

Lernergebniskontrolle und Leistungsbeurteilung

1. Was Lehrer, Schüler und Eltern dazu sagen

„Ich mache prinzipiell keine Tests und auch keine mündlichen Einzelprüfungen – höchstens, wenn es um Sein oder Nichtsein geht. Sie nehmen nur wertvolle Zeit weg. Die Note in GW ergibt sich aus der Mitarbeit während des Unterrichts. Dabei halte ich mich im allgemeinen an die Gauß'sche Normalverteilung. Dadurch, daß ich die Klasse auch in Englisch habe, sind mir die allgemeinen Stärken und Schwächen meiner Pappenheimer ja gut bekannt.“

„Unser Professor stellt den Mädchen bei der Prüfung immer leichtere Fragen. Außerdem hilft er ihnen, wenn sie einmal stecken bleiben. Manche prüft er überhaupt nicht. Er sagt, auf die Mitarbeit kommt es an. Ich finde das ziemlich ungerecht, denn ich zeige mindestens ebenso häufig auf wie manche dieser Interesse heuchelnden Tussis.“

„Ich gebe jetzt grundsätzlich in GW keine Fünfer mehr. Früher war ich viel strenger. Das hat mir aber mehrmals Probleme mit der Direktion eingebracht, einmal sogar mit der von den Eltern (der Vater war Rechtsanwalt) mobilisierten Schulaufsicht. Wozu sich das antun?“

„Mündliche Einzelprüfungen? Viele Schüler wollen sie. Außerdem sind sie ein erfolgreiches Disziplinierungsmittel, überhaupt, wenn man sie etwas ritualisiert: Katalog herausnehmen, mit der schwierigen Frage anfangen, mit ein paar Bemerkungen das Selbstbewußtsein des Kandidaten abbauen. Zugeben wird das aber niemand.“

„Ich habe beim Notengeben immer ein schlechtes Gewissen. Ist das noch ein Befriedigend oder schon ein Genügend? Soll ich mehr die Mitarbeit berücksichtigen oder den Test? Karin ist bei diesen immer so nervös, getraut sich aber auch oft nicht, auf mündliche Fragen zu antworten, weil sie Angst hat, sich vor der Klasse zu blamieren.“

„Beim letzten GW-Test erreichte mehr als die Hälfte der Schüler in der Klasse meines Sohnes von insgesamt 34 Punkten nur weniger als 17. Die Frau Professor war von den Leistungen der Kinder schwer enttäuscht. Vielleicht hätte sie aber das schlechte Ergebnis auch als eine Rückmeldung über ihren Unterricht sehen und diesen selbstkritisch analysieren sollen.“

2. Allgemeines zur Lernergebniskontrolle und Leistungsbeurteilung¹

Zum Pflichtbereich jedes Lehrers gehört neben der Organisation des Lernprozesses auch das Erfassen des fachbezogenen Lernergebnisses, also des Erfolges oder Mißerfolges seines Unterrichts durch stundenimmanente sowie stundenübergreifende *Lernergebnis-*

¹) Lernerfolg wird gewöhnlich als *Leistung* bezeichnet. Die vor allem in den siebziger Jahren heftig geführte Diskussion über die Gültigkeit eines Leistungsprinzips in der Schule ist heute abgeflaut, obwohl es natürlich kontroverse Auffassungen darüber gibt, was unter schulischer Leistung zu verstehen ist (siehe dazu u.a. ausführlich LICHTENSTEIN-ROTHER 1971 und – kürzer – in der Zeitschrift „Schulheft“, Nr. 98, 2000: Die Kontroverse um den pädagogischen Leistungsbegriff, S. 34–37). Das Bundesministerium für Bildung, Wissenschaft und Kultur verwendet sowohl in den Lehrplänen als auch bei den speziellen Verordnungen im Zusammenhang mit dem Erfassen und der Bewertung des individuellen Lernerfolges die Bezeichnungen *Leistungsfeststellung* sowie *Leistungsbeurteilung*. Weil der erste Begriff aber eigentlich zwei Funktionen bezeichnet (siehe die Feedback-Fragen auf der nächsten Seite), benütze ich für ihn lieber das Wort *Lernergebniskontrolle*.

LEISTUNGSBEURTEILUNG

kontrollen. Diese informieren zunächst darüber, ob und wie weit der Schüler das Lernangebot aufgenommen und bewältigt hat. Dabei sollte sowohl **(a)** das „**Produkt**“ **des Lernens** (beispielsweise die Antwort auf gestellte Fragen, Stoffwiederholungen, die Interpretation oder der Entwurf einer Karte bzw. eines Diagramms, das Auswerten eines Bildes, das Erstellen eines Posters oder Tafelbildes, die Führung eines Merkheftes oder einer Arbeitsmappe, die Präsentation des in der Gruppe erarbeiteten Ergebnisses, der Sachinhalt eines Referats, die Beantwortung der Aufgaben eines schriftlichen Tests, Videoaufzeichnungen oder die Zusammenstellung einer Fotodokumentation etc.) als auch **(b)** das **Handeln des Schülers auf dem Weg zum Ergebnis** (er verkörpert zum Beispiel originell die von ihm übernommene Rolle in einem didaktischen Spiel, geht eigenständig beim Behandeln und Lösen eines komplexen Problems vor, kann Gruppenprozesse organisieren oder zeigt Kreativität bei Themenbearbeitungen) und **(c)** die **Beteiligung des Schülers am Lernprozeß** (zeigt besonderes Interesse, arbeitet ständig mit, bringt viele innovative Anregungen in den Unterricht, arbeitet gut mit den anderen Schülern zusammen, übernimmt freiwillig oft zusätzliche Arbeiten, ist immer verlässlich u.a.m.) in die Erfassung miteinbezogen werden.

Lernergebniskontrollen dienen nicht, wie es manche Lehrer und auch Schüler (leider) oft sehen, bloß zur **Notenfindung**. Sie haben für Lehrer, Schüler und Eltern vor allem eine wichtige **Rückmeldefunktion**. Den Eltern geben sie Auskunft über den schulischen Fortgang ihres Kindes und bilden bei Beratungsgesprächen die Grundlage zur Verbesserung. Den einzelnen Schüler können sie im Vergleich mit früheren Leistungen über seinen Lernfortschritt informieren und damit sein Selbstwertgefühl stärken. Beim Vergleich der von ihm erbrachten mit der von ihm erwarteten Leistung (die ihm bekannt sein soll) zeigen sie ihm seinen momentanen Leistungsstand und machen ihn eventuell auf „Lücken“ aufmerksam. Vor allem wenn damit keine entmutigenden Selektionssanktionen verbunden sind, kann das Aufzeigen von „Lücken“ und das Bewußtmachen, daß sie das Weiterlernen schwer behindern, die Schüler anspornen, diese im Nachholverfahren (allein oder mit der Hilfe von Mitschülern) zu schließen.

Für den Lehrer ist die Lernergebniskontrolle das unbedingt notwendige „**Feedback**“ **über seine Unterrichtsarbeit**, das ihn in seinem pädagogischen (!) und didaktisch-methodischen Vorgehen bestärkt oder auffordert, dieses zu überdenken und zu modifizieren.

- „Konnte den Schülern die Relevanz der gesetzten Ziele bewußt gemacht werden?“
- „Wurden die angestrebten Ziele in der abgelaufenen Unterrichtseinheit erreicht?“
- „Bewährten sich dabei die eingesetzten Sozialformen, Arbeitsmittel und Arbeitsweisen?“
- „Verstanden auch „schwächere“ Schüler die Aufgabenstellungen?“
- „Besäßen die Schüler das notwendige methodische und sachliche Vorwissen zur Bearbeitung der Aufgabenstellungen?“
- „Reichte die zur Verfügung gestellte Zeit für die Auseinandersetzung mit dem Lerngut aus?“

Solche und ähnliche Fragen sollte sich jeder Lehrer immer wieder stellen und die Ursache schlechter Lernergebnisse nicht nur beim Schüler, sondern auch bei sich selbst bzw. bei seinem Unterricht suchen. Leider geschieht das nicht oft.

Lernergebniskontrollen liefern letztendlich auch die Grundlagen für die vom Gesetzgeber verlangte **Leistungsbeurteilung** der Schüler durch den Lehrer.² Diese erfolgt nach wie vor auf einer fünfstufigen Notenskala (Abb. 1). Das Problem dabei ist, daß es sowohl bei den mündlichen Lernergebniskontrollen als auch bei Leistungsbeurteilungen, die auf Schülerbeobachtung beruhen, oft ziemlich schwierig ist, Beurteilungen und Einschätzungen, die vielfach nicht eindeutig meßbar sind, quantifizierbar und damit auch für andere faßbar zu machen und unter eine fünfstufige Notenskala zu stülpen. Auch die in Abbildung 1 gezeigte Tabelle hilft nur wenig – ganz abgesehen davon, daß man über die in dieser verwendeten Formulierungen diskutieren kann.

Der subjektive Spielraum des Lehrers bei der Entscheidung für eine bestimmte Note auf Grund mündlicher bzw. auf Schülerbeobachtung beruhender Lernergebniskontrollen ist jedenfalls sehr groß. Dazu kommt weiters, daß die Notengebung durch eine ganze Reihe von Störfaktoren beeinflusst werden kann, zum Beispiel den sogenannten „Halo-Effekt“ (engl. *halo*: Heiligenschein). Dieser besagt, daß sich ein Lehrer bei der Leistungsbewertung oft von anderen als den geforderten fachspezifischen Kriterien leiten läßt (von Persönlichkeitsmerkmalen, den Leistungen in anderen Fächern, dem Elternhaus, dem Verhalten des Schülers u.a.m). Auch die Einflußnahme der Kollegen bzw. der Schulleitung

Abb. 1: Kriterien zur Leistungsbeurteilung nach der fünfstufigen Notenskala

	Sehr gut	Gut	Befriedigend	Genügend	Nicht genügend
a) Erfassung und Anwendung des Lehrstoffes b) Durchführung der Aufgaben	Anforderungen werden in weit über das Wesentliche hinausgehendem Ausmaß erfüllt	Anforderungen werden in über das Wesentliche hinausgehendem Ausmaß erfüllt	Anforderungen werden in den wesentlichen Bereichen zur Gänze erfüllt	Anforderungen werden in den wesentlichen Bereichen überwiegend erfüllt	Anforderungen werden nicht einmal in den wesentlichen Bereichen überwiegend erfüllt
c) Eigenständigkeit	muß deutlich vorliegen (wo dies möglich ist)	merkliche Ansätze (wo dies möglich ist)	Mängel bei b) werden durch merkliche Ansätze ausgeglichen		
d) selbständige Anwendung des Wissens und Könnens	muß vorliegen (wo dies möglich ist)	bei entsprechender Anleitung (wo dies möglich ist)			

Quelle: Informationsblätter zum Schulrecht, Teil 3, S. 21.³

²) Siehe u.a. die Verordnung über die Leistungsbeurteilung in Pflichtschulen sowie mittleren und höheren Schulen vom 24. Juni 1974, Bundesgesetzblatt Nr. 371, in der Fassung der Verordnungen BGBl. Nr. 439/1977, Nr. 413/1982, Nr. 216/1984, Nr. 395/1989, Nr. 492/1992 und BGBl. II Nr. 35/1997. Der aktuelle Verordnungstext ist auch zu finden unter der Internetadresse <<http://www.bmuk.gv.at/gesetze/index.htm>>.

³) Die „Informationsblätter zum Schulrecht“ (Wien 2000) enthalten sowohl den Gesetzestext als auch seine Kommentierung. Die Ausführungen über die Leistungsfeststellung und Leistungsbeurteilung befinden sich in Teil 3. Man kann sie unter folgender Adresse abrufen <<http://www.bmuk.gv.at/fsserv.htm>>, zuerst *Schulrecht*, dann *Informationsblätter* anklicken. Der Text steht als pdf-Datei zum Download bereit.

LEISTUNGSBEURTEILUNG

(insbesondere beim Jahresabschlußzeugnis), der vor allem Junglehrer mit befristeten Anstellungsverträgen gelegentlich ausgesetzt sind, muß in diesem Zusammenhang erwähnt werden.

Zweifelsohne handelt es sich bei der Leistungsbeurteilung um ein sehr sensibles Tätigkeitsfeld des in einer Doppelrolle, nämlich als Pädagoge, der fördern soll, und als Zensor, der rangordnen und selektieren soll, agierenden Lehrers. Das schafft bei manchen Kollegen Unbehagen, und zwar umso mehr, wenn ihnen (auf Grund langjähriger reflektierter Erfahrung) nicht nur die mangelnde Objektivität und Vergleichbarkeit, sondern auch die Diagnoseschwäche der traditionellen Leistungsbeurteilungsverfahren bewußt wird.⁴ Aus dieser Sicht ist die Suche nach und das Ausprobieren von alternativen Formen der Leistungsbeurteilung verständlich (vgl. dazu VIERLINGER 1999; SERTL 2000; BRUNNER und SCHMIDINGER 2001; JABORNEGG 1997; LISSMANN 2000).

In den anglo-amerikanischen Ländern bemüht man sich schon seit längerer Zeit, neue Instrumente und Methoden schulischer Effizienzkontrolle zu entwickeln, die besser greifen sollen als die traditionellen. Der in diesem Zusammenhang verwendete Begriff „assessment“ meint „... *the process of gathering, interpreting, recording and using information about pupils' responses to an educational task*“ (HARLEN et al. 1992, S. 217–218). Als an den Zielen ausgerichtetes Feedback über Lernstrategien und Lernmittel übernimmt dort „assessment“ in verschiedenen Formen in der Schulpraxis eine wichtige Funktion beim Aufbau von Lernkompetenzen der Schüler und von Ergebniskontrollen des Unterrichts (vgl. GRIMWADE 1996; LAMBERT 1996). Dabei geht heute der Trend allgemein in die Richtung, Leistungsbeurteilung als einen gemeinsam mit den Schülern verantworteten Prozeß der Kriterienfindung und Kriterienanwendung im Rahmen der Lernergebniskontrolle zu verstehen.

3. Mündliche Verfahren der Lernergebniskontrolle und Leistungsbeurteilung in „Geographie und Wirtschaftskunde“

Es ist wichtig, die Begriffe Lernergebniskontrolle und Leistungsbeurteilung auseinanderzuhalten (und auch die Schüler darüber zu informieren, um das Vertrauensverhältnis zu festigen), denn – wie oben bereits gesagt – soll nicht jede Kontrolle gleich mit einer Note für die erbrachte Leistung bewertet werden. Die Möglichkeiten, das Lernergebnis in Geographie und Wirtschaftskunde festzustellen und zu bewerten, sind äußerst vielfältig und man sollte möglichst viele davon auch für die Beurteilung heranziehen. Im folgenden werden zunächst einige „traditionelle“ Verfahren aufgezählt, wobei mit den mündlichen begonnen wird.

3.1 Im Rahmen der Wiederholung von Lerninhalten

Zu Beginn einer Unterrichtseinheit (oder an ihrem Ende) fassen beispielsweise im Rahmen der Wiederholung mehrere Schüler oder ein einzelner auf Grund von Lehrerfragen

⁴ Zur Fragwürdigkeit der Notengebung vergleiche man u.a. K. INGENKAMP (1995), F. EDER (1997), R. VIERLINGER (1999) oder J. THONHAUSER (1993).

oder in Form eines selbständigen Berichtes die Ergebnisse der vergangenen Stunde zusammen bzw. erklären wichtige Begriffe noch einmal – mit oder ohne Atlas bzw. mit Hilfe des Tafelbildes oder der Notizen, die sie in ihr Heft (ihre Lernmappe) eingetragen haben. So ergibt sich Gelegenheit, Wichtiges noch einmal zu erörtern und zu festigen, Unklarheiten auszuräumen, Richtigstellungen vorzunehmen, und man kann auch an die neue Unterrichtseinheit anknüpfen.

Beliebt sind in der Sekundarstufe I Wiederholungen in Form eines Spiels, wobei zwei Teams in einem Frage- und Antwortgespräch um Punkte kämpfen. Die Schüler wurden in der Stunde vorher aufgefordert, sich mit den Zielen (!) und dem Inhalt der zu wiederholenden Unterrichtseinheit zu beschäftigen und sich Fragen bzw. Aufgaben darüber zu überlegen, die sie natürlich auch selbst richtig beantworten können sollten. Dabei kann man gut erkennen, wie weit sie das in dieser Unterrichtseinheit Angebotene aufgenommen und verstanden haben. Überdies ist es sehr lehrreich, festzustellen, was sie für wichtig hielten und welche ihrer Fragestellungen sie als leicht oder als schwierig beurteilen.⁵

3.2 Im Rahmen der Neudurchnahme von Lerninhalten

Auch die mündlichen Äußerungen der Schüler bei der Neudurchnahme von Lerninhalten sollten zur Lernergebniskontrolle herangezogen werden. In einem aufbauenden GW-Unterricht merkt man im Lehrer-Schüler-Gespräch sehr schnell, ob und wie weit die für das Neue notwendigen und bereits früher erarbeiteten Kenntnisse und Fertigkeiten bei den einzelnen Schülern vorhanden sind oder fehlen und kann dann durch Fragen, Impulse, Hinweise etc. entsprechende Maßnahmen treffen. Wenn beispielsweise in der 7. Klasse (11. Schulstufe) die Schüler aufgefordert werden, zu unterschiedlichen Auffassungen in der Pensionsfrage Stellung zu nehmen, brauchen sie für eine seriöse Auseinandersetzung u.a. die Kenntnis der Prognosen über die zukünftige Bevölkerungsentwicklung und zu deren Verständnis wiederum die Kenntnis bestimmter demographischer Prozesse und Begriffe. Auch die Fertigkeit, ein Häufigkeitspolygon zur Altersstruktur („Alterspyramide“) zu interpretieren und die sich daraus ergebenden Konsequenzen für die Zukunft abzuleiten, wäre sehr nützlich.

Selbstverständlich wird der Lehrer sowohl bei der Wiederholung als auch bei der Neuerarbeitung die punktuellen Leistungen der einzelnen Schüler bewerten, indem er während des Unterrichtsverlaufes Lob und Tadel verteilt. Nach der Stunde sollte er jedoch dann vor allem qualitative und besonders originelle Äußerungen einzelner Schüler mittels eines Symbols – nicht mit Noten – in sein Schülerblatt/-buch eintragen. Wenn er dieses dann von Zeit zu Zeit überprüft, wird er merken, welche Schüler häufige, welche seltene, welche gar keine Eintragungen aufweisen. In der Folge wird er versuchen, die letztere Gruppe beim Wiederholen und Neubearbeiten stärker heranzuziehen. Die laufenden Eintragungen im Schülerblatt/-buch sind wichtig, denn sie minimieren die sog.

⁵) Die Regeln und die Durchführung eines solchen zur Wiederholung und Festigung eingesetzten Spiels sind genau beschrieben in den Erläuterungen zu dem Spiel „Ein transnationales Unternehmen“ in W. SITTE et al. (1999): *Didaktische Spiele für den offenen Unterricht in GW*. Wien, Verlag Ed. Hölzel. Siehe auch <<http://gw.eduhi.at/didaktik/woess/ou/ou.htm>>.

LEISTUNGSBEURTEILUNG

„Erinnerungstäuschung“ und helfen ganz wesentlich bei der am Jahresende bzw. bei einer Zwischenkonferenz verlangten Notenfindung – auch beim Elternsprechtag sind sie ein nützliches Argumentationsinstrument. Allerdings muß sich der Lehrer im klaren sein, daß seine Wahrnehmung der Schülerleistungen im Rahmen der Wiederholung und der Neudurchnahme nur Zusatzfunktion bei der fachlichen Gesamtbeurteilung hat.

Abgesehen davon, daß vom einzelnen Schüler meist nur (oft beliebig gestreute) Teilmomente des Unterrichts wiedergegeben werden, gibt es noch zahlreiche andere Faktoren, die die Schülerbeurteilung beeinflussen können, wie beispielsweise den sog. „Logischen Fehler“,⁶ den „Erwartungsfehler“,⁷ die sprachliche Ausdrucksweise des Schülers,⁸ den „Fragestil“ des Lehrers etc. Außerdem stellt sich bei der Beurteilung der mündlichen Beteiligung das Problem, wie eine rege, jedoch von simplen Antworten geprägte Beteiligung im Verhältnis zu seltenen, jedoch substantziellen Beiträgen eines Schülers bewertet werden soll.

3.3 Bei Referaten

Zur mündlichen Leistung zählen auch die von Schülern gehaltenen Referate. Sie sollten in der Sekundarstufe I im allgemeinen nicht länger als zehn Minuten, in der Sekundarstufe II nicht länger als 15 Minuten dauern (was der Lehrer bei der Themenwahl des Schülers berücksichtigen muß) und es sollte sich eine Diskussion daran anschließen. Beurteilungskriterien in GW (die alle Schüler kennen müssen) sollten in erster Linie Sachrichtigkeit, Schwerpunktsetzung sowie die Art der Unterstützung des Vortrags durch Karten, Diagramme, Tabellen und Bilder sein – nicht das rhetorische Verhalten, obwohl natürlich auch auf die Korrektheit der Sprache zu achten ist. Auf der Sekundarstufe II kann bei entsprechenden Gegebenheiten auch die vom Schüler selbständig erarbeitete Kritik an manchen Aussagen, Auffassungen bzw. Meinungen bewertet werden. Eine verbale (auch das Positive herausstreichende) Beurteilung ist der Ziffernbenotung vorzuziehen. An ihr können auch die Klassenkollegen beteiligt sein, doch muß man dabei auf deren Sympathien und Antipathien gegenüber dem Mitschüler achten. Da die Schüler die Ausarbeitung ihrer Referate im allgemeinen außerhalb der Schule erledigen, ist es (besonders im Zeitalter des Internet) wichtig, daß man sich über ihren eigenständigen Anteil daran informiert.

⁶) Dazu gehören Interferenzen zwischen dem wahrgenommenen Schülerverhalten (z.B. „meldet sich nicht spontan“ oder „ist unfähig, bei Aufruf gleich zu antworten“) und dem Urteil über den Schüler („ist uninteressiert“ oder „ist unkonzentriert“). Vgl. KLEBER 1976, S. 41.

⁷) Dieser Fehler tritt vor allem dann auf, wenn man den Schüler jahrelang unterrichtet hat und glaubt, ihn genau zu kennen.

⁸) Geradezu den klassischen Beweis dafür erbrachte die von BIRKEL und PRITZ (1980) durchgeführte Untersuchung einer mündlichen Abiturprüfung in Geographie, die mit der Note „3“ beurteilt worden war. Die Prüfung wurde mit der Videokamera gefilmt, danach ein Drehbuch angefertigt, das den gesprochenen Text der Prüfung mit allen Gesten und Verweisungen zur Karte von derselben Person einmal in flüssig gesprochenen 16 Minuten und ein zweites Mal in etwas stockender Redeweise in 21 Minuten wiedergab. 101 Geographielehrer beurteilten entweder die langsame oder die schnelle Fassung mit Noten von 1 bis 5. Der Durchschnitt der schnellen Fassung lag bei 2,57, während das Notenmittel für das langsame Sprechtempo 3,44 betrug – und das bei absolut gleicher fachlicher Examensleistung!

3.4 Im Rahmen der formellen Einzelprüfung

Die mündliche Prüfung als formelle Einzelprüfung⁹ wird vor allem zur Leistungsbeurteilung eingesetzt. In der *Verordnung über die Leistungsbeurteilung* (siehe Fußnote 2) bzw. in den *Informationsblättern zum Schulrecht* (siehe Fußnote 3) wird die Mitarbeit im Unterricht bei den Wiederholungen und der Neudurchnahme von Lerninhalten so hoch eingeschätzt, daß die formelle Einzelprüfung nur dann (einmal im Semester) durchgeführt zu werden braucht, wenn (a) der Lehrer ohne sie kein sicheres Urteil über die Leistung des Schülers abgeben zu können glaubt oder (b) wenn der Schüler es selbst wünscht.

Der Termin der Prüfung ist mindestens zwei Unterrichtstage vorher bekanntzugeben. Ihre Höchstdauer soll laut Leistungsbeurteilungsverordnung in der Sekundarstufe I nur zehn, in der Sekundarstufe II nur 15 Minuten betragen, doch werden kleine Überschreitungen toleriert. Geprüft darf nur das werden, das im Unterricht durchgenommen wurde (sein Umfang bzw. klare Eingrenzungen sind bekanntzugeben). Bei der Auswahl des Prüfungsinhalts entscheide man nach dem Prinzip der Bedeutsamkeit. Gewöhnlich wird die Prüfung aus zwei (wegen der zur Verfügung stehenden Zeit aber höchstens aus vier) voneinander unabhängigen Fragen bestehen. Über das Ergebnis sollte gesprochen werden und es darf nicht allein die Semester- oder Jahresendnote ergeben, sondern muß in die gesamte fachliche Beurteilung des Schülers einfließen.

Man verlange nicht nur einfaches „Reproduktionswissen“ („Nenne vier charakteristische Merkmale des Hochgebirges“), sondern auch höherstufige Leistungen, beispielsweise „Anwendungswissen“ („Sie¹⁰ sollen an jemanden 100 Euro bezahlen. Welche Möglichkeiten der Zahlung haben Sie, wenn sowohl Sie als auch der Empfänger ein Konto benützen bzw. wenn weder Sie noch der Empfänger ein Konto benützen?“). Solche Fragen sollte man dann allerdings auch höher bewerten. Eine von den Anforderungen her anspruchsvoller gestaltete, formelle mündliche Prüfung besitzt im Rahmen der Leistungsfeststellung besondere Relevanz, wenn der Schüler auch kontroverse Ansichten vergleicht und bewertet (zum Beispiel bei wirtschaftspolitischen Sachverhalten oder strittigen Raumordnungsfragen) oder wenn Verständnis für ökonomische Gesetzmäßigkeiten bzw. problemlösendes Denken bei der Aufgabenstellung festgestellt werden kann. Der Beurteiler sollte daher bei formellen mündlichen Prüfungen auch derartige Fragen bzw. Aufgabenstellungen vorbereiten. Was spricht eigentlich dagegen, Fragen schriftlich auf DIN-A6-Karten zu notieren, sie in leichte, mittelschwierige und schwierige einzuteilen und sie dann den Prüfling auf Grund seiner Wahl ziehen zu lassen.

In GW ist es oft angebracht, konkrete „Unterlagen“ in die Prüfung miteinzubeziehen – das können Karten, Bilder, Diagramme, Presseauszüge, Geräte und Sachgegenstände etc. sein (Beispiele: „Zeigen Sie, (a) wie man das GPS-Gerät benützt und (b) erklären Sie seine Funktionsweise“ oder „Vergleiche die beiden Kartenausschnitte und beschrei-

⁹) Auf die mündliche Matura wird hier nicht eingegangen, da sie einen eigenen Stichwortartikel benötigt.

¹⁰) Die unterschiedlichen Anredeformen sollen die Zuordnung der Fragen zur Sekundarstufe I bzw. II andeuten.

LEISTUNGSBEURTEILUNG

be die Landschaftsveränderungen“). Auch den PC und das Internet kann man in die Prüfung einbauen („Holen Sie sich (1.) aus dem Internet das heutige Wettersatellitenbild von Europa und analysieren Sie (2.) die Wetterlage“). Selbstverständlich wird der Lehrer die Fragen und stichwortartig auch ihre Beurteilung in seinem Katalog notieren. Topographiefragen ohne Benützung stummer Karten (z.B. als Overhead-Transparente) oder ohne Verwendung von Tafelskizzen sind im allgemeinen abzulehnen („Nenne vier Flüsse Afrikas“). Aus dem gleichen Grund ist es auch nicht sinnvoll, topographische Kenntnisse mit Kreuzworträtseln zu prüfen.

4. Tests

Obwohl es auch bei Tests zahlreiche Bewertungsprobleme gibt, weisen sie bei entsprechend richtiger (!) Gestaltung doch einen geringeren subjektiven Spielraum der Beurteilung auf als alle Formen der mündlichen Überprüfung. Zudem ist bei ihnen die Beweisführung gegenüber dritten Personen (Schülern, Eltern, Direktor, Schulinspektor) leichter. Viele Lehrer greifen deshalb im Zusammenhang mit der Leistungsbeurteilung zu diesem Verfahren.

4.1 Standardisierte Tests

Der wissenschaftlich abgesicherte *standardisierte Test* wird gewöhnlich von Experten mit hohem methodischen Aufwand nach den Regeln der klassischen Testtheorie konstruiert (siehe u.a. die umfangreiche Darstellung von H. LUKESCH 1998). Mit seiner Hilfe sollen curricular geplante Lernvorgänge möglichst gültig und zuverlässig erfaßt werden, um Schulleistungen objektiv vergleichen zu können. Solche Tests werden gerne in der pädagogischen und psychologischen Forschung verwendet. Im alltäglichen Unterricht spielt der standardisierte Test allerdings keine Rolle. Weil die Lehrer aber von Zeit zu Zeit damit belästigt werden, sollen im folgenden einige unser Fach betreffende Probleme am Beispiel der 1995 in 50 Staaten an 600 Schulen mit zirka 18.000 Schülern und 2000 Lehrern in Mathematik, Physik/Chemie, Biologie und Umweltkunde sowie Geographie und Wirtschaftskunde durchgeführten „Third International Mathematics and Science Study“ (TIMSS) kurz aufgezeigt werden. In Österreich wurde die Durchführung dieser Studie vom damaligen Unterrichtsministerium einem erziehungswissenschaftlichen Universitätsinstitut übertragen.

Der erste kritische Einwand gegen diese Studie betrifft die Zuordnung unseres Schulfaches zu den Naturwissenschaften, der zweite die durchgehend falsche Bezeichnung als „Geographie“ und nicht als „Geographie und Wirtschaftskunde“. Man könnte meinen, daß hier vielleicht nur eine gewisse Schludrigkeit der Bearbeiter vorliegt. Beim Betrachten der Liste mit „geographischen“ Themen,¹¹ auf der die GW-Lehrer diejenigen an-

¹¹) U.a. *Eigenschaften der Erde* (Erdschichten, Landformen, Gewässer, Atmosphäre, Gesteine und Boden, Eisformen), *Prozesse, die auf der Erde ablaufen* (Wetter und Klima, physikalische Kreisläufe, Aufbau und Zerfall am Beispiel von Vulkanen und Erdbeben, die geologische Zeiteinteilung, Fossilien), *Die Erde im Universum* (Beziehungen zwischen Erde, Sonne

kreuzen sollten, die sie im Schuljahr 1994/95 in der 3. und 4. Klasse „gelehrt“ haben, taucht jedoch der Verdacht auf, daß die Bearbeiter des Tests keine Ahnung vom GW-Lehrplan 1985/86 und seinem didaktischen Hintergrund hatten. Ist niemandem (an der Universität, im Ministerium, bei der Schulaufsicht und auch bei den betroffenen Lehrern) bewußt geworden, daß man den österreichischen Unterrichtsgegenstand „Geographie und Wirtschaftskunde“ nicht so ohne weiteres mit dem Gegenstand Geographie in anderen Staaten vergleichen kann, weil wir mit der Einbeziehung der Wirtschaftserziehung 1962 einen ganz entscheidenden Schritt in einen Bereich der gesellschaftlich ausgerichteten Allgemeinbildung gemacht haben, bei dem der erdräumlich und ökonomisch agierende Mensch im Mittelpunkt des Unterrichts steht? Wenn der neue Bildungsauftrag für unser Schulfach lautet, „Motive und Auswirkungen, Regelmäßigkeiten und Probleme menschlichen Handelns in den beiden teilweise eng miteinander verflochtenen Aktionsbereichen *Raum* und *Wirtschaft* sichtbar und verständlich zu machen“, kann dieses nicht den Naturwissenschaften zugeordnet werden. Die Kritik an diesem TIMSS-Test geht aber noch weiter und betrifft auch die inhaltliche Seite der Items sowie deren unpräzise Formulierung (u.a. Items B1, B5, C7).

Sollte es unserer Schulbehörde einmal einfallen, zentral erstellte standardisierte Tests zur Leistungsüberprüfung von GW an allen Schulen einzusetzen, um damit der Beliebigkeit des Unterrichtens entgegenzusteuern und eine Normierung der Anforderungen zu erreichen, so werden hoffentlich nicht nur Erziehungswissenschaftler, sondern auch ausgewiesene Fachdidaktiker und praktizierende Lehrer bei der Erstellung mitwirken.

4.2 Der informelle Test

Der *informelle Test* (engl. „*informal*“ – frei gestaltet, nicht standardisiert) wird vom unterrichtenden Lehrer erstellt. Daher bezeichnet man ihn in der englischsprachigen Literatur auch als „*teacher made*“ *test*. Der Konstruktionsaufwand muß sich dabei „an den Grenzen orientieren, die Ausbildung, Verfügbarkeit technischer Hilfsmittel und (vor allem) die Arbeitszeit dem Lehrer auferlegen“ (INGENKAMP 1985, S. 130). Die pädagogischen und fachlichen Bedürfnisse des Lehrers bestimmen die Funktion und die Form des informellen Tests. In GW wird er bisher meist nur zur Ermittlung einer sich über einen größeren Unterrichtsabschnitt erstreckenden und daher gegenüber der punktuellen Prüfung repräsentativeren Leistungsfeststellung bzw. -beurteilung herangezogen. Sehr wenige Lehrer verwenden ihn allerdings als *Feedback-Instrument* für ihren Unterricht. Der informelle Test wird im Vergleich zur mündlichen Prüfung deshalb als objektivere Leistungsfeststellung betrachtet, weil alle Schüler bei gleichen äußeren Bedingungen in der gleichen zur Verfügung stehenden Zeit dieselben Aufgaben zu lösen haben und ein für alle gültiges (vom Lehrer allerdings subjektiv gesetztes) Punktesystem die

und Mond, Planeten und das Sonnensystem, Dinge außerhalb unseres Sonnensystems, Entstehung des Universums), *Geschichte der Naturwissenschaften* (berühmte Naturwissenschaftler, klassische Experimente, historische Entwicklung von wissenschaftlichen Ideen, die industrielle Revolution und klassische Erfindungen), *Das Wesen der Naturwissenschaften* (Wissenschaftliche Erkenntnisse und Methoden), *Messen* (Gebrauch von Apparaturen, Durchführung von Routineexperimenten, Datenerhebung). Quelle: Fragebogen für Geographie und Wirtschaftskunde, Seite 2. TIMSS, Austrian IEA Research Center, 1995.

LEISTUNGSBEURTEILUNG

Antworten bewertet. Außerdem spart er wertvolle Unterrichtszeit und gibt dem Lehrer Rechtssicherheit bei Einsprüchen gegen seine Noten.

4.3 Hinweise zur Anlage informeller Tests

- 1) Man lege die Testaufgaben (Items), die sich auf eine längere Unterrichtsperiode beziehen, nicht zufällig und willkürlich fest, sondern leite sie vor allem von denjenigen tatsächlich abgehandelten Zielstellungen und Lerninhalten dieser Unterrichtsperiode ab, die man sowohl für das private, soziale und spätere berufliche Leben der Schüler als auch für deren weiteren Lernfortschritt bedeutsam hält. Selbstverständlich wird man die Schüler eine Woche vorher über den geplanten Test informieren und eventuell Hinweise über Rekapituliermöglichkeiten geben. Auch sollte den Schülern die Arbeit mit verschiedenen Aufgabentypen vertraut sein.
- 2) Der Test soll nicht nur Aufgaben enthalten, die bloß die einfache Reproduktion von Faktenwissen verlangen, sondern immer auch **Aufgaben mit höheren kognitiven Anforderungen**. Bei diesen ist im Rahmen der Bewertung eine entsprechend höhere Punktezahl vorzusehen. In der Literatur wird empfohlen, zur Erstellung der Items eine *Spezifikationstabelle in Form einer Matrix* anzulegen, in welcher der „Inhaltsaspekt“ mit dem „Verhaltensaspekt“ kombiniert wird (LUKESCH 1998, S. 503f). Vertikal trägt man die Lerninhalte der Ziele ein und horizontal die verschiedenen erwünschten Anforderungsniveaus. So wird festgelegt, auf welcher Höhe der Anforderungen welche Inhalte eines Ziels im Test erfaßt werden sollen. Als Suchinstrument dient eine der gebräuchlichen Taxonomien.¹² Aus Zeitgründen wird der im Unterrichtsalltag stehende Lehrer seine Matrix gegenüber der theoretisch möglichen allerdings vereinfachen (vgl. Abb. 2). Er soll sich auch nicht dazu verleiten lassen, sehr viele Items mit sehr hohem Anforderungsniveau zu formulieren, denn das ist genauso falsch wie das ausschließliche Abprüfen einfacher kognitiver Leistungen.
- 3) Bei der formalen Gestaltung von Testitems sollte man Einförmigkeit vermeiden, immer **verschiedene Aufgabentypen** verwenden und in GW nicht nur Texte, sondern auch Kartenskizzen, Kartenausschnitte, Diagramme, Abbildungen, Tabellen etc. (so wie im operativen Unterricht gehandhabt) einsetzen. Man unterscheidet Aufgabenstellungen mit *gebundener* und solche mit *freier* Aufgabenbeantwortung. Zu den Erstgenannten zählen **Multiple-Choice-Aufgaben** (Mehrfachwahl-Aufgaben). Bei ihnen steht im sogenannten Aufgabenstamm eine direkte Frage oder ein unvollständiger Satz und der Schüler soll die richtige(n) der vorgegebenen Antwortalternati-

¹² Außer der bekannten sechsstufigen Taxonomie von BLOOM (siehe Stichwort „*Taxonomie*“) wird manchmal auch die vierstufige von H. ROTH (Pädagogische Psychologie des Lehrens und Lernens. 7. Aufl., Hannover 1963) benützt: *Reproduktion des Wissens* (Gelerntes aus dem Gedächtnis wiedergeben), *Reorganisation des Wissens* (eigene Verarbeitung eines Sachverhalts), *Transfer* (Übertragung des Wissens auf ähnliche Fälle), *Problemlösen* (Wissen selbst produzieren, um neue Sachverhalte zu untersuchen und zu beurteilen). Vgl. dazu auch H. KÖCK (1980, S. 169). Diese Taxonomien erfassen nur den kognitiven Bereich. Die Abspaltung eines instrumentellen Bereichs ist in Geographie und Wirtschaftskunde nicht sinnvoll (siehe Stichwort „*Zielorientierung*“).

Abb. 2: Beispiel einer einfachen Spezifikationstabelle („Inhalts-/Verhaltensmatrix“) zum Thema: Gehen die Sozialleistungen zurück? (11. Schulstufe)

Inhaltsaspekt	Verhaltensaspekt		
	Reproduzieren	Beschreiben und erklären	Selbständig denken und handeln
Sozialquote, Transferzahlungen	X	X	
Sozialversicherung, Familienlastenausgleichsfonds, Wohlfahrtsstaat	X	X	
Einrichtungen und Leistungen des sozialen Netzes		X	
Zusammenhang zwischen sozialen Maßnahmen und gesellschaftspolitischen Vorstellungen		X	
Probleme bei der Finanzierung		X	
Kontroverse Aussagen zum Kindergeld vergleichen und dazu Stellung nehmen			X

ven kennzeichnen (vgl. Abb. 3). Je mehr *Distraktoren* (falsche Antwortalternativen) vorhanden sind (mindestens vier oder fünf), desto geringer ist die Ratewahrscheinlichkeit. Diese kann auch durch eine angeschlossene offene Aufgabenstellung in Form einer *Kurzantwortfrage* – bei Beispiel (b) in Abbildung 3 beispielsweise: „*Begründen Sie Ihre Entscheidung*“ – eliminiert werden. Für deren Beantwortung gibt es dann zwei weitere Punkte. Man verwende nur Distraktoren, die dem intellektuellen Niveau der Schüler entsprechen – bei Beispiel (a) zum Beispiel nicht „*der Bundespräsident*“. Wichtig ist, daß richtige Lösungen nicht immer an der gleichen Stelle positioniert sind, sondern ihren Platz in der Reihe der Antworten wechseln. Multiple-Choice-Aufgaben lassen sich sowohl zum Überprüfen *einfacher* Fragestellungen – Beispiel (a) – als auch zum Erfassen *komplexer* Leistungen – Beispiel (b) – einsetzen. Tests mit ihnen sind relativ schnell durchzuführen und einfach zu korrigieren. Probleme macht allerdings oft das Finden geeigneter Distraktoren.

Zu den gebundenen (geschlossenen) Aufgabentypen gehören weiters **Zuordnungsaufgaben** (Abb. 4). Auch sie erlauben ein rasches Überprüfen von Kenntnissen. Macht ein Schüler am Anfang einen Fehler, stellen sich allerdings automatisch Folgefehler ein. Um Restzuordnungen auszuschalten, sollte es immer mehr Elemente geben als zugeordnet werden können. Das erste Beispiel in Abbildung 4 stammt aus dem von N. J. GRAVES herausgegebenen „*New Unesco Source Book for Geography Teaching*“ (Verlag Longman 1982, S. 337).¹³

Kurzantwortaufgaben (vgl. Abb. 5) gehören zu Aufgabenstellungen mit freier Antwortformulierung. Der Schüler kann die Antwort nicht aus vorgegebenen Antwort-

¹³⁾ In diesem Buch findet man auch Beispiele mit höheren Anforderungsniveaus.

Abb. 3: Beispiele für Multiple-Choice-Aufgaben

a) Die laufenden Geschäfte einer Aktiengesellschaft führt/führen

- der Vorstand
- der Aufsichtsrat
- nur der technische Direktor
- ausschließlich allein der kaufmännische Direktor
- der kaufmännische und der technische Direktor gemeinsam

1 P.

b) Welche zwei der unten angeführten Staaten haben auf Grund ihrer ökonomischen Entwicklung die besten Aussichten, zuerst in die EU aufgenommen zu werden? *

- A
- D
- F
- B
- E
- G
- C

6 P.

	1999	2000	2001	2002
A: BIP/Kopf im Vergleich zum EU-Durchschnitt (100 %)		49		
Arbeitslosenquote	17,5	18,5	17,5	18,0
Einkommen/Kopf in Kaufkraftstandards (EU-15: 100)			48	
Bruttoauslandsverschuldung in % d. BIP	53,4	54,3	47,4	48,0
Inflationsrate	10,6	12,2	7,2	6,5
B: BIP/Kopf im Vergleich zum EU-Durchschnitt (100 %)		41		
Arbeitslosenquote	12,0	13,9	16,5	16,0
Einkommen/Kopf in Kaufkraftstandards (EU-15: 100)			26	
Bruttoauslandsverschuldung in % d. BIP	41,7	40,3	38,5	39,3
Inflationsrate	7,3	10,1	6,2	5,5
C: BIP/Kopf im Vergleich zum EU-Durchschnitt (100 %)		55		
Arbeitslosenquote	7,1	6,4	6,0	5,5
Einkommen/Kopf in Kaufkraftstandards (EU-15: 100)			52	
Bruttoauslandsverschuldung in % d. BIP	60,4	68,0	64,1	59,0
Inflationsrate	10,0	9,8	9,0	6,5
D: BIP/Kopf im Vergleich zum EU-Durchschnitt (100 %)		69		
Arbeitslosenquote	13,6	12,2	11,7	11,5
Einkommen/Kopf in Kaufkraftstandards (EU-15: 100)			70	
Bruttoauslandsverschuldung in % d. BIP	27,4	33,6	33,1	38,0
Inflationsrate	6,1	8,9	7,5	6,0
E: BIP/Kopf im Vergleich zum EU-Durchschnitt (100 %)		60		
Arbeitslosenquote	8,9	9,0	8,4	8,0
Einkommen/Kopf in Kaufkraftstandards (EU-15: 100)			58	
Bruttoauslandsverschuldung in % d. BIP	42,6	40,8	37,2	38,7
Inflationsrate	2,1	3,9	4,1	4,2
F: BIP/Kopf im Vergleich zum EU-Durchschnitt (100 %)		50		
Arbeitslosenquote	19,5	21,5	23,0	23,8
Einkommen/Kopf in Kaufkraftstandards (EU-15: 100)			52	
Bruttoauslandsverschuldung in % d. BIP	49,0	56,7	55,0	53,6
Inflationsrate	2,1	3,9	4,1	4,2
G: BIP/Kopf im Vergleich zum EU-Durchschnitt (100 %)		24		
Arbeitslosenquote	14,0	18,1	18,0	17,6
Einkommen/Kopf in Kaufkraftstandards (EU-15: 100)			22	
Bruttoauslandsverschuldung in % d. BIP	82,3	84,1	77,0	76,0
Inflationsrate	0,3	10,0	7,9	6,0

*) absichtlich ohne konkrete Nennung der Staaten, damit die Schüler nicht Zeitungswissen reproduzieren

Abb. 4: Beispiele für Zuordnungsaufgaben

a) The following types of farming are described in the right-hand list. Write the appropriate letter in the boxes provided next to descriptions.

<p>A. Intensive Farming</p> <p>B. Plantation agriculture</p> <p>C. Extensive farming</p> <p>D. Shifting cultivation</p> <p>E. Market gardening</p> <p>F. No name given</p>	<p>1. Land is cultivated for a few years and abandoned when yields begin to decrease.</p> <p>2. The output per hectare is high, but the output per man may be low. Much fertilizer is used and careful cultivation occurs.</p> <p>3. One crop is usually grown on a large estate with a view to a predominantly export market.</p> <p>4. One crop is grown, usually grain, year in year out.</p> <p>5. The output per man is high but the output per hectare may be low.</p> <p>6. Animals are kept and crops grown, the animals produce milk or meat and the crops are usually fodder crops.</p> <p>7. Many crops are grown to satisfy a not too distant urban population.</p>	<table border="1" style="border-collapse: collapse; width: 60px; height: 100px;"> <tr><td style="width: 30px; height: 20px;"></td><td style="width: 30px; height: 20px;"></td></tr> <tr><td style="width: 30px; height: 20px;"></td><td style="width: 30px; height: 20px;"></td></tr> <tr><td style="width: 30px; height: 20px;"></td><td style="width: 30px; height: 20px;"></td></tr> <tr><td style="width: 30px; height: 20px;"></td><td style="width: 30px; height: 20px;"></td></tr> <tr><td style="width: 30px; height: 20px;"></td><td style="width: 30px; height: 20px;"></td></tr> <tr><td style="width: 30px; height: 20px;"></td><td style="width: 30px; height: 20px;"></td></tr> <tr><td style="width: 30px; height: 20px;"></td><td style="width: 30px; height: 20px;"></td></tr> <tr><td style="width: 30px; height: 20px;"></td><td style="width: 30px; height: 20px;"></td></tr> <tr><td style="width: 30px; height: 20px;"></td><td style="width: 30px; height: 20px;"></td></tr> <tr><td style="width: 30px; height: 20px;"></td><td style="width: 30px; height: 20px;"></td></tr> </table>																				

Σ POINTS

Quelle: GRAVES 1982, S. 337.

b) Bei diesem Beispiel müssen Klimadiagramme und die Bezeichnungen der Klimaregionen einer Europakarte zugeordnet werden. Aufgabe D verlangt eine Kurzantwort.

<p><i>A. Kennzeichne fünf der sieben Klimadiagramme (auf der nächsten Seite) mit den fettgedruckten Buchstaben (OK, KK, TK, MK, ÜK).</i></p>	<p><i>B. Schreibe die Namen der fünf Klimaregionen in die Karte (auf der nächsten Seite).</i></p>
--	---

<p>Ozeanisches Klima <input style="width: 30px; height: 20px;" type="text"/> (3 P.)</p> <p>Kontinentales Klima <input style="width: 30px; height: 20px;" type="text"/> (3 P.)</p> <p>Tundrenklima <input style="width: 30px; height: 20px;" type="text"/> (3 P.)</p> <p>Mittelmeerklima <input style="width: 30px; height: 20px;" type="text"/> (3 P.)</p> <p>Übergangsklima <input style="width: 30px; height: 20px;" type="text"/> (3 P.)</p> <p><input style="width: 30px; height: 20px;" type="text"/></p>	<p>Ozeanisches Klima <input style="width: 30px; height: 20px;" type="text"/> (1 P.)</p> <p>Kontinentales Klima <input style="width: 30px; height: 20px;" type="text"/> (1 P.)</p> <p>Tundrenklima <input style="width: 30px; height: 20px;" type="text"/> (1 P.)</p> <p>Mittelmeerklima <input style="width: 30px; height: 20px;" type="text"/> (1 P.)</p> <p>Übergangsklima <input style="width: 30px; height: 20px;" type="text"/> (1 P.)</p> <p><input style="width: 30px; height: 20px;" type="text"/></p>
---	---

C. Wo regnet es im Juli am wenigsten? Gib die Abkürzung der Station an. (2 P.)

D. Erkläre kurz, wieso es in der Sierra Nevada trotz 37° nördl. Breite mehrere Monate genügend Schnee zum Schifahren gibt. (4 P.)

Punktesumme:

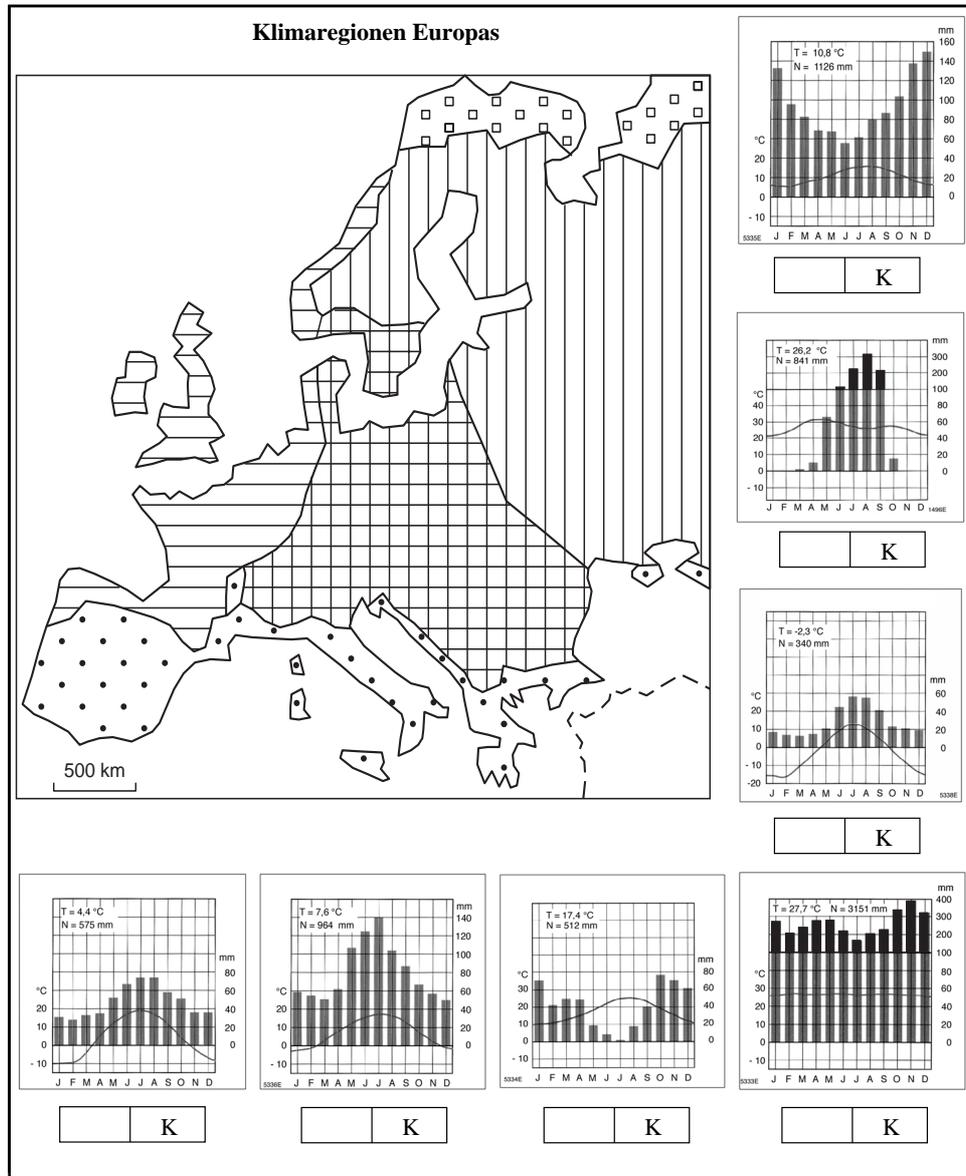


Abb. 5: Beispiele für Kurzantwortaufgaben mit freier Antwortformulierung

- Geben Sie fünf charakteristische Konjunkturmerkmale einer Rezessionsphase an. (5 P.)
- Mit welchen zwei wirtschaftspolitischen Maßnahmen kann die öffentliche Hand versuchen, die Konjunktur wieder anzukurbeln? (2 P.)
- Mit welchem geldpolitischen Instrument kann die Zentralbank einer Volkswirtschaft die Regierung dabei unterstützen? (1 P.)
- Wie wirkt dieses? (4 P.)

ten auswählen, sondern muß sie selbst finden und formulieren (Probleme: Rechtschreibfehler; Bewertungsspielräume bei fehlender Eindeutigkeit; hoher Korrekturaufwand). Man formuliere die Frage deshalb so, daß eine exakte Antwort möglich ist bzw. verlange nur wenige Worte bzw. einen einzigen Satz.

Ergänzungsaufgaben sollte man in GW vor allem zum Vervollständigen einer Tabelle, eines Diagramms, einer Graphik oder einer Karte einsetzen. Das Beispiel in Abbildung 6 überprüft einen komplexen Zusammenhang und reicht über die Kontrolle von Faktenwissen hinaus.

Abb. 6: Beispiel einer Ergänzungsaufgabe

Die Produktzyklentheorie geht davon aus, daß sowohl Sachgüter als auch Dienstleistungen auf dem Markt verschiedene Entwicklungsphasen durchlaufen.

(a) Zeichnen Sie im oberen Abschnitt des Diagramms den Verlauf der Nachfragekurve ein (2 P.) und
 (b) im unteren Teil den der Ertragskurve (4 P.)

↑ N A C H F R A G E	Entwicklung	Markter- schließung	Sättigung	Spätphase	<input type="text"/> (2 P.)
					- E R T R Ä G E <input type="text"/> (4 P.)

Aus sprachlichen und sachlichen Gründen ist von **Lückentexten** abzuraten. Das folgende Beispiel wurde in der Sekundarstufe II tatsächlich eingesetzt!

(Klima ist die typische Zusammenfassung der
 und Zustände und während in
 charakteristischer der). 7 Punkte

Die **Richtig-Falsch-Aufgabe** besteht nur aus einer Feststellung und zwei möglichen Antworten; beispielsweise „*Ungarn will der EU beitreten: O falsch O richtig*“. Der Schüler hat durch Ankreuzen, Unterstreichen oder Niederschreiben der Worte JA bzw. NEIN zu beurteilen, welche der beiden Aussagen richtig oder falsch ist. Da

LEISTUNGSBEURTEILUNG

es nur zwei Wahlantworten gibt, stehen die Chancen für den Unwissenden, die richtige Antwort zu erraten, 50:50. Dieser Nachteil ist auch durch eine zusätzlich geforderte Begründung der Auswahl nur schwer auszugleichen. Denn wie reagiert man, wenn die richtige Antwort angekreuzt, die Begründung jedoch falsch ist oder umgekehrt?

- 4) Die Aufgabenstellungen sollen immer klar und so unmißverständlich wie möglich das sagen, was der Schüler zu machen hat. Dabei verpacke man nicht mehrere Aufgaben in einen Satz, sondern zerlege sie in (a), (b) etc. Der Beantwortungsspielraum ist so einzugrenzen, daß die Lösung oder Nichllösung klar ersichtlich ist.
- 5) Bei der Konstruktion der Items probiere man (vor allem als noch wenig erfahrener Lehrer) selbst die Lösung aus und multipliziere die dazu benötigte Zeit mit dem Faktor drei oder vier. So kann man die Bearbeitungszeit leichter abschätzen und danach die **Anzahl der Items** festlegen. Bei einer größeren Anzahl und verschiedenen Anforderungsniveaus hat jeder Schüler mehr Möglichkeiten, seine Fähigkeiten zu beweisen, und der Lehrer kann die Leistung zuverlässiger (weil weniger zufallsabhängig) feststellen. Bei der Aufgabenzusammenstellung ist die Korrekturökonomie zu beachten und bei der Gestaltung des Testblattes (am besten mit dem Computer) ein übersichtliches Layout. Platz für die Antworten soll immer direkt unter oder neben den durchnummerierten Aufgabenstellungen zur Verfügung gestellt werden.
- 6) Man beginne mit nicht zu anspruchsvollen Aufgaben. Die Schüler sollen zunächst in einer „warming up-Phase“ mit der Prüfungssituation zurechtkommen. Im mittleren Testbereich können dann die schwierigsten Items angesiedelt werden. Gegen Ende, wenn die Konzentration nachläßt, sollten die Anforderungen wieder geringer werden. Manche Autoren schlagen hier auch zusätzliche Aufgaben für leistungsstärkere Schüler vor. Es gibt Kolleginnen und Kollegen, die schreiben auch ermunternde Bemerkungen auf das Testblatt, um Streß bzw. Angst bei den Schülern abzubauen, beispielsweise:

*Meine Lieben, Ihr müßt zwar bei diesem Test etwas weniger reproduzierend arbeiten, dafür aber mehr in kombinatorischer Hinsicht denken. Ich habe aber da volles Vertrauen zu Euren Fähigkeiten. In diesem Sinne wünsche ich Euch „good and successfull brain-storming and no headache“!
Und wenn es wo eine Blockade gibt, nur net d' Nerv'n wegschmeißen!*
- 7) **Topographie** sollte immer nur zusammen mit stummen Karten geprüft werden! Die höchste Punktezahl, die bei allen topographischen Fragen erreicht werden kann, sollte maximal 25 Prozent der möglichen Gesamtpunktezahl des Tests ausmachen. Sonst erzielt ein Schüler bloß mit richtigen Topographieantworten ein positives Gesamtergebnis.

4.4 Zur Bewertung informeller Tests

Sowohl die mögliche Gesamtpunktezahl als auch die bei jedem Item zu erreichenden Punkte müssen ebenso wie die auf Grund der richtigen Lösungen vom Schüler erreichte

Punktezahl auf dem Testblatt ersichtlich sein. Auf diese Weise kann sich der Schüler während der Arbeit über seinen aktuellen Punktestand informieren. Häufig findet man dort dann (leider) auch eine Notenzuteilung. Zu ihrer Ermittlung gibt es in der Literatur verschiedene, in der Praxis jedoch nur bedingt einsetzbare Vorschläge (KLAUER 1987, S. 187ff). Unzulässig ist es, sich bei der Vergabe der Noten nach der Gauß'schen Normalverteilung zu orientieren, denn diese „ergibt sich nur dann für eine Zufallsvariable, wenn diese von sehr vielen, voneinander unabhängigen Bedingungen abhängt“ (LUKESCH 1998, S. 452), was für Testleistungen nicht zutrifft.

Die oft vorgeschlagene Praxis, daß für eine positive Leistung mindestens 50 Prozent des Gesamtpunktwertes der richtigen Lösungen erreicht werden muß, wird von vielen zu recht kritisiert bzw. sogar abgelehnt (vgl. u.a. BURDIS et al. 1983, S. 52f oder VIERLINGER 1999, S. 44f). Lehrer, die überzeugt sind, daß positive Verstärkungen eher wirksam sind als negative Beurteilungen, verschieben daher den größeren Teil der Notenkurve¹⁴ sehr subjektiv oft auf deren „Sonnenseite“.¹⁵ Manche Lehrer tragen beim Korrigieren keine Noten auf dem Testblatt ein und begnügen sich bloß mit der Feststellung, ab welchem Punktwert die Leistung positiv war; außerdem loben sie einzelne Schüler, die besonders gute Leistungen erbracht haben. Sie halten nicht viel von einer Beurteilungsarithmetik und der Übergewichtung des Tests bei der Bewertung der Schülerleistung und rechtfertigen sich mit §3 Abs. 2 der schon erwähnten Leistungsfeststellungsverordnung.¹⁶ Noten ergeben sich nicht durch ein Rechenverfahren, sondern müssen *verantwortlich* gegeben werden (KIRCHBERG 1997, S. 350).

4.5 Korrektur, Rückgabe und Auswertung von Tests

Bei der Korrektur sollte man sich zuerst einen Überblick verschaffen und einige Tests von leistungsstarken und leistungsschwachen Schülern vergleichen. Damit erkennt man die Leistungsbreite und der Lehrer kann sich Gedanken machen, ob die Aufgaben verständlich formuliert und die ihnen zugeordneten Punkte richtig festgelegt waren. Es ist keine Schande, in Einzelfällen Veränderungen vorzunehmen. Allerdings muß man das den Schülern bei der Besprechung begründen. Die genaue Korrektur kann man folgendermaßen durchführen: Zuerst bei allen Schülerarbeiten nur Aufgabe 1, dann Aufgabe 2, Aufgabe 3 usw. – durch kann man besser vergleichen – oder eine Arbeit nach der anderen vollständig durchsehen und korrigieren.

¹⁴) Wenn man dem Testergebnis Noten zuteilen will, trage man beispielsweise auf der x-Achse eines rechtwinkligen Koordinatendiagramms die Punktwerte und auf der y-Achse jeweils die Anzahl der Schüler, die sie erreichten, ein und verbinde sie mit einer Linie. In der so erhaltenen Notenkurve der Klasse legt man subjektiv das „cutting score“ (die Grenze zwischen positiven und negativen Arbeiten) fest. Die Strecke von ihm zum höchsten Punktwert wird dann entweder in vier gleiche oder in Richtung zu diesem (auch subjektiv) in immer kleiner werdende Abschnitte eingeteilt.

¹⁵) Daher muß, wenn man schülerorientiert unterrichtet und den Test auch als Rückmeldung zum eigenen Unterricht verwendet, davon abgeraten werden, sowohl das „cutting score“ als auch die Zuteilung der Punkte zu Noten **vor** der Durchführung des Tests festzusetzen.

¹⁶) Dort steht: „Die unter Abs. 1 lit. c genannten Formen der Leistungsfeststellung dürfen nie für sich allein ... die alleinige Grundlage einer Semester- oder Jahresbeurteilung sein.“

LEISTUNGSBEURTEILUNG

In der Klasse bespricht man selbstverständlich die Aufgabenlösungen, wobei Wichtiges wiederholt bzw. vertieft wird. Auch dabei bekommt der Lehrer eine Rückmeldung über seinen Unterricht. Negative Leistungen von Schülern werden nicht vor der ganzen Klasse, sondern einzeln besprochen.

Wenn die Schüler (schrittweise in allen Unterrichtsgegenständen) dazu erzogen wurden, daß sie in der Lage sind, ihre Lernprozesse und Lernergebnisse selbstkritisch zu betrachten,¹⁷ dann kann man sie auch bei der Beurteilung mitwirken lassen – wobei die endgültige Entscheidung jedoch der Verantwortlichkeit des Lehrers obliegt.

5. Die direkte Leistungsvorlage (das Portfolio)

Eine der vorgeschlagenen Alternativen zur Verbesserung der Leistungsfeststellung ist die *Portfoliomethode*. Der Ausdruck kommt aus dem Italienischen und meint dort eine Mappe mit Bildern und Graphiken. Im Zusammenhang mit der Leistungsfeststellung versteht man unter „Portfolio“ die Sammlung verschiedener repräsentativer Dokumente, die den individuellen Lernweg des Schülers konkret belegen und zeigen, wieweit er die im Fach angestrebten Zielstellungen (am Ende des Semesters, eines Jahres oder seiner Schulzeit) erreicht hat. Als Dokumente, die der Schüler – (meist) in einer Mappe oder (seltener) in einem Pappkarton – aufbewahrt, kommen u.a. in Frage: ausgefüllte Arbeitsblätter, erledigte informelle Tests, ausführlich geführte „Mitschrift-Unterlagen“; selbst entworfene Kartenskizzen, Diagramme, Transparente, Plakate; Referatsunterlagen, Projektergebnisse, Protokolle über Betriebserkundungen oder Exkursionen sowie von Felduntersuchungen; Fotos, Videoaufnahmen, Zeichnungen, Tonbandinterviews, PC-Disketten, CD-ROMs, Modelle; Berichte über besuchte Ausstellungen, gesehene TV-Sendungen, gelesene Fachliteratur und unternommene Reisen etc.

Wichtig ist, daß der Schüler seine besten Arbeiten im Dialog mit dem Lehrer selbst aussucht und die Auswahl auch begründet. Auf diese Weise kann er im Lauf der Schulzeit die Fähigkeit zur realistischen Selbsteinschätzung seiner Arbeiten lernen. Es braucht bei der Anlage einer Portfoliosammlung kein dickbauchiges Konvolut zu entstehen; entscheidend ist, daß repräsentative Belegstücke das Erreichen fachlicher Kompetenzen zeigen. Um „Schwindeln“ zu verhindern, kann der Lehrer die im schulischen Betrieb zustande gekommenen Dokumente mit seiner Chiffre (oder mit einem Stempel) abzeichnen bzw. kommentieren und über die außerschulisch entstandenen Schülerarbeiten Stichproben-Nachweise verlangen.

Da die Ziffernbeurteilung nicht so schnell abgeschafft werden wird,¹⁸ kann man derzeit (abgesehen bei Schulversuchen) die Portfoliomethode nicht als Ersatz für die Ziffernbeurteilung einsetzen. Man kann sie aber in allen Schularten und -stufen zur Leistungsfest-

¹⁷) Ob das bei allen möglich ist, muß allerdings angezweifelt werden.

¹⁸) Der ehemalige Deutsche Bundespräsident HERZOG hat im Herbst 1997 beim Berliner Bildungskongreß erklärt: „Wer die Noten aus den Schulen verbannt, schafft Kuschelecken, aber keine Bildungseinrichtungen, die auf das nächste Jahrtausend vorbereiten.“ (zitiert nach VIERLINGER 1999, S. 7). In Österreich wäre zur allgemeinen Notenabschaffung eine Zweidrittelmehrheit im Nationalrat notwendig.

stellung mitheranziehen. Dabei helfen Beurteilungsraster. BRUNNER und SCHMIDINGER (2001) haben in ihrem Buch, das viele praktische Anregungen für die Arbeit mit Portfolios in der Sekundarstufe I enthält, solche vorgeschlagen. VIERLINGER, der sich schon lange gegen die Ziffernbenotung einsetzt, bringt in seinem sonst jedoch sehr lesenswerten Buch (1999, im Anhang B) leider ein aus fachdidaktischer Sicht schlechtes Portfolioeispiel zu Geographie und Wirtschaftskunde. Vor dem Einsatz der Portfoliomethode ist sehr viel Aufklärung bei Schülern und Eltern notwendig. Angehende Lehrer sollten im Rahmen ihrer Ausbildung über diese Methode nicht nur theoretisch informiert, sondern (zumindest teilweise) auch nach ihr beurteilt werden, wie bereits THONHAUSER (2000) einmal vorgeschlagen hat.

Literatur

- ANTONI, W. et al. (1989): *Forschendes Lernen im Sach- und Realienunterricht*. Wien/Linz: BMUKS und Kopten, 220 S. – Austrian IEA Research Center (1995): *Third International Mathematics and Science Study (TIMSS)*. Austria. Salzburg: Institut für Erziehungswissenschaft der Universität Salzburg. – BECKER, G. E. (1998): *Lernleistungen auswerten und beurteilen*. Weinheim: Beltz, 145 S. – BIRKEL, P. und V. PRITZ (1980): Sprechflüssigkeit und Vorinformationen als validitätsmindernde Faktoren bei mündlichen Prüfungen. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 12, S. 284–289. – BRUNNER, I. und E. SCHMIDINGER (2001): *Leistungsbeurteilung in der Praxis. Der Einsatz von Portfolios im Unterricht der Sekundarstufe I*. Linz: Veritas, 96 S. – BURDIS, H. et al. (1983): *Der Informelle Test in der allgemeinbildenden höheren Schule*. Wien: Österr. Bundesverlag, 67 S. (= Beiträge zur Lehrerfortbildung an allgemeinbildenden Schulen 23). – EDER, F. (1997): *Neue Lehrpläne – Folgerungen für die Leistungsbeurteilung*. In: *Erziehung und Unterricht* 147, S. 415–429. – GRIMWADE, K. (1996a): *Practical Approaches to Assessment, Record-Keeping and Reporting*. In: BAILEY, P. und P. FOX (Hrsg.): *Geography Teachers' Handbook*. Sheffield: Geographical Association, S. 203–209. – GRIMWADE, K. (1996b): *Assessment at Key Stage 3*. In: BAILEY, P. und P. FOX (Hrsg.): *Geography Teachers' Handbook*. Sheffield: Geographical Association, S. 211–215. – GRAVES, N. J. (Hrsg.) (1982): *New Unesco Source Book for Geography Teaching*. Harlow: Longman, 394 S. – HARLEN, W. et al. (1992): *Assessment and the Improvement of Education*. In: MOON, B. und S. MAYES (Hrsg.): *Teaching and Learning in the Secondary School*. London: Routledge, S. 273–286. – INGENKAMP, K. (1985): *Lehrbuch der Pädagogischen Diagnostik*. Weinheim: Beltz, 301 S. – INGENKAMP, K. (Hrsg.) (1995): *Die Fragwürdigkeit der Zensurengebung*. 9. Auflage (1. Aufl. 1971). Weinheim: Beltz. – JABORNEGG, D. (1997): *Das Portfolio – Möglichkeiten und Grenzen einer alternativen Prüfungsform*. In: DUBS, R. und R. LUZI (Hrsg.): *Schule in Wissenschaft, Politik und Praxis*. St. Gallen: IWP, S. 411–425. – KIRCHBERG, G. (1997): *Lernkontrollen und Leistungsbewertung*. In: HAUBRICH, H. et al.: *Didaktik der Geographie konkret*. 3. Neubearbeitung, München: Oldenbourg, S. 327–366. – KLAUER, K. J. (1987): *Kriteriumsorientierte Tests*. Lehrbuch der Theorie und Praxis lehrzielorientierten Messens. Göttingen: Hogrefe. – KLEBER, W. (1979): *Beurteilung und Beurteilungsprobleme*. Weinheim: Beltz, 186 S. – KÖCK, H. (1980): *Theorie des zielorientierten Geographieunterrichts*. Köln: Aulis, 205 S. – LAMBERT, D. (1996): *Issues in Assessment*. In: BAILEY, P. und P. FOX (Hrsg.): *Geography Teachers' Handbook*. Sheffield: Geographical Association, S. 187–201. – LICHTENSTEIN-ROTHER (Hrsg.) (1971): *Schulleistung und Leistungsschule*. Bad Heilbrunn: Klinkhardt. – LISSMANN, U. (2000): *Beurteilung und Beurteilungsprobleme bei Portfolios*. In: JÄGER, R. S. (Hrsg.): *Von der Beobachtung zur Notengebung*. Landau, S. 283–329 (= Empirische Pädagogik 5). – LUKESCH, H. (1998): *Einführung in die pädagogisch-psychologische Diagnostik*. Regensburg: Roderer, 733 S. (= Psychologie in der Lehrerbildung 3). – SACHER, W. (1999): *Tests und Klausuren in der Schule*. In: *Pädagogik* 51

LEISTUNGSBEURTEILUNG

(4), S. 43–47. – SCHANZ, G. (1973): Der Einsatz informeller Tests im Erdkundeunterricht. In: Geographische Rundschau 25, S. 22–29. – Noten nicht zu umgehen. Alternative Formen der Leistungsbeurteilung auf dem Prüfstand (2000). Themenheft der Zeitschrift „Schulheft“. Wien, 189 S. (= Schulheft 98). – SERTL, M. (2000): Wir entkommen den Zwängen nicht, indem wir sie verfeinern. Neue Formen der Leistungsbeurteilung als neue Etappe in der „zurichtenden Modernisierung“ des Menschen. In: Schulheft 98, Wien, S. 9–29. – THONHAUSER, J. (1993): Probleme der Leistungsbeurteilung und Empfehlungen zu ihrer Lösung in der neueren Fachliteratur. In: Erziehung und Unterricht 143, S. 490–497. – THONHAUSER, J. (2000): Gesammeltes Wissen anstatt übersprungener Prüfungshürde. In: Schulheft 98, Wien, S. 178–188. – VIERLINGER, R. (1999): Leistung spricht für sich selbst. Direkte Leistungsvorlage (Portfolios). Statt Ziffernzensuren und Notenfetischismus. Heinsberg: Dieck, 143 S. und 3 Anlagen.

Manuskript abgeschlossen: 2001

Wolfgang Sitte