

DOWNLOAD



Florian Kern

Mündliche Prüfungen im Allgemeinen

Mündliche Prüfungen Physik – Klasse 7–10
Hintergrundwissen
zum Prüfungsformat

Downloadauszug aus
dem Originaltitel:

AOL
verlag



Das Werk als Ganzes sowie in seinen Teilen unterliegt dem deutschen Urheberrecht. Der Erwerber des Werkes ist berechtigt, das Werk als Ganzes oder in seinen Teilen für den eigenen Gebrauch und den **Einsatz im eigenen Unterricht** zu nutzen. Die Nutzung ist nur für den genannten Zweck gestattet, **nicht jedoch für** einen schulweiten Einsatz und Gebrauch, für die Weiterleitung an Dritte (einschließlich, aber nicht beschränkt auf Kollegen), für die Veröffentlichung im Internet oder in (Schul-)Intranets oder einen weiteren kommerziellen Gebrauch.

Eine über den genannten Zweck hinausgehende Nutzung bedarf in jedem Fall der vorherigen schriftlichen Zustimmung des Verlages.

Verstöße gegen diese Lizenzbedingungen werden strafrechtlich verfolgt.

**Download
zur Ansicht**

Mündliche Prüfungen im Allgemeinen

1. Mündliche Prüfungen

Schulische Abschlüsse, Prüfungen an der Hochschule aber auch in der zweiten Phase der Lehrerausbildung oder berufliche Abschlüsse sehen mündliche Prüfungen als einen Teil der Prüfungsleistung vor oder bestehen teilweise nur aus einer mündlichen Prüfung. In vielen Studiengängen dient die mündliche Prüfung, aufgrund der kumulativ zu erbringenden Leistungen während des Studiums, ganz allein am Ende des Studiums dem Abschluss.

In einer Art subjektiver Empirie könnte man nun all diese mündlichen Prüfungen, die man irgendwann einmal absolviert hat, vor dem geistigen Auge Revue passieren lassen und wird relativ schnell feststellen, dass sie eben nicht alle gleich oder sogar ähnlich waren, sondern sich in vielerlei Hinsicht unterscheiden haben. Diese Unterschiede waren in den meisten Fällen nicht nur vom geprüften Fach oder vom Inhalt der Prüfung abhängig, sondern oft auch von vielen anderen Faktoren, die zum Teil noch nicht einmal in der Hand der Beteiligten lagen. Auch spielen rechtliche Grundlagen in vielen Fällen eine bedeutende Rolle, definieren sie doch meist die Organisation, den Umfang, teilweise auch den Aufbau und die Struktur einer vorgegebenen mündlichen Prüfung. Nicht zuletzt erscheint rückblickend die Person des Prüfers elementar und die Art und Weise, wie derjenige die Prüfung gestaltet hat.

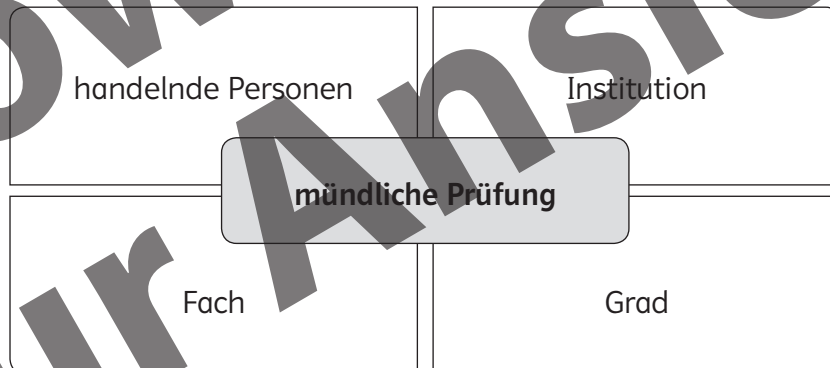


Abbildung 1: Beeinflussende Rahmenbedingungen für eine mündliche Prüfung

1.1 Definition

Wenn Prüfungen von ihrer Art her von so vielen externen Faktoren abhängig sind, dann ist es verwunderlich, dass es keinen einheitlichen Rahmen, keine Basisregularien für mündliche Prüfungen gibt. Das kann auch damit zusammenhängen, dass der Begriff nicht eindeutig definiert ist. Deswegen zunächst der Versuch einer Begriffsklärung:

Als mündliche Prüfung wird an dieser Stelle und in diesem Buch ein zeitlich abgeschlossenes Gespräch zwischen einem oder mehreren Prüfern und einem oder mehreren Prüflingen verstanden, welches der Erfassung von Leistung dient und in dem die wesentlichen und zentralen Aufgabenstellungen und Antworten in mündlicher Form, teilweise aber auch mit Hilfsmitteln erfolgen. Der Begriff „Prüfung“ kann in diesem Falle auf eine

Abschluss- oder Zwischenprüfung hinweisen, aber auch auf eine mündliche Prüfung, die als eine Form der Leistungsbewertung unter vielen anderen steht. Allen ist gemein, dass sie der Ermittlung eines Kompetenzstandes dienen und letztlich in eine Bewertung münden, die eine Note vorsieht.

Die meisten mündlichen Prüfungen sind so beschaffen, dass dem Prüfling eine gewisse Vorbereitungszeit eingeräumt wird. Diese wiederum ist, weil es anders nicht praktikabel wäre, nicht in mündlicher Form gestaltet, sondern erfolgt in den allermeisten Fällen schriftlich. Oft werden dem Prüfling für diese Vorbereitung die Kernfragen der anschließenden mündlichen Prüfung in schriftlicher Form zur Verfügung gestellt. Die obenstehende Definition muss also dementsprechend erweitert werden:

Die Vorbereitung zu einer mündlichen Prüfung, auch in Form der Vorabveröffentlichung der Kernfragen kann auch in schriftlicher Form erfolgen.



Abbildung 2 : Mündliche Prüfungen in der Schule

Abbildung 2 macht noch einmal deutlich, dass man grundsätzlich auch zwischen Prüfungen im schulischen Alltag und solchen für einen Abschluss unterscheiden sollte. Letzte Variante ist in den meisten Fällen durch rechtliche Normen in ihrer Variabilität hinsichtlich Zeit oder Vorbereitung beschränkt. Zudem ist anzumerken, dass eine pragmatische Sicht auf die unterschiedlichen Einsatzmöglichkeiten mündlicher Prüfungen dazu führt, dass man im schulischen Alltag aus organisatorischen Gründen eher auf Prüfungen ohne Vorbereitung zurückgreifen wird, während in Prüfungen, die einem Abschluss dienen, eher mündliche Prüfungen im Vordergrund stehen, die mit einer entsprechenden Vorbereitung versehen sind (auch wegen der o. g. Beschränkung durch die Normen).

Hat man sich mit mündlichen Prüfungen schon auseinandergesetzt, so wird man sich an dieser Stelle fragen, ob Präsentationsprüfungen, also Prüfungen dergestalt, dass vom Prüfling ein bestimmtes Thema vorbereitet wird, welches dann in der Prüfung präsentiert werden soll und zu dem dann anschließend von

den Prüfern Fragen gestellt werden, nicht auch zu den mündlichen Prüfungen zählen. Es soll in keinem Fall in Abrede gestellt werden, dass auch diese Art der Prüfung eine mündliche ist. Es ist aber eine Kategorie, die für dieses Buch zunächst keine große Relevanz hat. Aus diesem Grund wird sich dieses Buch damit nicht ausführlich beschäftigen, sondern sich an der obigen Definition orientieren. Weitere Möglichkeiten der mündlichen Leistungserfassung sind überblicksartig in der folgenden Abbildung zusammengetragen:

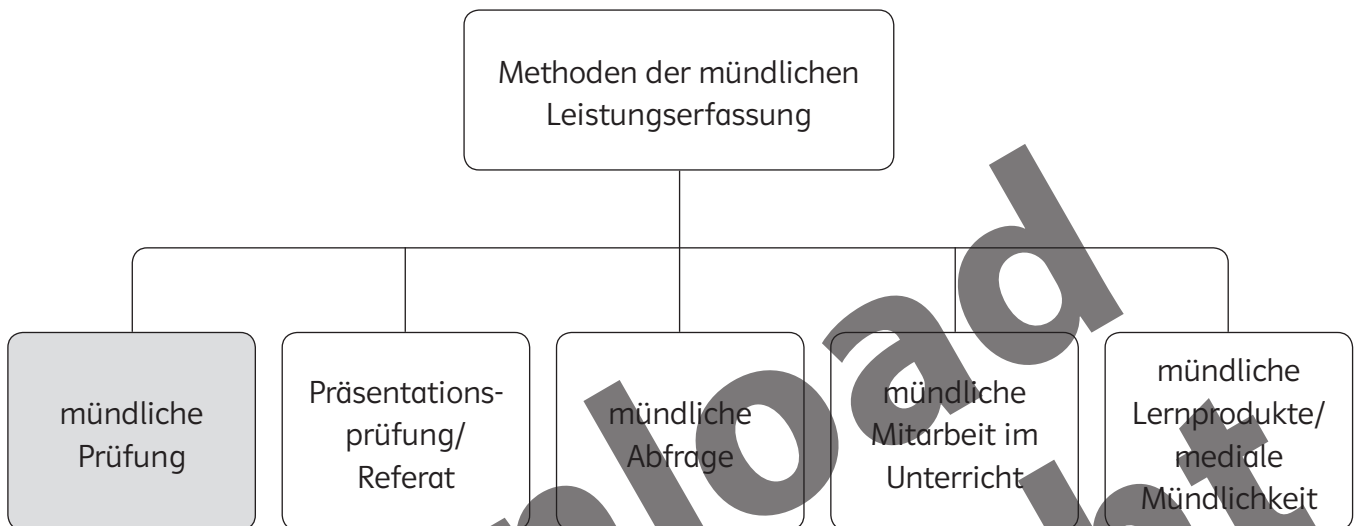


Abbildung 3: Möglichkeiten der Erfassung mündlicher Leistungen

Die Übersicht in der Abbildung 3 soll die Vielfalt der Methoden mündlicher Leistungserfassung deutlich machen und auch dazu anregen, möglichst viele dieser Methoden auszuprobieren, zu versuchen und anzuwenden.

1.2 Mündliche Prüfungen – früher und heute

Im Gegensatz zur weitverbreiteten Meinung hat die schriftliche Prüfung eine weitaus kürzere Tradition als die mündliche Prüfung (Birkel 1978, S. 14). Mündliche Prüfungen, historisch gerne als Diskussion oder sokratisches Gespräch umschrieben, fanden schon in der Antike statt und wurden in den klassischen Fächern nicht ausschließlich zur Überprüfung der erreichten Lern- und Lehrziele genutzt. Es ist zu vermuten, dass schriftliche Prüfungen eingeführt wurden, um das Prüfungsverfahren zu ökonomisieren.

Die lange Tradition der mündlichen Prüfungen, die bis zur Antike zurückreicht, lässt den Schluss zu, dass dort auch überprüft werden sollte, inwieweit der Prüfling (es ist davon auszugehen, dass es zu dieser Zeit tatsächlich nur männliche Prüflinge waren) zusammenhängend, rhetorisch geschickt und überzeugend argumentieren und diskutieren konnte. Die zunehmende Verknüpfung der rhetorischen, argumentativen und personalen Fähigkeiten mit weiteren fachlichen Inhalten machte die mündliche Prüfung nunmehr immer unhandlicher im Umgang und führte wohl Anfang der 19. Jahrhunderts zu einer Verlagerung zu schriftlichen Prüfungen hin. Die bis heute implizit vorhandene große Bedeutung personaler Kompetenzen, wie Auftreten, rhetorische Fähigkeiten, Überzeugungskraft etc., resultieren sicherlich auch aus dem historisch gewachsenen Verständnis der mündlichen Prüfung.

Nicht zuletzt könnte die Unterscheidung zwischen *schriftlichen (Haupt-)Fächern* und *nichtschriftlichen (Neben-)Fächern*, welche als Bezeichnungen in den heutigen Schulen immer noch üblich sind, auf die zunehmende Fachlichkeit sowie die ökonomische Form der schriftlichen Prüfung, als die alleinig objektive, hinweisen. Man hat wichtige Fächer (Hauptfächer) dann auch als schriftliche bezeichnet und die Leistungserfassung fast ausschließlich auf schriftliche Formen beschränkt. Diese Sichtweise spiegelt sich in vielen Diskussionen zur mündlichen Leistungserfassung wider. Schriftliche Prüfungen werden häufig als objektiver stilisiert, nicht zuletzt auch aus Tradition, nicht aus einer sachlich begründbaren Position.

Mündliche Prüfungen werden heutzutage oft als ein Korrektiv zu einer schriftlichen Prüfung verstanden, in den meisten Fällen wohl eher als ein positives. Wie sonst ist es zu erklären, dass es in vielen Bundesländern bei Abschlussprüfungen möglich ist, sich nach einer missglückten schriftlichen Prüfung noch zusätzlich für eine mündliche Prüfung in dem betreffenden Fach zu melden, um die erhaltene Note zu verbessern. In einigen Ländern ist eine mündliche Prüfung sogar vorgeschrieben, wenn Vornote und Ergebnis der schriftlichen Prüfung in einer definierten Differenz zueinander liegen, eine sogenannte Abweichungsprüfung. Was uns als selbstverständlich erscheint, ist nicht wirklich hinterfragt. Denkbar könnte es in diesem Falle ja auch sein, dass man an einer zweiten schriftlichen Prüfung teilnehmen müsste, wenn die Abweichungen allzu groß sind. Die Normengeber scheinen sich wohl etwas dabei gedacht zu haben, eventuell die Erfassung weiterer Kompetenzbereiche. Genaueres kann man den meisten Prüfungsordnungen nicht entnehmen. In diesem Zusammenhang könnte auch diskutiert werden, inwieweit sich eine mündliche Prüfung als ökonomischer (einfacher durchzuführen) oder als individueller (besser auf den einzelnen Prüfling und seine Stärken und Schwächen angepasst) darstellt und ob eine solche Einschätzung implizit Einzug in die Prüfungsordnungen verschiedener Bundesländer gehalten hat.

1.3 Schulfächer und mündliche Prüfungen

Im Kanon der Schulfächer gibt es unterschiedliche Traditionen in Bezug auf mündliche Prüfungen und deren Umsetzung. Während in den modernen Fremdsprachen und zum Teil auch im Fach Deutsch mündliche Prüfungen abseits von Abschlüssen schon seit geraumer Zeit Teil des Unterrichtsalltags sind, zählen sie in den anderen Fächern – und dazu zählen sowohl die Mathematik, als auch die Naturwissenschaften – üblicherweise nicht zum Kanon der Leistungsbewertungen, die im regulären Unterricht genutzt werden. Abgesehen vielleicht von mündlichen Abfragen zu Beginn einer Stunde. Hierbei scheint es sehr von der Form dieser Abfragen abhängig zu sein, ob man diese im Kontext der oben angeführten Definition als mündliche Prüfung bezeichnen kann oder nicht.

Die unterschiedliche Häufigkeit im Vergleich der Schulfächer und die Erfahrung in den Fremdsprachen und zum Teil in Deutsch könnte unter anderem auch daran liegen, dass die eigentlichen Grundkompetenzen der mündlichen Prüfung,

- die sprachliche Ausdrucksfähigkeit,
- die rhetorische Qualität und
- die Fähigkeit zum Argumentieren,

schon seit jeher zum Kompetenzrepertoire der modernen Fremdsprachen und des Faches Deutsch zählen. Für viele andere Fächer – zumindest bis zur Etablierung der Bildungsstandards – zählten diese Fähigkeiten eher nicht zu den grundlegend zu vermittelnden Fähigkeiten. Die Bildungsstandards im Fach Physik für den Mittleren Schulabschluss vom 16.12.2004 betonen die fachliche Kommunikation als Kompetenz. Ein lohnendes Argument also, auch in den nichtschriftlichen Fächern, wie Physik, auf einen hohen Stellenwert der Mündlichkeit zu achten und sich für mündliche Prüfungen zu interessieren.

2. Gütekriterien der Leistungserfassung

Gütekriterien, die in der Testtheorie eine Rolle spielen, also Objektivität, Reliabilität und Validität, nehmen spätestens seit PISA in der öffentlichen, aber vor allem auch in der akademischen Diskussion über Schule und die Leistungsbewertung an den Schulen einen breiten Raum ein. Das ist auf der einen Seite sicherlich von Vorteil, erhöht doch diese Diskussion die Chance auf eine faire und transparente Leistungsbeurteilung von Schülern. Auf der anderen Seite allerdings verführt der stete Verweis auf die Gütekriterien dazu, dass diese als absolutes Maß in den Blick genommen werden. Das ist sicherlich nicht bei allen Leistungsbeurteilungen möglich und sinnvoll.

2.1 Die Gütekriterien – Zusammenhänge

Wir kommen an dieser Stelle nicht um eine kurze Definition und Einordnung der Gütekriterien herum. Diese Definition soll einen praxistauglichen Charakter haben.

Eine Leistungserfassung ist dann objektiv, wenn sie in dreierlei Hinsicht Voraussetzungen erfüllt oder versucht zu erfüllen: Die Rahmenbedingungen der Prüfung müssen eine unabhängige und neutrale Leistungserfassung ermöglichen. Sie soll inhaltlich so gestaltet sein (und das betrifft auch die Bewertungskriterien), dass eine objektive Beurteilung, bzw. Bewertung der Leistung ermöglicht wird. Die Interpretation der Ergebnisse, und das bedeutet in unserem Fall vor allem die Rückschlüsse auf vorhandene Kompetenzen, muss unabhängig von den bewertenden Personen sein.

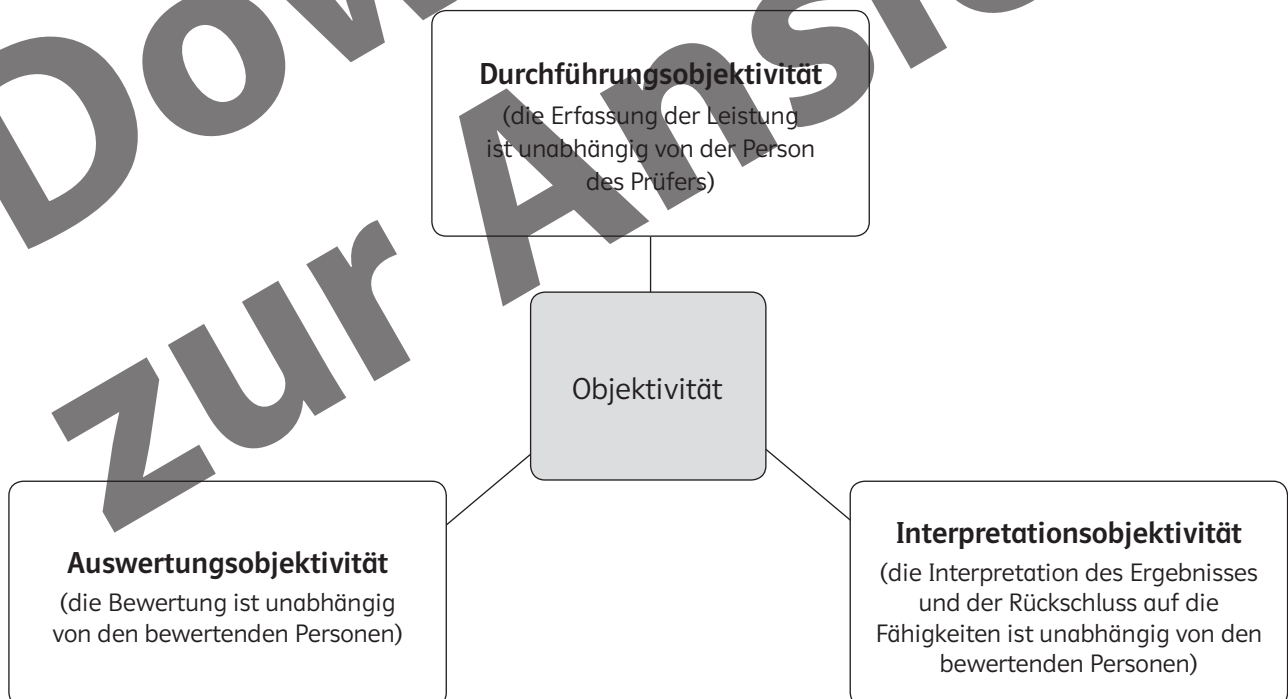


Abbildung 4: Objektivitätskategorien in der mündlichen Prüfung

Während die Notwendigkeit und die Einhaltung der Durchführungsobjektivität normalerweise sofort einleuchtend erscheint, hängt eine hohe Auswertungsobjektivität und eine hohe Interpretationsobjektivität insbesondere von den gestellten Fragen ab (sind diese überhaupt objektivierbar?) und vom auf die Fragen

bezogenen Erwartungshorizont. Hierbei ist implizit enthalten, dass gleiche Antworten auch gleich interpretiert werden müssen, also interpretationsobjektiv sein müssen.

Reliabel ist eine Leistungserfassung dann, wenn sie gewährleisten kann, dass sie zu zwei unterschiedlichen Zeitpunkten die gleiche Leistung erfassen kann, mithin zur gleichen Bewertung führt. Die Reliabilität ist von der Objektivität abhängig, wenn eine Prüfung nicht objektiv ist, dann kann sie auch nicht reliabel sein.

Die Validität als Kriterium ist von den beiden vorhergehenden abhängig, ohne dass eine Prüfung reliabel ist, kann sie auch nicht valide sein. Das Kriterium Validität ist für eine mündliche Prüfung im Fachunterricht dementsprechend stark fachlich gebunden. Obwohl es als selbstverständlich erscheint, ist es schwierig, umzusetzen und wird aus diesen Gründen auch am häufigsten in der schulischen Praxis unbewusst missachtet. Valide ist eine Leistungserfassung genau dann, wenn sie das abprüft, was sie zu prüfen vorgibt.

Die scheinbare Validität vieler schriftlicher und mündlicher Prüfungen hält einer genaueren Prüfung meist nicht stand. Dies soll hier kurz an einem Beispiel aus der Mittelstufenphysik deutlich gemacht werden: In Schulbüchern oder Arbeitsheften findet man zum Themenbereich Zusammenhang zwischen Spannung, Stromstärke und Widerstand (Stichwort: Ohm'sches Gesetz) kurze Aufgaben oder Tabellen, in denen bestimmte Größen gegeben sind und andere wiederum gesucht sind. Meistens soll der Prüfling mithilfe einer Formel, die er vielleicht auswendig kann oder die zuvor im Unterricht behandelt wurde, die gesuchten Größen berechnen (siehe Abbildung 5).

Berechne die fehlenden Größen:

Spannung	Stromstärke	Widerstand
	2 A	10 Ω
230 V		2 k Ω
7,3 V	0,1 A	

Abbildung 5: Beispielaufgabe zur Validität

Unabhängig von der didaktischen Qualität dieser Aufgabe könnte man jetzt behaupten, die Aufgabe wäre in der Lage, zu messen bzw. festzustellen, ob ein Prüfling in der Lage ist, mit den drei Grundgrößen der Elektrizitätslehre umzugehen. Dieses Ziel verfolgt diese Aufgabe jedoch nicht wirklich: Die Aufgabe dient zur Messung, ob es jemandem gelingen kann, aus einer vorgegebenen Formel und Werten für die einzelnen Variablen der Formel mithilfe von Gleichungsumformungen den Wert einer fehlenden Variablen zu berechnen.

Es ist also eine Aufgabe, die zunächst einmal mathematische Fähigkeiten erforderlich macht und mit der Elektrizitätslehre lediglich durch die Einkleidung (oder Anwendung) und mögliche Größenordnungen zu tun hat. Natürlich sind solche Berechnungen durchaus wichtig für den Umgang mit Größen der Elektrizitätslehre. Abgeprüft wird aber an dieser Stelle allerdings nur dieses und das Verständnis der physikalischen Größen bleibt außen vor.

Anders wäre es, wenn man Messwerte beispielsweise mit einer Abbildung eines Versuchsaufbaus versehen würde und ohne weitere Angaben das Ablesen bzw. die weitere Berechnung der einzelnen Größe fordert. Alternativ könnte auch über ein Diagramm zur Versuchsauswertung ein verständnisvoller Zusammenhang hergestellt werden.

In Abbildung 6 sollen die Zusammenhänge der drei Gütekriterien noch einmal deutlich werden. Der Pfeil bedeutet in diesem Zusammenhang „ist notwendige Bedingung“, also: ohne Objektivität keine Reliabilität und ohne Reliabilität keine Validität in der Leistungserfassung.



Abbildung 6: Die Abhängigkeit der Gütekriterien (Pfeil bedeutet „ist notwendige Bedingung“)

Bevor im nächsten Abschnitt die Einhaltung dieser Kriterien bei mündlichen Prüfungen diskutiert werden soll, ist in Bezug auf die drei Gütekriterien Folgendes festzuhalten: Es ist fast unmöglich, eine in allen Belangen hohe Güte in Bezug auf diese drei Kriterien in der schulischen Praxis zu erreichen. Schulische Möglichkeiten der Leistungserfassung werden tendenziell immer subjektiv, nicht immer reliabel und in Teilen nicht valide bleiben, aber es ist erforderlich, sich diese Kriterien immer wieder vor Augen zu führen und zu versuchen, eine möglichst hohe Güte anzustreben, insbesondere was das erste und das letzte Kriterium betrifft. Das Anstreben dieser Erfassungs- und Beurteilungsqualität sichert die Hinwendung zu einer fairen, gerechten und vor allem leistungsabhängigen Bewertung.

„Je größer die Beurteilungsschwierigkeiten sind und je stärker zwar formal, aber nicht inhaltlich nach dem Leistungsprinzip zugeteilt wird, desto größer wird der Anteil nicht-leistungsthematischer Zuteilungskriterien. Es werden sogenannte ‚extrafunktionale Rollenelemente‘ wie passive Anpassungsbereitschaft und Loyalität gegenüber Organisationen mit sozialer Macht als Zuteilungskriterien verwendet“ (Kleber, Meister, Schwarzer 1982, S. 67).

Heruntergebrochen auf die alltägliche Leistungserfassung bedeutet dies eine wachsende Ungerechtigkeit, sobald die Kriterien nicht in einer hohen Güte angestrebt werden.

2.2 Gütekriterien und mündliche Prüfungen

Angewandt auf mündliche Prüfungen erscheinen einzelne Aspekte der Gütekriterien eventuell in unerreichbarer Ferne zu rücken. Im folgenden Abschnitt sollen die Kriterien noch einmal in Bezug auf mündliche Prüfungen betrachtet werden.

Bei der Beantwortung dieser Frage spielen vor allem die Unterschiede zwischen schriftlichen und mündlichen Prüfungen eine Rolle. Es werden in jeder Prüfungsform, ohne dass man es möchte, immer wieder Faktoren bei der Bewertung der Prüfungsleistung eine Rolle spielen, die nichts mit der eigentlichen fachlichen Leistung des Prüflings zu tun haben. Doch welche Charakteristika machen es bei einer mündlichen Prüfung schwierig, das Ziel der Einhaltung der Kriterien zu erreichen? Dazu soll folgende Übersichtstabelle dienen.

schriftliche Prüfung	mündliche Prüfung
<ul style="list-style-type: none"> • manifeste Äußerungen • fest vorgegebene Fragen- und Auftragsstruktur • Beurteilung und Bewertung erfolgt nach dem Prüfungs-(Erfassungs)vorgang • Beeinträchtigung durch personale Faktoren eher schwierig (z. B. Handschriften) 	<ul style="list-style-type: none"> • flüchtige Äußerungen • teilweise spontane Fragen- und Auftragsstruktur • Beurteilung und Bewertung erfolgt (teilweise) während des Prüfungs-(Erfassungs)vorganges • Beeinträchtigung durch personale Faktoren eher einfach (z. B. Auftreten)

Abbildung 7: Unterschiede zwischen schriftlicher und mündlicher Prüfung

2.2.1 Objektivität

Hinsichtlich der Beurteilung der Objektivität gilt es zwischen zwei Ebenen der mündlichen Prüfung zu unterscheiden, dem Verlauf der Prüfung und der Leistungsbewertung am Ende der Prüfung. Einige Faktoren beeinflussen beide Ebenen und die drei im ersten Abschnitt genannten Kategorien der Objektivität: Durchführungsobjektivität, Auswertungsobjektivität und Interpretationsobjektivität. Die meisten der gleich genannten Problemfelder sind nicht disjunkt in eine der zwei Ebenen bzw. drei Kategorien einzuordnen, sondern betreffen meist mehrere gleichzeitig. Man kann sich die Zusammenhänge an der Matrix in Abbildung 8 deutlich machen.

	Durchführungsobjektivität	Auswertungsobjektivität	Interpretationsobjektivität
Verlauf der Prüfung	Kommunikation		
Beurteilung der Prüfung	Ungesagtes Flüchtigkeit	Verallgemeinerungen	Beurteilung

Abbildung 8: Einordnung der Handlungsfelder in die Zusammenhänge bezüglich der Objektivität

So legt zum Beispiel die eigene Erfahrung nahe, dass die Persönlichkeit des Prüflings einen maßgeblichen Einfluss auf die Beurteilung und die Bewertung einer mündlichen Prüfung ausübt. Nicht umsonst sind die Regale der Buchhandlungen voll von Literatur zum richtigen und geeigneten Auftreten in einer mündlichen Prüfung. Dennoch bleibt auch festzuhalten, dass der Prüfling die Durchführungsobjektivität beeinflussen kann, indem er z. B. durch sein Auftreten der Prüfung einen anderen Verlauf geben kann, als ursprünglich geplant war. Man sieht schon an diesem Beispiel, dass die Trennung häufig nicht so einfach ist. Dass man vielmehr aber auch während einer Prüfung darauf achten muss, dass einem die Kriterien für eine objektive Durchführung nicht abhanden kommen. Die Vielfalt der möglichen Störungen ist groß. Die wichtigsten Handlungsfelder hinsichtlich der Objektivität sollen hier in mehreren kurzen Abschnitten aufgelistet werden.

Kommunikation

Eine mündliche Prüfung ist immer eine asymmetrische Kommunikationssituation, die über eigene Regeln und Rituale verfügt (Birkel 1978, S. 28). Dieser Begriff der asymmetrischen Kommunikation, der insbesondere durch Watzlawick geprägt wurde, soll hier kurz erläutert werden, weil er für viele mündliche Prüfungsgespräche prägend und durchaus geeignet ist, eine reflektierte Herangehensweise an die eigene Prüfungspraxis zu ermöglichen.

In seinen fünf Axiomen der Kommunikation beschreibt Watzlawick (Watzlawick, Beavin, Jackson 1982, S. 68 ff.) im fünften Axiom die symmetrische und die komplementäre Interaktion. „Symmetrische Beziehungen zeichnen sich also durch Streben nach Gleichheit und Verminderung von Unterschieden zwischen Partnern aus, während komplementäre Interaktionen auf sich gegenseitig ergänzende Unterschiedlichkeiten basieren“ (Watzlawick, Beavin, Jackson 1982, S. 69).

Auch ohne große Kenntnis der Kommunikationstheorie ist hier offensichtlich, dass es sich bei der Kommunikation zwischen Prüfendem und Prüfling in einer mündlichen Prüfung um die Kommunikation in einer komplementären Beziehung handeln muss. Diese besteht normalerweise auch schon länger und entsteht nicht erst am Datum der Prüfung. Durch diese, gesellschaftlich und institutionell, festgelegte Rollenverteilung ergeben sich im Prüfungsgespräch, also während der Kommunikation, aber auch schon zuvor oder danach Situationen bzw. Konflikte, die eine Objektivität der Prüfung beeinflussen können.

Als Beispiel sei an dieser Stelle genannt, dass es theoretisch dem Prüfer obliegt, in der Prüfung ein Thema zu wechseln, einen neuen Punkt aufzurufen oder in einem Thema eine Spezialisierung anzustreben. Dass kann er wiederum aber nur tun, ohne häufig konkret zu wissen, an welchem Punkt seiner Kompetenzdarlegung der Prüfling gerade ist. Anders gesagt: Eventuell wird das Thema bereits gewechselt, während der Prüfling noch zu den wichtigsten Punkten im Gespräch kommen wollte und es gelingt ihm an dieser Stelle nicht mehr, sein Wissen und Können auszubreiten.

Diese besondere Erwähnung der vorhandenen asymmetrischen Kommunikationssituation in einer mündlichen Prüfung soll nun nicht dazu dienen, zu versuchen, dieser Situation Herr zu werden, oder sie eventuell symmetrischer zu gestalten. Das wird nicht gelingen, ist diese Asymmetrie doch institutionell zementiert. Es soll aber, und das gilt für alle hier aufgeführten Handlungsfelder, Anlass zur Selbstkritik und zur Reflexion geben, um Faktoren, die die Objektivität beeinflussen, zumindest in den Blick zu nehmen und kritisch zu betrachten, auch wenn man sie nicht vollkommen ausschalten kann.

Flüchtigkeit

Eine weitere Gefahr für die Objektivität, welche auch die Beziehung zwischen der Prüfung und ihrer Bewertung deutlich macht, stellt die Flüchtigkeit der mündlichen Prüfung dar. Mündliches wird normalerweise nicht festgehalten, sondern ist nach der Wahrnehmung und der Beobachtung zunächst einmal nicht reproduzierbar, es sei denn, man zeichnet das Gespräch auf. Hier, wie in vielen andern Fällen soll noch einmal deutlich werden, dass es sich um ein Buch für die Praxis in der schulischen Realität handelt. Natürlich kann man Prüfungen, eventuell zu wissenschaftlichen Untersuchungszwecken, aufzeichnen, dennoch ist es nicht ergiebig, das an dieser Stelle zu diskutieren, denn niemand wird das in der Schulpraxis tun, weil der Aufwand in keinem Verhältnis zum gewünschten Ertrag steht. Diese Flüchtigkeit der mündlichen Kommunikation hat mittelfristig einen Einfluss auf die Bewertung, aber auch kurzfristig auf den Prüfungsverlauf und sollte demnach bedacht werden. Auf die Objektivität entsteht insoweit ein Effekt, als das der intersubjektive Analyseprozess dessen, was an Leistung bzw. an Kompetenzen gezeigt wurde, wesentlich schwieriger ist als in einer schriftlichen Arbeit.

Verallgemeinerungen

Es besteht die Gefahr der Verallgemeinerung von Eindrücken. Diese Unsicherheit ist auch davon abhängig, wie gut Prüfer und Prüfling miteinander bekannt sind und welche Vorerfahrungen, insbesondere bei den Prüfern, einem Eindruck, der in irgendeiner Art und Weise vom Prüfling hinterlassen wird, zugrunde liegt. Dasselbe gilt übrigens auch in umgekehrter Richtung, inwieweit also die Vorerfahrungen des Prüflings mit dem Prüfer einen wesentlichen Einfluss auf dessen Verhalten in der Prüfung ausüben können.

Ungesagtes und Implizites

Eine weitere Möglichkeit, Objektivität zu gefährden, entsteht durch in der Prüfung nicht wahrnehmbare oder nur beschränkt wahrnehmbare Dinge, also Implizites. Dies betrifft zum einen inhaltliche Fragen, die z. B. durch einen sogenannten „didaktischen Vertrag“ zwischen Prüfendem und Prüfling irgendwann einmal beschlossen wurden. Dieser Vertrag beinhaltet beiderseitige Interpretationen von Fragen bzw. Antworten, die naturgemäß vom vorhergehenden Unterricht und der bisherigen Kommunikation miteinander abhängig

sind. Zum anderen können aber auch nonverbale Haltungen, Gesten oder auch Äußerungen zu einer Verunsicherung auf beiden Seiten führen.

Die Unfähigkeit zur Nichtkommunikation betrifft auch Rituale, die – von wem auch immer vorgegeben – sozusagen als allgemeine Verhaltensrituale in einer Prüfung einer Schule oder auch eines Landes implizit vorgegeben sind und deren Begründungen nicht mehr hinterfragt werden, die aber auch aus diesem Grunde nicht mehr transparent gemacht werden. Bourdieu (2001, S. 123 f.) spricht in solchen Fällen vom gemeinen Menschenverstand: „Der gemeine Menschenverstand ist ein Fonds von allen geteilter Überzeugungen, der in den Grenzen des jeweiligen sozialen Universums eine grundlegende Übereinstimmung über den Sinn der Welt und einen Bestand von (stillschweigend akzeptierten) Gemeinplätzen sichert [...]“. Er beschreibt damit die Problematik, die durch eine ritualisierte Tradition auf vielen schulischen Geschehen lastet und aus der es schwierig ist zu entkommen. Umso wichtiger ist eine beständige Reflexion dieser Problematik, um eine zwar nie vollständig herstellbare aber dennoch immer zu verbessernde Objektivität zum Wohle aller herzustellen.

Das bedeutet auf der praktischen Seite stets im Auge zu behalten, wie kommuniziert wird und ob die beobachtete Kommunikation bewusster Teil des Prüfungsgeschehen ist oder nur Teil einer tradierten Kommunikationsstruktur.

Beurteilung

Bei der Beurteilung ist es entscheidend, darauf zu achten, dass die Person des Prüfers keinen Einfluss auf die Beurteilung hat. Hieraus kann ein Plädoyer für die Erarbeitung sicherer und transparenter Beurteilungskriterien abgeleitet werden. Dies trifft auch auf die Interpretationsobjektivität zu, mit der sichergestellt werden muss, dass die Indikatoren eindeutig validen Merkmalen zuzuordnen sind. In der Bewertung wiederum muss die Zuordnung zu den Notenstufen nachvollziehbar sein. Das bedeutet allerdings nicht, dass ein Algorithmus zugrunde gelegt wird, vielmehr muss deutlich werden, was z. B. eine gute oder eine sehr gute Leistung ist und welche Kriterien dafür angelegt werden. Ein Problem ist hierbei oft die große Bandbreite bei Beurteilungen im mittleren Bereich. Eine Thematik, die Lehrkräften auch von anderen mündlichen Leistungen bekannt ist und die auch nicht einfach aufzulösen sein wird.

Hingewiesen sei in diesem Zusammenhang auf die in der Notengebung übliche Tendenz zur Mitte, mitunter auch als Milde-Härte-Fehler oder Vermeidungseffekt bezeichnet, der dafür Sorge trägt, dass sich gerade bei der Beurteilung mündlicher Leistungen viele Benotungen im mittleren Bereich bewegen. Ein Grund mehr für klare Kriterien zur Bewertung und beobachtbare Indikatoren in diesem Zusammenhang. Erwähnt werden soll auch die Problematik, dass der Unterschied zwischen zwei Notenpunkten (z. B. im 15-Punkte-System) nicht als gleich interpretiert wird. So sehen viele Lehrkräfte den Abstand zwischen 10 Punkten (noch gut) und 9 Punkten (voll befriedigend) als größer an als zwischen 7 Punkten (noch befriedigend) und 8 Punkten (befriedigend). Eine Reihe von Hinweisen auf weitere allgemeine Bewertungsprobleme findet sich in Paradies, Wester, Greving (2009).

Zusammenfassend lässt sich sagen, dass sich eine hundertprozentige Objektivität hinsichtlich der Beurteilung durch verschiedene Prüfer wohl nicht zuverlässig erzeugen lässt, auch nicht bei einer optimalen Vorbereitung. Tröstlich ist an dieser Stelle, dass dies auch bei schriftlichen Prüfungen nicht der Fall ist (vgl. Ingenkamp 1976; Birkel 2005). Prinzipiell gilt es, dieses Problem zu berücksichtigen und alle Vorbereitungen daraufhin auszurichten, eine möglichst große Objektivität herzustellen. Im späteren Kapitel, wenn es um praktische Hinweise zur Durchführung der Prüfung und um eine gute Beurteilung und Bewertung gehen soll, wird dieser Aspekt nochmal aufgegriffen.

2.2.2 Reliabilität

Die Reliabilität einer mündlichen Prüfung zu gewährleisten, ist schwieriger: Unter der Reliabilität im Sinne eines Gütekriteriums für eine mündliche Prüfung versteht man, dass die mündliche Prüfung zu zwei verschiedenen Zeitpunkten die (angenommene identische) Leistungsfähigkeit eines Prüflings zu messen in der Lage ist (und auch in der Bewertung eben zu diesem identischen Urteil gelangt). Man erkennt sogleich, dass dieses Kriterium aufgrund des singulären und situativen Charakters einer mündlichen Prüfung noch schwieriger einzuhalten sein wird als die Objektivität. Zumal nicht klar ist, ob die Performanz, also das Zeigen verschiedener fachlicher Kompetenzen, im mündlichen aber auch im schriftlichen Bereich immer gleich und von konstanter Qualität ist, auch wenn die Leistungsfähigkeit sich nicht verändert hat. „Die Reliabilität bezieht sich also auf die Präzision, mit der in der mündlichen Prüfung die Leistungsfähigkeit eines Prüflings gemessen wird, unabhängig davon, ob das, was gemessen und beurteilt wird, auch tatsächlich die Leistungsfähigkeit des Prüflings ist. Das muss im Zusammenhang mit der Frage nach der Validität der mündlichen Prüfung diskutiert werden“ (Birkel 1978, S. 68). Die Reliabilität ist also eine notwendige Voraussetzung der Validität einer mündlichen Prüfung (wie auch anderer Prüfungen) und sollte unter dem Aspekt in den Blick genommen werden, dass man mithilfe einer hohen Validität auch eine erträgliche (nicht zu verifizierende) Reliabilität herstellen kann.

Untersuchungen zur Reliabilität mündlicher Prüfungen sind schwierig und immer einer Reihe von Nebenbedingungen unterworfen (Erinnerungseffekt, motivationale Probleme etc.), die eine echte Beurteilung erschweren. Vergleichende Untersuchungen nutzen daher auch oft Audio- oder Videoaufzeichnungen, die dann innerhalb bestimmter Zeiträume erneut beurteilt werden. Das beleuchtet allerdings nur eine Seite des Prüfungsprozesses, nämlich die Beurteilung, nicht aber den Prüfungsvorgang selbst. Dennoch sollte man sich besonders, wenn die mündliche Prüfung keine Abschlussprüfung ist, des Gütekriteriums bewusst sein und sich klarmachen, dass dieses Kriterium zumindest implizit ein Garant für eine gerechte Beurteilung darstellt.

2.2.3 Validität

Validität ist ein Gütekriterium, das jedem Prüfenden auch in einer mündlichen Prüfung sofort als notwendig einleuchtet. Valide ist eine Prüfung genau dann, wenn sie auch misst, was sie vorgibt, zu messen. Der einfachste Validitätsbereich scheint zunächst die inhaltliche Validität zu sein, also die Beobachtung, ob die genutzten Fragen oder Impulse auch das abdecken, was an Inhalten geprüft werden soll. Hierbei gilt es aber auch zu unterscheiden, inwieweit das zu Prüfende reines Wissen ist oder in Kombination mit Können auch als Performanz einer Kompetenz beobachtbar sein soll. Dieser Frage wird später noch einmal fachspezifisch nachgegangen.

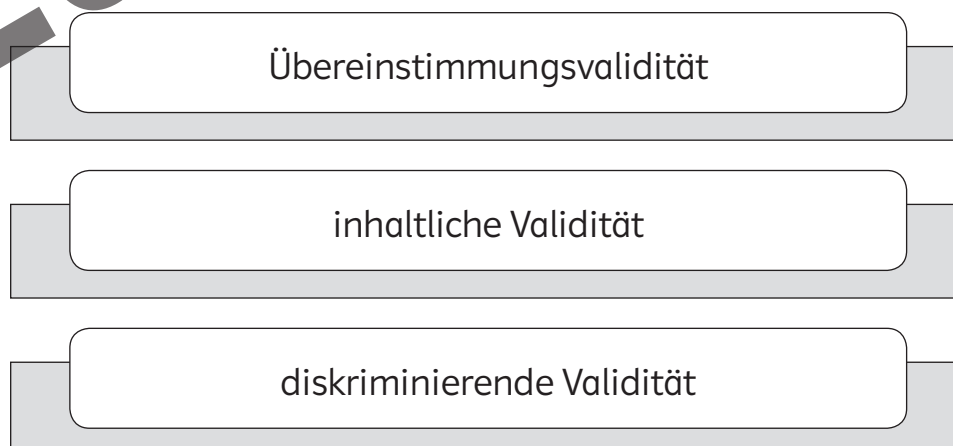


Abbildung 9: Validitätskategorien in einer Prüfung

Die Validität ist wie die anderen Gütekriterien von vielen Faktoren abhängig. Neben dem Prüfling und dem Prüfenden sind es vor allem die Fragen bzw. Impulse, die die *inhaltliche Validität* beeinflussen. Zu dem letzten Punkt muss in Bezug auf die Physik genau herausgearbeitet werden, welche Fragen für eine mündliche Prüfung in besonderem Maße geeignet sind. Dabei ist insbesondere der Einsatz von Operatoren empfehlenswert. Diese liefern die Fragestellung betreffend möglichst große Transparenz. Darüber hinaus gilt es, ebenfalls im Zusammenhang mit der Validität, die Frage zu beantworten, inwieweit ersichtlich ist, dass die in der mündlichen Prüfung erzielte Beurteilung der fachlichen Leistung mit den anderen Beurteilungen in diesem Fach korreliert: die sogenannte *Übereinstimmungsvalidität*. Die *diskriminierende Validität* wird oftmals vernachlässigt, ist aber ebenfalls von Relevanz in einer Prüfung. Durch sie wird gewährleistet, dass die Fragen und Arbeitsaufträge aufgrund ihrer inhaltlichen Beschaffenheit auch geeignet sind, zwischen Prüflingen mit guter oder schlechter Leistung zu unterscheiden.

Prüfling und Prüfender sind Faktoren, die nicht nur im Fach Physik einen hohen Einfluss auf die Validität einer mündlichen Prüfung ausüben können. Das liegt vor allem daran, dass in einer mündlichen Prüfung eben durch spontane Entwicklungen Situationen entstehen können, in denen das zu Messende angesichts anderer Phänomene in den Hintergrund gerät. Das soll an folgendem Beispiel deutlich werden: In einer mündlichen Prüfung ist davon auszugehen, dass man ein besonderes Interesse daran hat, kommunikative Kompetenzen zu messen, die mit dem Fach verbunden sind. Wenn nun aber hinsichtlich dieser kommunikativen Kompetenzen ein Übergewicht in die Richtung der personalen kommunikativen Kompetenzen entsteht, besteht die Gefahr, dass sich ein Verlust von Validität entwickelt, weil die fachlichen Kompetenzen durch die personalen überlagert werden.

Es wird deutlich, dass Validität angesichts der Diskussion, die hier für das Fach Physik geführt wird, ein eher schwieriges Gütekriterium werden kann. Wir gehen davon aus, dass es uns gelingt, in einer mündlichen Prüfung Kompetenzbereiche beurteilen zu können, die in anderen Beurteilungsformen nicht beurteilt werden können.

2.2.4 Zusammenfassung

Angesichts der vielen offenen Fragen, die in Bezug auf die Gütekriterien und auch der Qualität mündlicher Prüfungen hier erläutert worden sind, erscheint es schon bemerkenswert, dass es nicht eine Reihe von Untersuchungen zu diesem Thema gibt. Bedauerlicherweise ist das nicht der Fall, sondern viele Erkenntnisse sind mittlerweile schon weit mehr als 30 Jahre alt und sollten in Zukunft, auch angesichts veränderter schulischer Diskussionen, z. B. zum Thema Kompetenzerwerb und Leistungskultur, einer dringenden Revision unterzogen werden (Birkel 1978, S. 206 f.). Birkel, der im Rahmen einer Arbeit mündliche Abiturprüfungen im Fach Deutsch untersucht hat, ist in seiner Analyse mündlicher Abschlussprüfungen, insbesondere an Gymnasien, zu einer eher skeptischen Sicht auf mündliche Prüfungen gelangt. Allerdings ist sicherlich nicht jede von ihm getroffene Feststellung ohne weiteres auf andere Klassenstufen oder auch auf andere Fächer übertragbar. Allerdings sind aber Hinweise übertragbar, die eine Objektivierung mündlicher Prüfungen aufgrund der von ihm gewonnenen Ergebnisse möglich machen könnten.

2.3 Als Prüfungsformat ungeeignet?

Nachdem nur ein Auszug möglicher Unwägbarkeiten in Bezug auf die Einhaltung der Gütekriterien, vor allem aber der Objektivität einer mündlichen Prüfung hier vorgestellt werden kann, drängt sich die Frage auf, ob eine mündliche Prüfung, angesichts all der bisher schon genannten Mängel als Format der Leistungserfassung überhaupt geeignet ist. Diese Frage ist bewusst gestellt, basiert sie doch auf der Annahme, dass die bisherigen Möglichkeiten der Leistungserfassung die Gütekriterien in übermäßiger Form erfüllten.

In einem pragmatischen Ansatz hingegen muss zwischen den neuen Ideen und den bisher existierenden abgewogen werden, in diesem Falle unter der Prämisse, ob die Nachteile, die durch eine mündliche Prüfung als Leistungserfassungsinstrument entstehen, eventuell durch Vorteile auf einer anderen Seite aufgewogen werden.

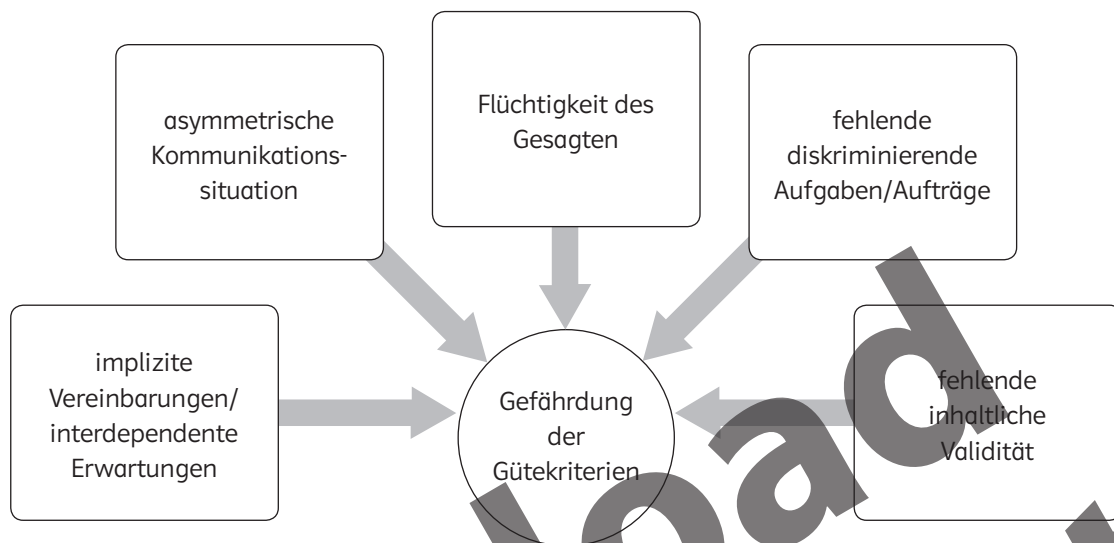


Abbildung 10: Gefährdung der Gütekriterien mündlicher Prüfungen

Die Frage, warum es eigentlich mündliche Prüfungen geben muss, wenn es doch auf der schriftlichen Ebene so viele Möglichkeiten der Leistungserfassung gibt, ist durchaus berechtigt und wurde auch im oben angeschnittenen historischen Kontext noch nicht zufriedenstellend beantwortet.

Die Antwort auf diese Frage muss in zwei unterschiedlichen Kapiteln stattfinden, auf einer allgemeinen Ebene, also hier und auf einer fachspezifischen Ebene, also speziell für das Fach Physik. Bedauerlicherweise existieren nur wenige Befunde, die einen Aufschluss darüber ermöglichen, inwieweit eine Erhöhung der Validität der Leistungsmessung durch eine mündliche Prüfung gewährleistet werden kann. Aber alleine durch die Struktur dieser Art der Leistungserfassung lassen sich Vorteile ableiten, die in einer schriftlichen Leistungserfassung nicht in diesem Maße existieren.

Eine gute Grundlage für Begründungen liefert immer die Betrachtung der Kritikpunkte am Vergleichsobjekt. Ist eine schriftliche Prüfung objektiver als eine mündliche? Diese Frage kann, in dieser allgemeinen Art gestellt, mit gutem Gewissen verneint werden. Denn auch bei einer schriftlichen Prüfung gibt es eine Reihe von Fallstricken hinsichtlich der Objektivität, die es zu beachten gilt. So sind auch dort Rahmenbedingungen schwierig, z. B. das korrekte und valide Formulieren von Aufgaben: In einer schriftlichen Prüfung gibt es keine Nachfragen oder keine weiteren Erläuterungen.

Nicht umsonst werden alle in der Schule angewandten Verfahren zur Leistungsfeststellung von Testtheoretikern oder Bildungsforscher eher abschätzig als „subjektive Verfahren“ bezeichnet (vgl. Heller 1984, 28 f.). Dies sollte Sie jetzt aber nicht den Glauben in Ihre schulische Leistungsbeurteilung verlieren lassen, sondern eher den Blick dafür schärfen, dass selbst Verfahren, die man bisher für objektiv gehalten hat (wie eine Klassenarbeit oder schriftliche Überprüfung), niemals in Gänze objektiv im eigentlichen Sinne sein können. Das an sich ist aber als Begründung für die mündliche Prüfung nicht ausreichend, denn etwas Neues ist nicht deswegen gut, weil es ebenso viele Fehler wie das Alte hat.

Die Begründung für diese Art der Leistungserfassung liegt auf der Hand, wenn man sich die Vielfalt der zu beherrschenden Kompetenzen in den verschiedenen Fächern anschaut. Und dabei wird deutlich, dass man mit schriftlichen Prüfungen nur einen Bruchteil dieser zu erwerbenden Kompetenzen erfassen bzw. abprüfen kann. Manche sind auch gar nicht abzuprüfen.

Mündliche Prüfungen sind also notwendig, um Kompetenzbereiche zu erfassen, die mithilfe anderer Erfassungsinstrumente nicht berücksichtigt werden können.

Was das wiederum genau für die Physik bedeutet, soll später noch einmal deutlich gemacht werden.

2.4 Allgemeine Hinweise zum Erreichen der Gütekriterien

- Die Kompetenzen und Lernziele, die überprüft werden sollen, müssen vorher festgelegt werden. Diese Forderung ist nicht identisch mit dem Festlegen der Fragen bzw. des Themas. Es ist entscheidend, sich zunächst darüber Gedanken zu machen, welche Kompetenzbereiche man in der mündlichen Prüfung in den Blick nehmen will und wo die jeweiligen Schwerpunkte liegen.
- Mündliches ist flüchtig. Aus diesem Grund muss ein Katalog erstellt werden, aus dem Kriterien und Indikatoren hervorgehen, die ein Erreichen der Ziele widerspiegeln. Diese Kriterien und Indikatoren finden sich in Bewertungsrastern wieder, die man nutzen sollte, um mündliche Prüfungen zu beurteilen und letztendlich auch zu benoten. In einem Kriterienkatalog muss festgehalten werden, welche Leistungen wie bewertet werden. Kompetenzen, Lernziele und Kriterien müssen zueinander kompatibel sein.
- Die Kriterienkataloge müssen transparent sein, also auch für die Schülerhand zugänglich. Sie müssen mit den betroffenen Schülern diskutiert werden, aber auch innerhalb der Fachgruppe an der Schule, wenn man plant, mündliche Prüfungen durchzuführen. Diese Art der kommunikativen Validierung ist ein probates Mittel, frühzeitig auf Probleme hinzuweisen.
- Prüfungen sollten von mehreren Prüfern beurteilt werden, die ihre Leistungsbewertung unabhängig voneinander festhalten. Das Ergebnis sollte mithilfe aller Bewertungen ausgelotet werden, um so verschiedene Sichtweisen in den Blick nehmen zu können (Birkel 1978, S. 205). Diese scheinbar einfache Forderung entfaltet ihre Kraft zur Objektivierung genau dann, wenn diese Bewertung vor der Diskussion schriftlich und individuell festgelegt wird. Denn schon alleine der Vorschlag einer Bewertung durch einen Prüfer direkt nach einer Prüfung birgt die Gefahr der Beeinflussung des anderen Prüfers in sich. Diese Forderung ist aber wohl aus ökonomischen Gründen nur in Abschlussprüfungen umsetzbar. In Prüfungen, in denen lediglich ein Beurteiler zur Verfügung steht, kann man den Schüler in die Beurteilung miteinbeziehen.
- Die Sprache ist von entscheidender Bedeutung in einer Prüfung, insbesondere die Fachsprache ist in der Physik von immenser Bedeutung. Das gilt auch für die verwendeten Begriffe in den Frage- oder Impulssequenzen. Hierbei muss darauf geachtet werden, dass die verwendeten Operatoren klar und deutlich sind und dass sie vom Prüfling in der gleichen Weise verstanden werden, wie vom Prüfenden. Das ist eine Forderung, die auch für schriftliche Prüfungen im besonderen Maße gültig ist. In mündlichen Prüfungen ist sie so zentral, weil viele Äußerungen nicht dezidiert geplant erfolgen. Es ist sinnvoll, sich eine Operatorliste anzulegen, die man auch in der Prüfung nutzen kann.

Trotz aller redlichen Bemühungen wird es nicht gelingen, eine mündliche Prüfung in allen Belangen objektiv und valide zu gestalten. Es gilt in jedem Fall, die Vor- und Nachteile diesbezüglich abzuwägen und sich zu fragen, ob der etwaige Verlust von Güte durch den Gewinn der Messung alternativer Kompetenzbereiche aufgewogen werden kann. So muss hingenommen werden, dass ein Fazit am Ende stehenbleibt: „Durch den interaktiven und möglicherweise adaptiven Charakter mündlicher Prüfungen ist bei diesen Leistungsfeststellungen und darauf aufbauenden Beurteilungen nicht der gleiche Grad der Objektivität wie bei schriftlichen Prüfungen zu erreichen“ (Kirk 2004, S. 39).

3. Messen, Beurteilen, Bewerten

3.1 Beobachtung und Wahrnehmung

Wahrnehmen umschreibt, was durch Sinneseindrücke zu uns dringt. Das können visuelle und olfaktorische Eindrücke, aber natürlich auch sensorische oder andere Eindrücke sein. In einer mündlichen Prüfung werden viele Dinge wahrgenommen, sowohl durch den Prüfer als auch durch den Prüfling.

Zentral ist es, die wichtigen und die unwichtigen Wahrnehmungen voneinander zu trennen, sodass nicht durch die unwichtigen Wahrnehmungen ein sogenannter Eindruck entsteht. Schaut man sich dieses Wort nämlich genauer an, so ahnt man schon, dass damit ein Urteil auf einer Detailebene umschrieben wird, das sich später schwer wieder auf einer objektiven Ebene von jener Beobachtung trennen lässt, welche für die Bewertung der Prüfung entscheidend sein dürfte.

Wahrnehmen kann man nicht verhindern, allerdings kann man Wahrnehmung in eine richtige Richtung lenken, Wichtiges also von Unwichtigem trennen, nicht nur wahrnehmen, sondern beobachten. Oft wird die Beobachtung den rein visuellen Eindrücken zugewiesen. So soll Beobachten an dieser Stelle allerdings nicht verstanden werden. Es umschreibt vielmehr jene strukturierte Wahrnehmung von Sinneseindrücken, die sich nicht nur an einer vorher festgelegten Unterscheidung von relevant und irrelevant für die Beurteilung orientiert, sondern die implizit eine diskriminierende Relevanz hinsichtlich der zu beurteilenden Kriterien beinhaltet. Heller und Hany (2002, S. 90) postulieren allgemein für schulische Leistungen: „Voraussetzung für die Messung schulischer Leistungen ist, dass diese sich in beobachtbarem Verhalten [...] manifestieren. Die Leistungssituation ist so zu gestalten, dass sich unterschiedliche Leistungspotentiale auch in unterschiedlichem Verhalten zeigen. Diese Verhaltensunterschiede bilden ein ‚empirisches Relativ‘ (eine strukturierte Sammlung von Verhaltensweisen) [...]“. Bei einer mündlichen Prüfung müssen wir es mit einer systematischen Beobachtung zu tun haben. Diese Beobachtung des Verhaltens wird bei einer mündlichen Prüfung in erheblicher Weise durch den interaktiven Charakter der Prüfung beeinflusst.



Abbildung 11: Wahrnehmen und Beobachten

Daher kann man folgern, dass der mündlichen Prüfung eine strukturierte Planung zugrunde gelegt werden muss, die neben einem Fragenplan oder Impulsen ebenfalls ein Bewertungsraster beinhaltet, in welchem man die beobachteten und relevanten Verhaltensweisen auf einfache Weise systematisch aufführt und wiederfinden kann. Das Vorhandensein eines Bewertungsrasters entbindet den Prüfer allerdings nicht davon, auch ein Protokoll zu führen, in welchem die leistungsrelevanten Merkmale der Prüfung nachvollzogen werden können.

3.2 Merkmale der Leistungserfassung in mündlichen Prüfungen

Will man die Möglichkeiten der Leistungserfassung durch eine mündliche Prüfung nutzen, muss klar sein, welche Unterschiede eine mündliche Prüfung im Gegensatz zu einer schriftlichen Prüfungsform hinsichtlich Beurteilung und Bewertung aufweist. Damit kann es gelingen, einzelne Kompetenzbereiche der Lernenden herauszuarbeiten bzw. verfälschende Elemente einer Prüfung auszuschließen.

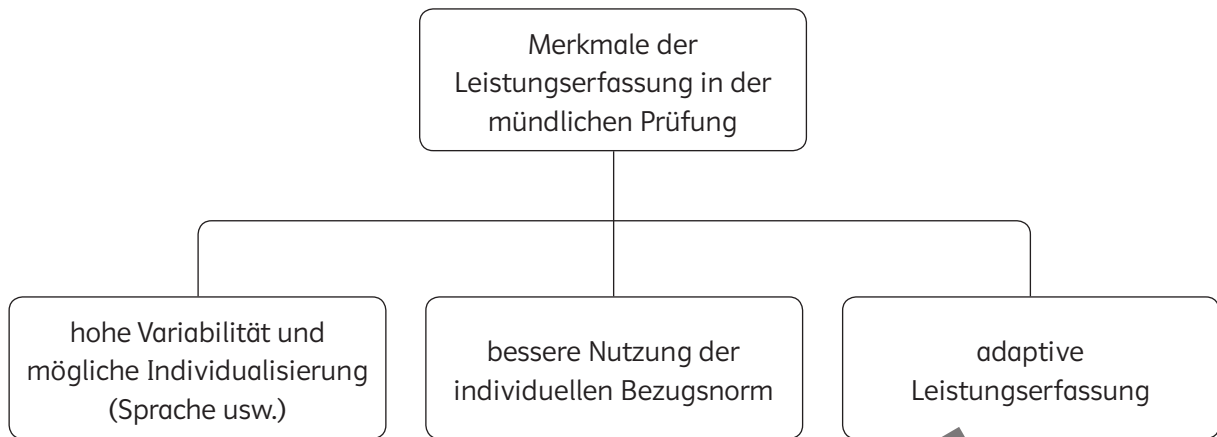


Abbildung 12: Leistungserfassung in mündlichen Prüfungen

Durch den temporären aber auch den variablen Charakter ist die Möglichkeit einer Individualisierung in einer mündlichen Prüfung sicherlich stärker gegeben als in einer schriftlichen Prüfung. Der Prüfende kann in bestimmten Punkten nachfragen, nachhaken oder alternative Formulierungen wählen. Dies soll angesichts der mitunter nicht immer gelungenen Sprache in manchen schriftlichen Prüfungen als ein Vorteil herausgestellt werden.

In der Literatur wird vor allem aber auch die bessere Möglichkeit einer Betonung der individuellen Bezugsnorm durch die Beurteilung mündlicher Leistungen, insbesondere auch mündlicher Prüfungen herausgestellt (Kirk 2004, S. 42 ff.). Da die mündliche Prüfung eine individualisierte Prüfung ist, ist es möglich, im Rückgriff auf bereits erbrachte Leistungen, auf individuelle Stärken oder Schwächen des Schülers einzugehen.

Dementsprechend besitzen mündliche Prüfungen in einem hohen Maße die Möglichkeit adaptiv Leistung zu messen, d. h., eine Leistungsmessung zu ermöglichen, die auf die Fähigkeiten und Fertigkeiten des Prüflings zugeschnitten wird, weil sie einen singulären Charakter besitzt (vgl. hierzu Jäger 2001, S. 201 f.).

Nicht unterschätzt werden sollte im Hinblick auf die Variabilität in einer mündlichen Prüfung die Möglichkeit, die in den schriftlichen Prüfungen verwendete oft eindeutige Bildungssprache zugunsten einer eher besser zu verstehenden Alltagssprache abzuschwächen und damit bestimmten Schülern einen besseren Zugang zu den erfragten Inhalten und Kompetenzen zu gewähren. Die Rolle von Bildungs- und Fachsprache speziell im Physikunterricht und dort in Prüfungen ist mittlerweile unstrittig und schlägt sich natürlich insbesondere dann nieder, wenn kommunikative Kompetenzbereiche in den Fokus geraten. Diese Rolle wird im zweiten Theorieteil des Buches noch detaillierter beleuchtet.

3.3 Wie wird in mündlichen Prüfungen gemessen?

Wie bei allen Leistungsbewertungen ist der Dreischritt



Abbildung 13: Von der Beobachtung zur Bewertung

die Leitlinie der Leistungsbeurteilung bzw. -bewertung. Das Besondere ist hierbei, dass über den Erwartungshorizont hinaus, der bei schriftlichen Prüfungen meist inhaltlich-fachlich hergestellt wird, durch die Beobachtung und den impliziten Abgleich mit Kriterien und Indikatoren, also dem Beschreiben des Beobachteten, eine Beurteilung und zuletzt auch eine Bewertung entsteht.

Das macht die Leistungsmessung in mündlichen Prüfungen sicherlich schwieriger als in schriftlichen Prüfungen, nicht aber unbedingt aufwendiger, denn einmal eingesetzte Bewertungsraster können einfach auf andere mündliche Prüfungen angepasst und so immer wieder genutzt werden.

Für die Raster empfiehlt es sich, immer gleichbleibende Beschreibungen zu wählen, um bei häufigeren Durchführungen in der Bewertung auch zu einer gewissen Routine zu gelangen.

Der Unterschied zwischen schriftlichen und mündlichen Prüfungen lässt sich am besten an diesem Schaubild verdeutlichen:



Abbildung 14: Unterschied zwischen schriftlicher und mündlicher Prüfung

Für mündliche Prüfungen gilt, wie für alle anderen Leistungsmessungen auch, dass eine hohe Abhängigkeit zwischen den zugrundeliegenden Bezugssystemen und den darauf fußenden Bewertungskriterien besteht. Prinzipiell soll zwischen folgenden Bezugssystemen unterschieden werden:

- a) Die Lehrkraft kann die Güte der Leistung an einem nicht explizit definierten Kategoriensystem messen. Es liegt dann ein subjektives Bezugssystem vor, das auch mit dem Begriff des absoluten Urteils umschrieben werden kann. Es handelt sich um eine Art des Expertenurteils.
- b) Die Lehrkraft kann die derzeitigen Leistungen eines Schülers an dessen eigenen früheren Leistungen messen: individuelles Bezugssystem.
- c) Die Lehrkraft kann die Leistungen der von ihm geführten Klasse als Bezugssystem verwenden: klasseninternes Bezugssystem. Bei Ausweitung auf eine Schule erhält man das schulinterne Bezugssystem.
- d) Die Lehrkraft kann für bestimmte Gruppen normierte Messverfahren verwenden. Sie misst dann die Leistungen eines Schülers an den Normtabellen für solche Gruppen: Normbezug.
- e) Die Lehrkraft verwendet ein operationalisiertes Kriterium, an dem die Leistung des einzelnen Lernenden gemessen wird: Kriteriumsbezug.

(vgl. Kleber 1982, S. 72 f.)

In dieser Beschreibung der Bezugsnormen liegt ein Dilemma zwischen Anspruch und Wirklichkeit begründet. Wenn man die Frage nach einer größtmöglichen Objektivität beantworten müsste, so würde man ohne Zögern sagen, dass d) und e) die idealen Bezugsnormen darstellen, am besten sogar eine Mischung aus beiden, wenn es möglich wäre. Im Falle einer selbstreflektierten Sicht auf das eigene Lehrerhandeln wird man bemerken, dass eigentlich a) die zumeist genutzte Bezugsnorm am besten beschreibt. In der Praxis wird es nie möglich sein, d) und e) in Vollkommenheit zu erreichen, allerdings ist es möglich, einen Blick auf diesen Idealzustand zu haben und zu versuchen, sich ihm soweit wie möglich zu nähern. Die Vorgehensweise und passende Hilfsmittel werden im nächsten Abschnitt erläutert.

Es gilt, in einer mündlichen Prüfung bewusst auszuschließen, dass eine andere als die kriteriale Bezugsnorm eine übergeordnete Rolle spielt. Eine hundertprozentige Sicherheit, dass nicht-kriteriale Bewertungsmaßstäbe in die Beurteilung einfließen, kann niemand geben. Das liegt in der Natur menschlicher Handlungen. Allerdings ist es ein Fortschritt, wenn dem Prüfer bewusst ist, dass die Bezugnahme auf eine soziale Gruppe („Hier kommt Bert, der schlechteste der Klasse.“) oder auf individuelle Entwicklungen („Als nächste schauen wir uns Marie an, die sich in den letzten Wochen erheblich verbessert hat.“) einen stark verfälschenden Effekt auf das Ergebnis ausüben kann.

3.4 Gestaltung von Bewertungsrastern

Man kann diejenigen Raster als Vorlage oder als Bearbeitungsgrundlage nehmen, die für die Beispielprüfungen diesem Buch beigelegt sind. Das würde Winter (2015, S. 149) entsprechen, der zu Kompetenzrastern ausführt (was auch solche zur Bewertung von Leistungserfassungen miteinschließt): „Brauchbare Kompetenzraster zu erstellen und mit ihnen angemessen zu arbeiten, erfordert viel fachliche und pädagogische Expertise sowie Zusammenarbeit mit anderen. Das heißt nicht, dass Lehrpersonen in diesem Bereich gar nicht aktiv werden können. In der Regel werden sie aber nur Exegese zu bestehenden guten Kompetenzrastern betreiben können und diese auf ihren Unterricht und bestimmte Aufgabenstellungen hin konkretisieren und ausformulieren.“

In diesem Sinne soll hier auch keine Anleitung gegeben werden, wie solche Raster herzustellen sind, aber deren Genese wird anhand zentraler Punkte erläutert. Die Entwicklung und Verwendung eines Bewertungsrasters gleicht einer Zwickmühle, zwischen dem Versuch einer objektiven, validen Beurteilung und Bewertung einer Prüfung einerseits sowie dem Wunsch nach einem handhabbaren, praktisch gut nutzbaren und effizienten Instrument der Leistungsbeurteilung und -bewertung andererseits. Beiden Seiten muss Rechnung getragen werden und Kompromisse sind dabei unvermeidlich. Diese gehen zu Lasten beider Seiten, sollen aber das übergeordnete Ziel, nämlich mithilfe der mündlichen Prüfung Kompetenzbereiche in den Blick zu nehmen, die in anderen Prüfungen so nicht abzubilden sind, nicht aus dem Blick verlieren.

Die in diesem Buch mitgelieferten Bewertungsraster bestehen aus einer 4 × 5-Matrix, in der vier graduelle Abstufungen hinsichtlich der gezeigten Leistung enthalten sind, die am oberen und am unteren Rand jeweils noch eine zusätzliche Abstufung enthalten. Die fünf Spalten wiederum bilden die wesentlichen Kompetenzbereiche der jeweiligen mündlichen Prüfung ab. Diese sind so gestaltet, dass die Bereiche für alle Prüfungen gleich sind, lediglich in den Inhalten gibt es leichte Anpassungen auf die jeweiligen Prüfungen. Die Bezüge zur Prüfung selbst werden dann in diesen einzelnen Feldern der Matrix konkretisiert und ausgeführt. Nimmt man die Beispieleraster zur Hand, so kann man rasch feststellen, dass sich die Formulierungen, vor allem aber die Adverbien stark ähneln, sogar identisch sind. Das soll vor allem dazu dienen, in einer gewonnenen Praxis keine Brüche zu erzeugen. So kann gewährleistet werden, dass die genutzten Begriffe auch von den handelnden Personen gleich verstanden und interpretiert werden. Die Nutzung gleicher wiederkehrender Adverbien zur Diskriminierung innerhalb eines Kompetenzbereiches erleichtert

die qualitative Abstufung der einzelnen Beschreibungen und bietet ebenfalls Sicherheit in der Kommunikation mit Kollegen. Es ist sinnvoll, immer die gleichen Adverbien zur Beschreibung der Güte der Handlungen und somit der Beobachtungen zu nutzen, um eine gewisse Routine im Umgang mit diesen Adverbien zu erzielen. Diese sollten in jedem Falle, auch bei einer Anpassung der Inhalte des Rasters beibehalten werden. Sie eignen sich übrigens nicht nur für das jeweilige Fach und können leicht angepasst bzw. übernommen werden.

Kriterien und Indikatoren

Ein Kriterium ist ein Fähigkeitsgrad oder ein Merkmal, das beschrieben wird, ein Indikator ist die Umschreibung dessen, was dieses Merkmal anzeigt. Ein Indikator versucht also, etwas messbar zu machen, was zunächst nicht messbar erscheint. Bei Kriterien kann man zwischen quantitativ und qualitativ zu messenden Kriterien unterscheiden. Also solchen, die man schlicht und einfach zählen kann, und solchen, für die man einen Gütemaßstab formulieren muss bzw. verschiedene Indikatoren, die die qualitative Ausprägung des Merkmals sichtbar machen. An einem Beispiel wird der Unterschied deutlich:

Das Merkmal „nutzt die Fachsprache angemessen“ ist ein manifestes Merkmal oder auch ein Kriterium für eine erfolgreiche mündliche Prüfung in einem Schulfach. Manifest ist es deswegen, weil ich es direkt beobachten kann. Das ist nicht bei allen Merkmalen der Fall. Indikatoren für dieses Kriterium können sowohl sein, dass ich beobachte, wie häufig die Fachsprache korrekt oder inkorrekt genutzt wird (quantitativ) oder indem ich beobachte, ob die Nutzung der Fachsprache in der Prüfung zu den Ausführungen passt und notwendig ist (qualitativ).

Die Beschreibung und das Beispiel machen relativ schnell deutlich, dass es auf der pragmatischen Ebene zweifelhaft ist, ob eine Unterscheidung zwischen den Begriffen Kriterium und Indikator wirklich didaktisch und ökonomisch sinnvoll ist. Oft ist es so, dass Kriterien und Indikatoren sich aufgrund der äußeren Zwänge überlagern und sogenannte Mischformen entstehen, die dann in den Bewertungsrastern ihren Niederschlag finden.

Unterschiede existieren für den Fall, dass das zu messende Merkmal zu großen Teilen latent und nicht manifest ist. Diese theoretischen Konstrukte, wie z. B. Intelligenz oder auch LRS, werden erst durch die Indikatoren sichtbar gemacht. Eine Messung auf diesen Ebenen zählt jedoch eher zum Berufsbild eines Psychologen und weniger zu dem einer Lehrkraft. Im Übrigen ist es natürlich so, dass implizit eine Manifestation der latenten Variablen (theoretischen Konstrukte), wie z. B. „naturwissenschaftlich begabt“, durch das zu bewältigende Curriculum vorgenommen wurde. Dieser Schritt wurde also bereits im Vorfeld erledigt und so kann man sich für die Prüfung und deren Bewertung auf die beobachtbaren Kriterien beschränken.

Wichtiger erscheint es in diesem Zusammenhang, dass darauf geachtet wird, dass die formulierten Kriterien (Indikatoren) auch tatsächlich zu beobachten sind, ergo operationalisierbar sind. Denn ein Verhalten, das man nicht beobachten kann, kann man auch nicht bewerten (siehe dazu die Anmerkungen zum Thema Beobachten in Kapitel 3.1). Klassische Fehler in diesem Zusammenhang sind Raster, in denen man Sätze findet wie „Der Schüler hat sich angestrengt.“ oder „Der Schüler hat richtig gedacht.“. Beide Tätigkeiten, welche durchaus fachlich positiv zu sehen sind, sind nicht zu beobachten. Es gibt keine verlässlichen Indikatoren, die garantieren können, dass diese Tätigkeiten tatsächlich stattgefunden haben. In diesem Zusammenhang sei noch einmal darauf hingewiesen, dass man nur Leistung messen kann, aber kein vermeintliches Leistungspotenzial.

Dennoch finden sich in den Beispielrastern Beschreibungen, die sich nicht sofort als operationalisierbar erschließen. Wie passt das zu dem oben Gesagten? Zum Beispiel folgende: „Die Bearbeitung ist weitgehend durchgängig geplant und wird mit fast allen Zwischenschritten erläutert.“ In diesen Fällen erschließt sich

die Beschreibung in der Bearbeitung der Aufgabe. Gleiches gilt auch für Beschreibung folgender Kategorie: „Die Fachsprache wird weitgehend durchgängig und gezielt genutzt.“ Auch hiermit ist gemeint, dass sich die Erfüllung der Beschreibung durch die Bearbeitung der Aufgaben erschließt. Dies sind eher latente Kriterien, die dann durch die Aufgabenbearbeitung (und die damit verbundenen Handlungen) manifest werden. Hier zeigt sich besonders, dass ein Raster nicht den Vorgaben der Testtheorie folgen kann, sondern in Teilbereichen auf Erfahrungen und geteilten Überzeugungen beruht und diese nutzen sollte. Wollte man alle Kriterien und Indikatoren in einem solchen Raster konkret ausformulieren, so würde das Volumen die Handhabbarkeit sicherlich überschreiten.

Zentral ist, dass die genannten Kriterien und Indikatoren valide sind, dass sie also in der Lage sind, wirklich diejenigen Merkmale zu messen, die man zu messen beabsichtigt. Als plakatives Negativbeispiel mag hierbei ein Indikator dienen, den manche Kollegen in Bewertungsrastern zur Mitarbeit benutzen: „Du rufst nicht dazwischen!“. Ohne viel erläutern zu müssen, ist dem Beobachtenden schnell klar, dass dieser Indikator nicht valide ist, denn er misst nicht die Mitarbeit des Schülers. Leider sind viele unpassende Indikatoren nicht so einfach zu erkennen.

Die meisten Fehler werden hinsichtlich der inhaltlichen Validität gemacht, also hinsichtlich der Tatsache, ob – bezogen auf das zu prüfende Fach – eine Übereinstimmung mit den fachlichen und didaktischen Zielen vorliegt. Erwähnt werden soll auch der Aspekt der diskriminierenden Validität bezogen auf Formulierungen und Indikatoren in Bewertungsrastern. Hier geht es um die Frage, ob das angegebene Kriterium bzw. der angegebene Indikator auch in der Lage ist, zwischen verschiedenen Leistungsstufen zu unterscheiden. Man kann sich für diese Frage in der Tat in den meisten Fällen mit den oben angeführten Adverbien behelfen, die eine Diskriminierung in den meisten Fällen ermöglichen. Will man selbst Kriterien und Indikatoren formulieren, so sollte man beizeiten die Frage stellen, ob eine Abstufung hinsichtlich der Leistungsfähigkeit ermöglicht werden kann.

In den Beispielrastern wird keine eindeutige Zuordnung der einzelnen Felder der Matrix zu einer Note gemacht. Das ist Absicht, denn die Note, also die Bewertung der mündlichen Prüfung, ist mehr als eine aus einem Algorithmus hervorgehende Punktschnee, die dort steht, sondern das abwägende pädagogische Urteil, welches aus einer Summe von Beobachtungen hervorgeht. Das erschließt sich auch, wenn man Beurteilung und Bewertung dergestalt ansieht, dass die Leistungsbeurteilung darin besteht, die beobachteten Handlungen konkret mit den vorgegebenen Kriterien in Beziehung zu setzen. Dies geschieht durch die vorgenommenen Markierungen im Raster. Leistungsbewertung wäre dann die Verdichtung dieser Beurteilung auf einer Skala, also etwas, was danach geschieht und nicht gleichzeitig.

In diesem Zusammenhang wird auch immer wieder die Frage gestellt, ob es sinnvoll ist, Fragen oder Arbeitsauforderungen mit einem Punktwert zu versehen, um dann am Ende die Punkte zusammenzählen und somit eine Bewertung schnell und ökonomisch erstellen zu können. Diese Möglichkeit ist aus mehreren Gründen kritisch zu sehen. Dagegen spricht unter anderem, dass durch die Bepunktung eine Wertigkeit innerhalb der Prüfung vorgegeben wird, die durch den spontanen Charakter und die Möglichkeiten der Varietät aber in der Prüfung selbst wieder ganz anders aussehen kann. Es analogisiert die mündliche Prüfung mit einer schriftlichen, was ohnehin nicht im Sinne der mündlichen Prüfung ist. Darüber hinaus können und sollen mündliche Prüfungen ja auch enden bevor ein ganzer Fragenkatalog abgearbeitet wurde. Auch aus diesem Grund erscheint eine Bepunktung wenig sinnvoll, wenn nicht sogar kontraproduktiv. Der letzte und wichtigste hier anzuführende Grund ist aber, dass in einer mündlichen Prüfung innerhalb der verschiedenen Fragen bzw. Aufgaben Abstufungen hinsichtlich der gezeigten Leistung deutlich werden sollen. Eine verbale, mündlich vorgetragene Antwort oder Lösung mit Begründung ist ja wesentlich komplexer zu bewerten und benötigt eine Reihe von Abstufungen, die mithilfe eines oder zweier Punkte nicht zu bewerten sind. Was dann passiert, ist ein Diskretisieren der Prüfung. Das bedeutet, dass die Prüfung als

kontinuierliches Gespräch in Teilpunkte zerlegt werden muss, aber ab einem bestimmten Teil keine weiteren Teilpunkte mehr sinnvoll sind. Die Alternative wäre, für jede Aufgabe wesentlich mehr Punkte vorzusehen, dabei wiederum ist eine Abstufung innerhalb der engen Zeitrahmen – die Prüfung läuft ja parallel weiter – auch schwierig umzusetzen.

3.5 Fehlerquellen in der Bewertung mündlicher Prüfungen

Fehlerquellen spielen in den mündlichen Prüfungen wie auch in anderen Prüfungen eine Rolle. Sicher können nicht alle hier erwähnt werden und einige wurden schon in vorhergehenden Abschnitten geschildert, dennoch ist auf Folgendes zu achten:

- Welcher Eindruck in welchem Teil der Prüfung gewonnen wird, ist entscheidend. So kann es positiv sein, dass eine starke Performanz am Ende der Prüfung den schwachen Beginn vollkommen überlagert und fast vergessen macht. Weniger häufig kann aber auch der umgekehrte Effekt auftreten: Eine schlechte oder gute Antwort zu Beginn legt den Rahmen für die kommenden Antworten. Man misst sozusagen am Anfangszustand.
- Insbesondere in einer mündlichen Prüfung besteht die Gefahr der Überlagerung, dass also weniger relevante Beobachtungen relevante Beobachtungen überlagern. Der einfachste Fall ist der versierte und rhetorisch perfekte Prüfling, dessen inhaltliche Einlassungen allenfalls mittelmäßig sind. Gerade diesen Fehler kann man auf einfache Art mit guten Bewertungsrastern eliminieren.
- Der Reiheneffekt ist ein fataler und zugleich häufig vorkommender Effekt, der besonders dann auftritt, wenn mehrere mündliche Prüfungen hintereinander erfolgen. Hierbei können zwei Nachteile (oder auch Vorteile) für den Prüfling entstehen. Einen besonderen Platz nimmt hierbei der erste Prüfling ein. Da für ihn kein direkter Vergleich mit einem Vorgänger existiert, ist eine Bewertung im mittleren Notenbereich häufig. Extreme Noten werden an dieser Stelle stets vermieden. Für den folgenden Prüfling jedoch entsteht eine eventuell fatale Situation, denn zu der kriterialen Bezugsnorm gesellt sich nun eine prüfungsindividuelle Bezugsnorm, die die Leistung des nachfolgenden an der des Vorgängers misst. Obwohl auf der objektiven Ebene nun tatsächlich durch die beginnende temporäre Routine eine Schärfung auf der inhaltlich-fachlichen Ebene eintritt, ist die zweite Bewertung in hohem Maße von der ersten abhängig und die noch folgenden Prüfungen von allen vorhergehenden.
- Der gesprochenen Sprache kommt in der mündlichen Prüfung ein besonderes Gewicht zu. Deswegen sind gerade Fehldeutungen oder auch ein Fehlhören aus den verschiedensten Gründen möglich und kritisch für eine gute Leistungsbewertung. Ein gezieltes Nachfragen oder Impulsgeben kann den Einfluss dieser Problematik einschränken.
- Erwähnt werden sollen noch zwei Effekte, deren Ursprung in der Länge der Prüfung und in der Anzahl der Prüflinge liegt. So tendieren Prüfer häufig dazu, nach einem längeren Zeitraum von Prüfungen, in einem sogenannten Ermüdungseffekt bessere Noten zu geben. Dementgegen sind Noten zu Beginn einer Reihe von mündlichen Prüfungen eher homogener und schwanken nicht so stark. Beide Effekte sind also ein Plädoyer dafür, nicht allzu viele mündliche Prüfungen hintereinander zu absolvieren.

Literatur

Allgemein

- Birkel, P. (1978): Mündliche Prüfungen. Zur Objektivität und Validität der Leistungsbeurteilung. Bochum: Ferdinand Kamp
- Birkel, P. (2005): Beurteilungsübereinstimmung bei Mathematikarbeiten? Journal für Mathematik-Didaktik, 1 (S. 28–51)
- Bourdieu, P. (2001): Meditationen – Zur Kritik der scholastischen Vernunft. Frankfurt a. M.: Suhrkamp
- Charbel, A. (2004): Top vorbereitet in die mündliche Prüfung. Nürnberg: Bildung und Wissen
- Grünig, B., Kaiser, G., Kreitz, R., Rauschenberger, H., Rinninsland, K. (1999): Leistung und Kontrolle. Weinheim: Juventa
- Heller, K. A. (Hrsg.) (1984): Leistungsdiagnostik in der Schule. Bern: Hans Huber
- Heller, K. A., Hany, E.A. (2002): Standardisierte Schulleistungsmessungen. In: Weinert (Hrsg.) (S. 87–101)
- Ingenkamp, K. (Hrsg.) (1976): Die Fragwürdigkeit der Zensurengebung. Beltz: Weinheim
- Jäger, R. S. (2001): Von der Beobachtung zur Notengebung. Ein Lehrbuch. Landau: Empirische Pädagogik
- Jürgens, E. (2000): Leistung und Beurteilung in der Schule. Sankt Augustin: Academia
- Kirk, S. (2004): Beurteilung mündlicher Leistungen. Bad Heilbrunn: Julius Klinkhardt
- Kleber, E., Meister, H., Schwarzer, C. (1982): Beurteilung und Beurteilungsprobleme. Weinheim: Beltz
- Miller, D. T., Ross, M. (1975): Self-serving biases in the attribution of causality: fact or fiction. Psychological Bulletin, 82 (S. 213–225)
- Paradies, L., Wester, F., Greving, J. (2009): Leistungsmessung und -bewertung. Berlin: Cornelsen Scriptor
- Rauschenberger, H. (1999): Umgang mit Schulzensuren. Funktionen – Entwicklung – Praxis. In: Grünig et al. (S. 11–99)
- Schumacher, C. (2016): Prüfungsangst in der Schule. Münster: Waxmann
- Söll, L. (1974): Gesprochenes und geschriebenes Französisch. Berlin: Erich Schmidt
- Stary, J. (o. J.). „Doch nicht durch Worte nur allein ...“ Die mündliche Prüfung. Zentrales Blogsystem der Freien Universität Berlin. http://blogs.fu-berlin.de/wp-includes/ms-files.php?path=/stary/&file=2011/05/H-2-1_Die-muendliche-Pruefung.pdf (Download: 02. September 2019)
- Watzlawick, P., Beavin, J. H., Jackson, D. D. (1982): Menschliche Kommunikation. Bern: Hans Huber
- Weinert, F. E. (Hrsg.) (2002): Leistungsmessung in Schulen. Weinheim: Beltz
- Winter, F. (2006): Leistungsbewertung. Baltmannsweiler: Schneider Hohengehren
- Winter, F. (2015): Lerndialog statt Noten. Weinheim: Beltz

Fachspezifische Literatur: Physik

- Bayer, R. (2009): Kompetenzorientierter Physikunterricht Basisvortrag. https://lehrerfortbildung-bw.de/u_matnatech/physik/gym/bp2004/fb1/plan/basis/index.html (Download: 02. September 2019)
- Behörde für Schule und Berufsbildung Hamburg (2018): Präsentationsprüfung im Abitur. Physik. www.hamburg.de/bsb/praesentationspruefung-beispielaufgaben/11545520/praesentationspruefung-beispielaufgaben/ (Download: 10. Oktober 2019)

- Girwidz, R. (2015): Medien im Physikunterricht. In: Kircher, E. (Hrsg.): Physikdidaktik. Theorie und Praxis. Berlin: Springer Spektrum (S. 228–245)
- Heinicke, S., Wodzinski, R. (Hrsg.) (2018): Sprachsensibel Physik unterrichten. Naturwissenschaften im Unterricht Physik, 165/166. Seelze: Friedrich
- Kircher, E., Girwidz, R., Häußler, P. (2015): Physikdidaktik. Theorie und Praxis. Berlin: Springer Spektrum
- Kultusministerkonferenz (Hrsg.) (2004): Einheitliche Prüfungsanforderungen in der Abiturprüfung Physik. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Physik.pdf (Download: 02. September 2019)
- Leisen, J. (Hrsg.) (2005): Sprache. Naturwissenschaften im Unterricht Physik, 87. Seelze: Friedrich
- Ministerium für Bildung und Kultur Saarland (2019): Allgemeine Prüfungsanforderungen für das Abitur im Fach Physik. www.saarland.de/dokumente/thema_bildung/APA_Physik_2019.pdf (Download: 10. Oktober 2019)
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2013): Kernlehrplan für die Sekundarstufe II. Physik. www.schulentwicklung.nrw.de/lehrplaene/upload/klp_SII/ph/GOST_Physik_Endfassung.pdf (Download: 10. Oktober 2019)
- Merzyn, G. (1994): Physikschulbücher, Physiklehrer und Physikunterricht. Kiel: IPN
- Posner, G. J., Strike, K. A., Hewson, P. W., Gertzog, W. A. (1982): Accomodation of a scientific conception: Towards a theory of conceptual change. In: Science Education, 66 (S. 211–227)
- Rincke, K., Leisen, J. (2015): Sprache im Physikunterricht. In: Kircher, E., Girwidz, R., Häußler, P.: Physikdidaktik. Theorie und Praxis. Berlin: Springer Spektrum (S. 635–655)
- Schecker, H. et al. (Hrsg.) (2018): Schülervorstellungen und Physikunterricht. Berlin: Springer Spektrum
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2004): Beschlüsse der Kultusministerkonferenz - Bildungsstandards im Fach Physik für den Mittleren Schulabschluss. München: Wolters Kluwer

Engagiert unterrichten. Begeistert lernen.

Weitere [Downloads](#), [E-Books](#) und [Print-Titel](#) des umfangreichen AOL-Verlagsprogramms finden Sie unter:

www.aol-verlag.de



AOL
verlag

Hat Ihnen dieser Download gefallen? Dann geben Sie jetzt auf www.aol-verlag.de direkt bei dem Produkt Ihre Bewertung ab und teilen Sie anderen Kunden Ihre Erfahrungen mit.

Impressum

Mündliche Prüfungen Physik • Klasse 7-10

Florian Kern unterrichtet Physik und Mathematik an einer Gemeinschaftsschule in Saarbrücken. Er arbeitet an Möglichkeiten der alternativen Erfassung von Leistungen, insbesondere „mündlichen“ Prüfungen.

Der Band ist unter Mitarbeit von Matthias Römer entstanden:

Matthias Römer ist seit über 20 Jahren Lehrer an Gesamt- und Gemeinschaftsschulen für Mathematik und Sozialkunde. Neben seiner Tätigkeit in der Schule führt er am Landesinstitut für Pädagogik und Medien Fortbildungen für Mathematiklehrerinnen und -lehrer durch und ist an die Universität des Saarlandes, Lehrstuhl für Mathematik und ihre Didaktik abgeordnet. Er hat einen Sohn.

Hinweis: Wir sprechen hier wegen der besseren Lesbarkeit von Schülern bzw. Lehrern in der verallgemeinernden Form. Selbstverständlich sind auch alle Schülerinnen und Lehrerinnen gemeint.

Das Werk als Ganzes sowie in seinen Teilen unterliegt dem deutschen Urheberrecht. Der Erwerber des Werkes ist berechtigt, das Werk als Ganzes oder in seinen Teilen für den eigenen Gebrauch und den Einsatz im Unterricht zu nutzen. Die Nutzung ist nur für den genannten Zweck gestattet, nicht jedoch für einen weiteren kommerziellen Gebrauch, für die Weiterleitung an Dritte oder für die Veröffentlichung im Internet oder in Intranets. Eine über den genannten Zweck hinausgehende Nutzung bedarf in jedem Fall der vorherigen schriftlichen Zustimmung des Verlages.

Sind Internetadressen in diesem Werk angegeben, wurden diese vom Verlag sorgfältig geprüft. Da wir auf die externen Seiten weder inhaltliche noch gestalterische Einflussmöglichkeiten haben, können wir nicht garantieren, dass die Inhalte zu einem späteren Zeitpunkt noch dieselben sind wie zum Zeitpunkt der Drucklegung. Der AOL-Verlag übernimmt deshalb keine Gewähr für die Aktualität und den Inhalt dieser Internetseiten oder solcher, die mit ihnen verlinkt sind, und schließt jegliche Haftung aus.

© 2020 AOL-Verlag, Hamburg
AAP Lehrerwelt GmbH
Alle Rechte vorbehalten.

Veritaskai 3 · 21079 Hamburg
Fon (040) 32 50 83-060 · Fax (040) 32 50 83-050
info@aol-verlag.de · www.aol-verlag.de

Redaktion: Kathrin Grüling
Lektorat: Uschi Pein-Schmidt, Sickete
Grafik/Layout/Satz: Satzpunkt Ursula Ewert GmbH,
Bayreuth

Coverfoto: © Monkey Business Images – stock.adobe.com

Bestellnr.: 10592DA1

Engagiert unterrichten. Begeistert lernen.

AOL
verlag