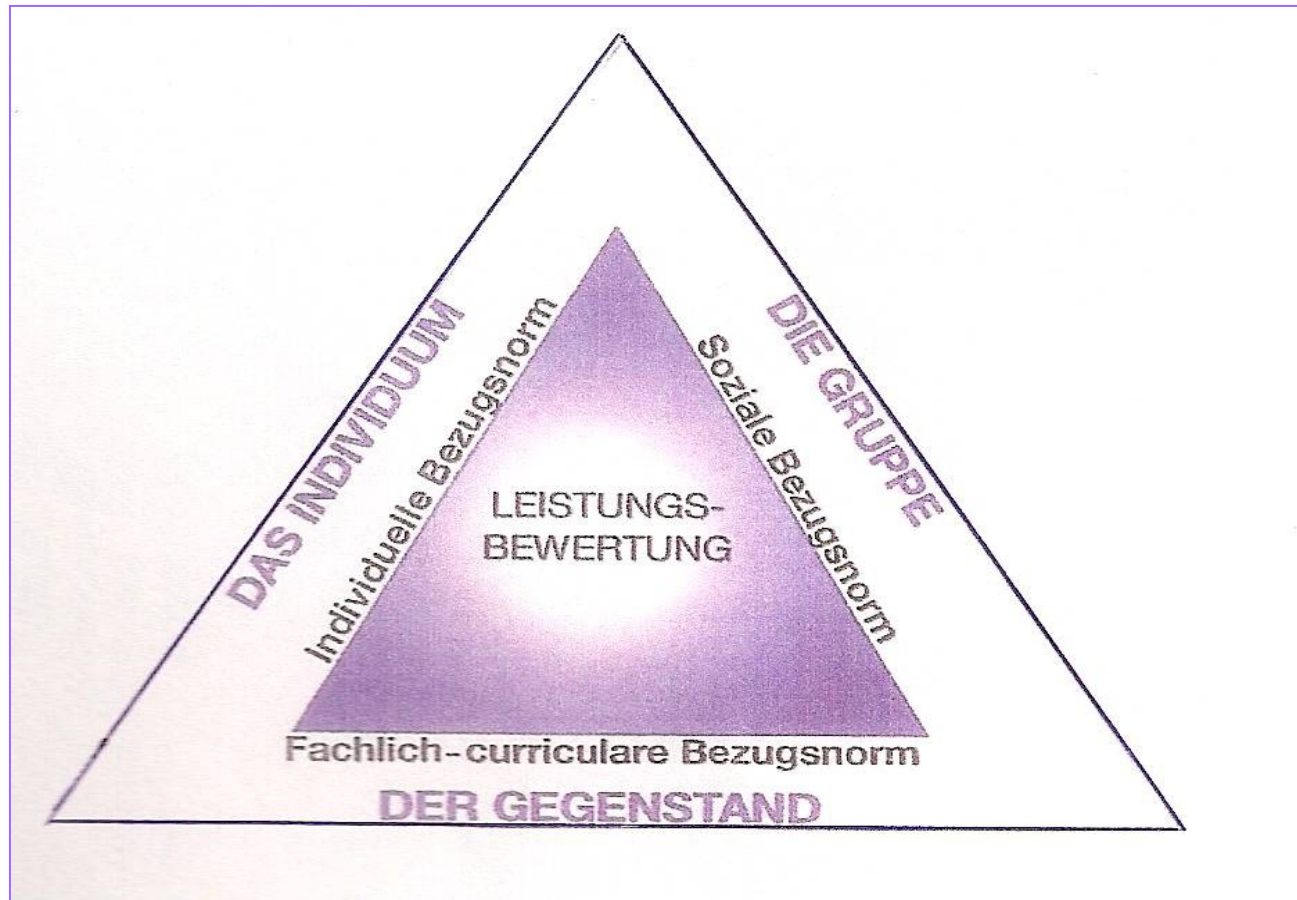


Abbildung 1: Vergleichsperspektiven bei drei Bezugsnormen zur Leistungsbewertung (nach Rheinberg & Krug, 1999, S. 41)

6. Bezugsnormen der Leistungsbeurteilung

(nach F.E. Weinert [2002])



Erläuterungen:

- **Soziale Bezugsnorm:**

Vergleich der individuellen Lernleistung mit den Leistungen der Referenzgruppe

- **Individuelle Bezugsnorm:**

Vergleich der individuellen Lernleistungen eines Schülers/ einer Schülerin mit seinen/ihren eigenen, zu früheren Zeiten erbrachten Leistungen

- **Kriteriumsorientierte (fachlich-curriculare) Bezugsnorm:**

Vergleich der individuellen Lernleistung mit dem Lernziel (Kriterium)

Realisierung der Bezugsnormen im schulischen Alltag:

Ausrichtung der Leistungsmessung und –bewertung an den curricularen Vorgaben sowie an der Bezugsnorm der Lerngruppe

Nachteile:

- Verzicht auf eine besonders für schwächere und stärkere Lerner wichtige differenzierte Leistungsrückmeldung
- Mangelnde Transparenz der individuellen Leistungsentwicklung
- Unzutreffende Einschätzung des absoluten Leistungsniveaus der unterrichteten Lerngruppe

Vorteile:

- Vergleichbarkeit des Lernentwicklungsstandes durch die Entwicklung von Bildungsstandards
- treffende Widerspiegelung der Leistungsunterschiede innerhalb der Lerngruppe

- **Fazit:**

Zu erstreben ist die...

- Orientierung der Leistungsbewertung an der kriterialen oder sachbezogenen Bezugsnorm (→ Formulierung von Bildungsstandards)
- Integration von qualitativen Verfahren der Leistungsrückmeldung unter Berücksichtigung des individuellen Lernentwicklungsstandes der Schülerinnen und Schüler (individuelle Bezugsnorm)

Anforderungen an ein Benotungsmodell:

- **Transparenz und logische Nachvollziehbarkeit** der Bewertung
- **Normvalidität** (Passung von zu messendem Leistungsmerkmal und entsprechender Bezugsnorm: Rangplatz in der Gruppe, Kriteriumsorientierung individueller Lernfortschritt)
- Ausrichtung der Leistungsbewertung an allgemein gültigen **Bildungsstandards**
- Integration **qualitativer Verfahren** in die Leistungsbewertung
- **Flexibilität in der Bewertung** und „pädagogisch günstige Voreingenommenheit“ (Weinert/ Schrader [1986])
- **Kritische Reflexion unvermeidlicher Messfehler** und deren gleichmäßige bzw. sinnvolle Verteilung auf die einzelnen Leistungsbereiche
- Ableitung **individueller Förderangebote** aus entsprechenden Maßnahmen der Leistungsbewertung

7. Ausgestaltung von Prüfungssituationen

A Allgemeine Grundsätze:

- **Proportionale Abbildung** = Prüfungsaufgaben stellen einen repräsentativen Querschnitt der vorangegangenen Unterrichtsinhalte dar
- **Variabilität** = abwechslungsreiche Ausgestaltung von Prüfungssituationen

B Auswahl, Umfang, Reihenfolge etc. der Aufgaben:

- * **Wechsel von geschlossenen, halb offenen und offenen Aufgabenformen**

- **Wechsel zwischen vielen, weniger umfangreichen Aufgabenformen und komplexeren Aufgabenstellungen in derselben Prüfung oder Konzeption jeweils unterschiedlicher Prüfungen**
- **Unterschiedlicher Schwierigkeitsgrad der einzelnen Aufgaben** mit anspruchsvollsten Aufgaben in der Mitte
- Erstellung einer **Musterlösung** vor der schriftlichen Arbeit
- Festsetzung der **Mindestkompetenz** (Note „ausreichend“)
- **Planung der äußeren Rahmenbedingungen** (Zeitpunkt, Dauer, zugelassene Hilfsmittel, Eröffnungs- und Schlussphase der Prüfung, Reaktion auf Täuschungsversuche)

Grundregeln zur Durchführung mündlicher Prüfungen

(in Anlehnung an Georg E. Becker [2007])

- Verzicht auf fragwürdige Rituale
- Konzentration auf den Kandidaten
- Kandidat/in mit ihrem Spezialgebiet beginnen lassen
- Anfangs den Prüfling referieren lassen
- Aktiv zuhören
- Bei unverständlichen Äußerungen um Präzisierungen bitten
- Unvollständige Beiträge weiterführen lassen
- Wechselspiel zwischen Anheben und Senken des kognitiven Niveaus
- (Prüfungsaufgaben in Anlehnung an die Taxonomie der kognitiven Fähigkeiten erstellen: Kenntnis – Verständnis - Kenntnisse übertragen – Analysieren – Beurteilen)

- Protokollieren des Prüfungsverlaufs
- Beurteilungsfehler reflektieren (z.B. Hof-Effekte, Erwartungsfehler, Fehler des ersten und letzten Eindrucks)

8. Formen der Leistungserbringung und –beurteilung im Rahmen der Neuen Lernkultur

Grundannahme:

Erweiterter Lern- und Leistungsbegriff: **Kompetenzorientierung** unter Berücksichtigung der **fachlich-inhaltlichen**, **methodisch-strategischen** und **sozialkommunikativen** Aspekte des Lernens



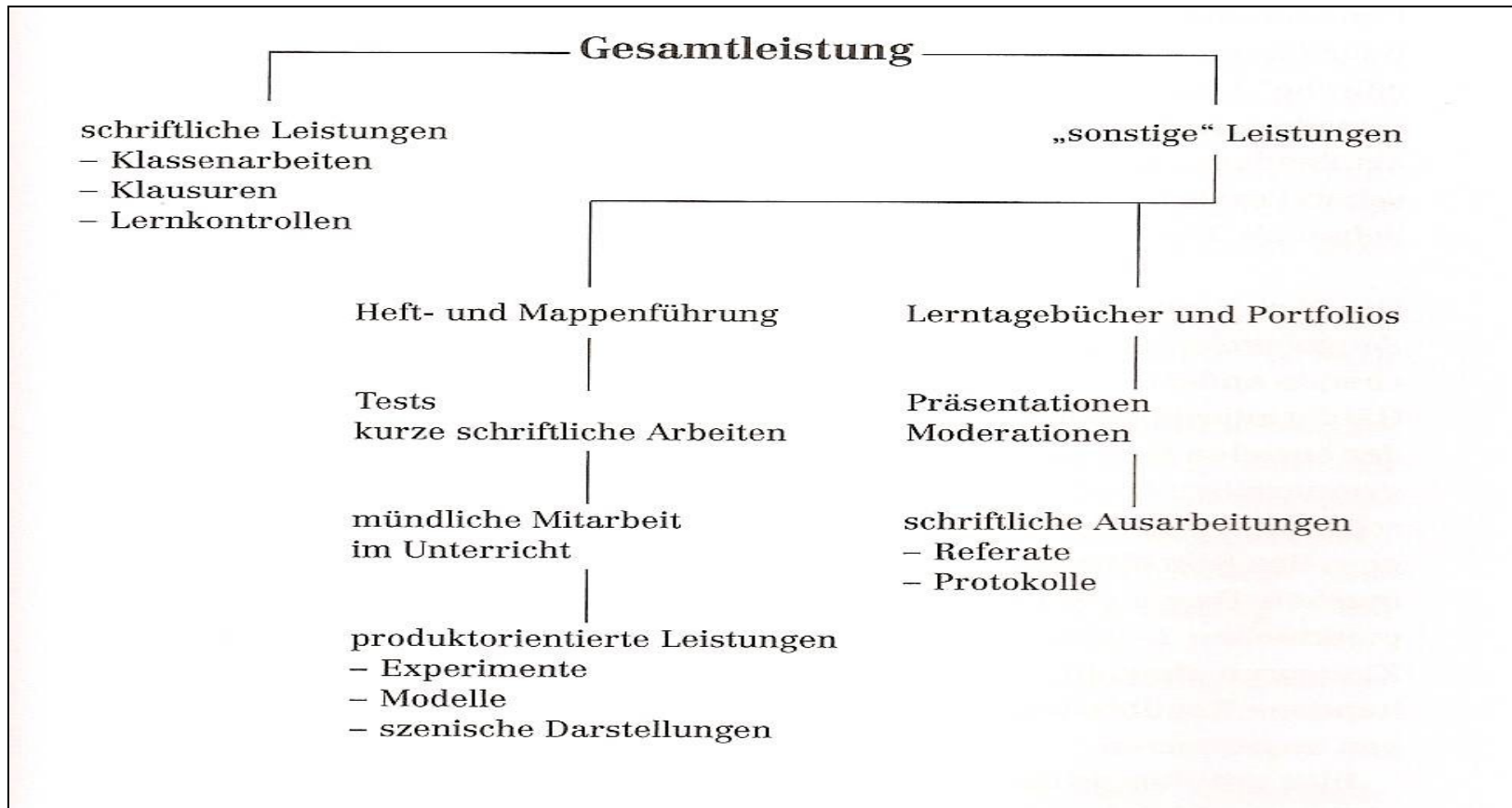
- **Fachlich-inhaltliche Anforderungen:** 1. Wissen, 2. Verstehen, 3. Erkennen, 4. Beurteilen
- **Methodisch-strategische Anforderungen:** Exzerpieren, Nachschlagen, Organisieren, Planen etc.
- **Sozial-kommunikative Anforderungen:** Zuhören, Argumentieren, Diskutieren, Kooperieren etc.

Konkretisierungen

- **„Direkte Leistungsvorlagen“** (z.B. eigene Kurzgeschichten, Rezensionen, Filmanalysen, Plakate, Werkstücke etc.) und **Portfolios**
- **Beurteilung von Lern- und Leistungsprozessen** in Form von *self-, peer- und teacher-assessment*
- **Lernpartnerschaften**
- **Projekte und Gruppenpräsentationen** mit der Notwendigkeit einer Bewertung der individuellen Lernleistung

Möglichkeiten der Zusammensetzung der Gesamtleistung

(nach Paradies/Wester/Greving [2005])



1. Persönliches Fazit

In der momentanen Situation des deutschen Schulwesens ist eine **Kombination** von traditioneller, an den **klassischen Gütekriterien der Leistungsmessung** orientierten Leistungsbeurteilung, und einer durch die **Lernenden mitbestimmten Form der Leistungsermittlung und -überprüfung** im Rahmen überwiegend **selbstgesteuerter Lernprozesse** erstrebenswert.

Literatur

- Becker, Georg E., Unterrichts auswerten und beurteilen. Handlungsorientierte Didaktik Teil III, Weinheim/Basel (Beltz) 2007.
- Helmke, Andreas, Unterrichtsqualität erfassen, bewerten, verbessern, 4. Aufl., Seelze (Kallmeyer), 2005, S. 84ff.
- Jürgens, Heiko, Leistung und Beurteilung in der Schule. Eine Einführung in Leistungs- und Bewertungsfragen aus pädagogischer Sicht, 6. Aufl., Sankt Augustin (Academia Verlag) 2005.
- Paradies, Liane, Franz Wester, Johannes Greving, Leistungsmessung und –bewertung, Berlin (Cornelsen) 2005.
- Sacher, Werner, Überprüfung und Beurteilung von Schülerleistungen, in: Ders./ Hans Jürgen Apel, Studienbuch Schulpädagogik, 2. Aufl., Bad Heilbrunn (Klinkhardt) 2005, S. 274- 297.
- Weinert, Franz. E. (Hrsg.), Leistungsmessungen in Schulen, 2. Aufl., Weinheim/ Basel (Beltz), 2002.
- Wengert, Hans Gert, Leistungsbeurteilung in der Schule, in: Gislinde Bovet/ Volker Huwendiek (Hrsg.), Pädagogik und Psychologie für den Lehrberuf, 4. Aufl., Berlin (Cornelsen) 2004, S. 294-319.
- Winter, Leistungsbewertung. Eine neue Lernkultur braucht einen anderen Umgang mit den Schülerleistungen, Baltmannsweiler (Schneider) 2004.