

Erstellen von schriftlichen, mündlichen und Multiple-Choice-Prüfungen

Wien, 24.02.2015

Inhaltsverzeichnis

Einleitung	3
Weitere Informationen und Links zum Thema Prüfen:	3
I. Orientierung an Studienzielen	4
Formulieren von Studienzielen	4
II. Prüfen	5
Konstruktion von Prüfungsfragen	5
Wissen/Erinnern	5
Können	6
Gütekriterien von Prüfungen	9
III. Beurteilen und Prüfungsmethoden	10
Bezugsnormen als Maßstab für die Beurteilung	10
Notenschlüssel	10
Prüfungsmethoden	11
Multiple Choice-Fragen und schriftliche Prüfungen mit einfachen offenen Fragen	11
Schriftliche Prüfungen (Essay-Fragen) und mündliche Prüfungen	11
IV. Beurteilungsfehler	12
Beurteilungsfehler vermeiden	13
Quellen und weiterführende Literatur	14

Einleitung

Leistungsüberprüfungen sind ein wesentlicher Bestandteil universitärer Lehre und beim Erreichen einer stärkeren **Studierendenzentriertheit** universitärer Lehre spielen qualitätsvolle Leistungsüberprüfungen eine tragende Rolle. Dazu ist es erforderlich, das **Wissen und Können der Studierenden** in den Mittelpunkt des Lehrens und Lernens sowie auch des Prüfens zu stellen. Für die konkrete Umsetzung in einer Lehrveranstaltung bedeutet dies, **Studienziele (Learning Outcomes)** nach Wissen und Können auszudifferenzieren und die Leistungsüberprüfung auf diese Ziele abzustimmen.

Das vorliegende Dokument fokussiert auf die **Prüfungsmethoden schriftliche Prüfung, Multiple-Choice-Prüfung und mündliche Prüfung**, die teilweise bereits im Curriculum festgelegt sind.

Das Dokument richtet sich vor allem an Lehrende und FunktionsträgerInnen im Bereich Lehre und soll Unterstützung bei der Beantwortung folgender Fragen bieten:

- ▶ Wie ist studienzielorientiertes Prüfen in der Praxis umsetzbar?
- ▶ Wie kann eine faire Beurteilung studentischer Leistungen gewährleistet werden?
- ▶ Wie kann die Qualität von Leistungsüberprüfungen grundsätzlich gesteigert werden?

Die **Funktionen von Leistungsüberprüfungen** unterscheiden sich je nach Lehr-/Lernkontext und Perspektive der beteiligten Personen (Lehrende bzw. Studierende). Leistungsüberprüfungen erfüllen einerseits die **Funktion eines Leistungsnachweises** im Sinn einer abschließenden Feststellung und Beurteilung von studentischem Wissen und Können. Andererseits können Leistungsüberprüfungen auch eine **lernfördernde Funktion** haben, abhängig von ihrer Form und der Differenziertheit der Rückmeldungen an die Studierenden.

Lehrende können die Ergebnisse der Leistungsüberprüfung als Feedback dafür nutzen, inwieweit Studierende die angestrebten Studienziele einer Lehrveranstaltung oder eines Moduls erreicht haben und entsprechende Anpassungen ihrer Lehre daraus ableiten. Für **Studierende** bestimmen Leistungsüberprüfungen das Verständnis des Faches mit und wirken steuernd auf das Lernen. Klar kommunizierte Leistungsanforderungen und **faire Beurteilungen** durch die Lehrenden können zur **Motivation der Studierenden** wesentlich beitragen und schaffen Klarheit bezüglich gegenseitiger Erwartungen.

Weitere Informationen und Links zum Thema Prüfen:

- ▶ Vom CTL ergänzend bereitgestellte **Infoblätter zu den einzelnen Prüfungsmethoden** berücksichtigen auch wichtige rechtliche Aspekte <https://ctl.univie.ac.at/services-zur-qualitaet-von-studien/pruefungen-im-studium/qualitaet-von-pruefungen/materialien/>
- ▶ Informationen zum Thema **Prüfungsstandards für Studienprogrammleitungen und Lehrende** finden sich im Dokument „**Standards von Prüfungen und schriftlichen Arbeiten**“, das als Diskussionsgrundlage im Intranet unter einem gleichnamigen Eintrag unter Themen A-Z zu finden ist.

- ▶ Ausführliche Informationen zu den **rechtlichen Aspekten** finden sich auf der **Website des Büros der Studienpräses**, insbesondere Checklisten zur Durchführung von Prüfungen bei nicht prüfungsimmanenten und bei prüfungsimmanenten Lehrveranstaltungen: <http://studienpraeses.univie.ac.at/informationmaterial/>.
- ▶ Ein zentrales Nachschlagewerk zu den **rechtlichen und administrativen universitätsweiten Rahmenbedingungen** für Leistungsüberprüfungen stellt das **Handbuch für Lehrende** dar (Download unter <https://ctl.univie.ac.at/materialien/handbuch-fuer-lehrende/>).

I. Orientierung an Studienzielen

Studienziele (Learning Outcomes) beschreiben die von den Lernenden erwarteten Ergebnisse der Lehre – sie sind Aussagen darüber, was Studierende am Ende eines Lernprozesses **wissen** (im Sinn von Kenntnissen) und **können** (im Sinn von Kompetenzen¹) sollen. Dabei werden fachliche Kompetenzen und überfachliche Kompetenzen² unterschieden. In der Regel verlagert sich in den Studien der Schwerpunkt mit zunehmendem Studienfortschritt von Wissen auf Können.

Bereits die Konzeption und Beschreibung eines Studiums, eines Moduls, einer Lehrveranstaltung erfordert, das Augenmerk auf die zu erreichenden Studienziele zu richten. Die explizierten Studienziele ergeben die Basis für die Wahl angemessener, **aktivierender Lehr-/Lernmethoden**.

Die vor Beginn einer Lehrveranstaltung bzw. eines Moduls explizierten **Studienziele bilden die Grundlage für die Leistungsüberprüfung**. Die formulierten Studienziele erleichtern die Entscheidung, welche **Art der Leistungsüberprüfung** sich für die Überprüfung der gesetzten Ziele eignet, die **Konstruktion von angemessenen Prüfungsfragen** sowie **adäquate Beurteilungen**. Die Studienziele sollten auch an die Studierenden klar kommuniziert werden.

Formulieren von Studienzielen

Eine **Hilfestellung bei der Formulierung von Studienzielen** bieten sogenannte (Lernziel-) Taxonomien wie etwa jene von Bloom (1956/1972) bzw. deren Revision durch Anderson und Krathwohl (2001). Sie erleichtern die Ausdifferenzierung der Studienziele nach verschiedenen **kognitiven Anspruchsniveaus** und in weiterer Folge auch das Formulieren von Prüfungsfragen im Sinn von Wissen sowie Können (Verstehen, Anwenden, Analysieren etc.). Das Verwenden von Verben, die sich an überprüfbaren Handlungen von Studierenden orientieren, kann hierbei unterstützend sein. In der folgenden Tabelle finden sich Beispiele für Verben zur Formulierung von Studienzielen:

¹ Zu den Begriffen „Lernergebnis“ und „Kompetenz“ vgl. ECTS-Leitfaden 2009, S. 13-15; http://wissenschaft.bmwf.gv.at/fileadmin/user_upload/Bologna/ECTS/ECTS_Leitfaden.pdf [Zuletzt aufgerufen am 18.02.2015]

² Zu den überfachlichen Kompetenzen zählen insbesondere Problemlösungsfähigkeit, die Fähigkeit einen wissenschaftlichen Bericht zu verfassen, kritisches Denken, Analysefähigkeit.

STUDIENZIELE					
Wissen	Können				
Wissen/Erinnern	Verstehen	Anwenden	Analyse	Kreieren/Synthese	Bewertung
reproduzieren	darstellen	durchführen	testen	zusammensetzen	beurteilen
angeben	beschreiben	berechnen	vergleichen	sammeln	argumentieren
benennen	bestimmen	benutzen	kontrastieren	organisieren	voraussagen
auflisten	demonstrieren	herausfinden	isolieren	konstruieren	wählen
schildern	ableiten	aufstellen	auswählen	konzipieren	evaluieren
bezeichnen	diskutieren	ausfüllen	unterscheiden	schreiben	begründen
...

Nachfolgend finden sich einige Beispiele für formulierte Studienziele im Sinn von Wissen und Können (in Anlehnung an die Lernzieltaxonomie von Bloom, 1956/1972, bzw. deren Revision durch Anderson und Krathwohl 2001).

II. Prüfen

Konstruktion von Prüfungsfragen

Für die Konstruktion passender Prüfungsfragen ist es (wie schon beim Formulieren der Studienziele) sinnvoll, sich die Frage zu stellen, auf welchem kognitiven Anspruchsniveau sich das zu überprüfende Studienziel befindet. Für ein differenziertes Erfassen des Wissens sowie des Könnens sollte bei der Fragenkonstruktion darauf geachtet werden, Fragen mit unterschiedlichem kognitivem Anspruchsniveau und Schwierigkeitsgrad zu erstellen.

Eine eindeutige Zuordnung der Prüfungsfragen zum jeweiligen kognitiven Anspruchsniveau ist oft nur schwer möglich und in der Praxis wohl auch nicht nötig. Grundgedanke hinter der Auflistung der folgenden Kategorien und Beispiele ist es, mögliche Formulierungen von Prüfungsfragen und Studienzielen zu zeigen.

Wissen/Erinnern

Die Studierenden sollen das gemäß Studienzielen erworbene **Wissen wiedergeben**. Die Kategorie Wissen unterscheidet sich von den anderen Kategorien der Taxonomie von Lernzielen darin, dass das **Erinnern** der hauptsächliche Prozess ist. Zum Wissen zählt unter anderem: konkrete Einzelheiten erinnern; die Bedeutung von Fachausdrücken/Terminologien kennen; Wissen über spezifische Fakten, Formeln, Grundprinzipien, Klassifikationen und Kategorien, Methoden, Theorien etc.

Wissen/Erinnern:

Beispiele für die Formulierung von Studienzielen und Prüfungsfragen

Mögliche Studienziele

- ▶ Die wichtigsten Begriffe der Buchhaltung definieren (terminologisches Wissen).
- ▶ Reproduzieren biologischer Tatsachen, die für ein Verständnis biologischer Prozesse wichtig sind (Wissen einzelner Fakten).
- ▶ Die Regeln der Interpunktion auflisten (Wissen von Konventionen).
- ▶ Grundlegendes Wissen über die Evolution des Menschen erwerben (Wissen über Trends und zeitliche Abfolgen).
- ▶ Die besonderen Eigenarten verschiedener Unternehmensformen darlegen (Wissen von Klassifikationen und Kategorien).
- ▶ Die Kriterien benennen, durch die eine zuverlässige Quelle von Informationen in den Sozialwissenschaften erkannt werden kann (Wissen von Kriterien).
- ▶ Techniken und Methoden auflisten und definieren, die von WissenschaftlerInnen benutzt werden, um grundlegende Probleme zu lösen (Wissen von Methoden).
- ▶ Die biologischen Gesetze der Fortpflanzung und Vererbung wiedergeben (Wissen von Prinzipien und Verallgemeinerungen).

Mögliche Prüfungsfrage

Mögliches Studienziel: Studierende können die drei Gehirnstrukturen nennen, die für Lernprozesse zentral sind. Wie diese Gehirnstrukturen zusammenhängen, muss in diesem Stadium noch nicht verstanden werden.

Welche Gehirnstrukturen sind für Lernprozesse von vorrangiger Bedeutung?

- 1) Hirnstamm
- 2) Limbisches System
- 3) Hypophyse
- 4) Großhirn

Können

Das **Prüfen von eigenständigen Denkleistungen**, die auf dem erworbenen Wissen aufbauen („Können“ oder auch „Transferleistungen“ genannt), entspricht folgenden kognitiven Anspruchsniveaus:

▶ Verstehen

Die **Bedeutung erlernten Wissens** soll erfasst werden. Die Studierenden wissen, worüber gesprochen wird und können von den darin enthaltenen Ideen Gebrauch machen. Ein Übertragen des Wissens auf andere Bereiche ist noch nicht erforderlich. Zum Verständnis zählen das Übersetzen, das Interpretieren und das Extrapolieren.

Verstehen:

Beispiele für die Formulierung von Studienzielen und Prüfungsfragen

Mögliche Studienziele

- ▶ Einen längeren Teil einer Information in kürzere oder abstraktere Ausdrücke fassen (Übersetzen).
- ▶ Verschieden Arten von Sozialdaten interpretieren (Interpretieren).

Mögliche Prüfungsfrage

Mögliches Studienziel: Studierende können anhand eines konkreten Beispiels die zugrunde liegende Art der Motivation bestimmen.

Seit Johanna von ihrem Vater zur Matura eine Auto versprochen wurde, wenn Sie in Englisch ein „sehr gut“ schafft, lernt Gerda jeden Tag zwei Stunden Englisch. Ihr Vater ist sehr froh, dass seine Tochter so hoch motiviert ist.

Von welcher Art der Motivation ist in diesem Beispiel die Rede?

- 1) Extrinsische Motivation
- 2) Intrinsische Motivation
- 3) Implizite Motivation
- 4) Emotionale Motivation

▶ **Anwenden**

Das Erlernete (angeeignetes und **verstandenes Wissen**) soll in **konkreten Situationen angewandt** werden können. „Anwendung“ unterscheidet sich von „Verständnis“ insofern, als Studierende nicht nur zeigen, dass sie beispielsweise Prinzipien grundsätzlich benutzen können, sondern darüber hinausgehend in konkreten Situationen auch richtig benutzen werden.

Anwenden:

Beispiele für die Formulierung von Studienzielen und Prüfungsfragen

Mögliche Studienziele

- ▶ Die Fähigkeit, die Gesetze der Trigonometrie in praktischen Situationen zu benutzen.
- ▶ Die Fähigkeit, Verallgemeinerungen und Schlüsse aus den Sozialwissenschaften auf aktuelle soziale Probleme anzuwenden.

Mögliche Prüfungsfrage

Mögliches Studienziel: Studierende können grundlegende statistische Berechnungen bei einer konkreten Fragestellung anwenden.

Bei einer Statistikprüfung haben Studierende folgende Punkte erreicht:

Punkte erreicht	4	6	7	8	9	10	11	13	14
Anzahl Studierende	1	3	5	8	7	5	2	1	1

Bestimmen Sie die den Modalwert, den Median, das arithmetische Mittel, die Varianz und die Standardabweichung.

▶ **Analyse**

Das vorgegebene Material wird in seine wesentlichen Teile zerlegt und Beziehungen zwischen den Teilen sollen erkannt werden. **Bedeutungen und Aussagen einer Information sollen ermittelt werden.**

Analyse:

Beispiele für die Formulierung von Studienzielen und Prüfungsfragen

Mögliche Studienziele

- ▶ Die Fähigkeit, Tatsachen von Hypothesen zu unterscheiden (Analyse von Elementen).
- ▶ Die Fähigkeit, logische Fehler in Argumenten zu bestimmen (Analyse von Beziehungen).

Mögliche Prüfungsfrage

Mögliches Studienziel: Studierende können Texte interpretieren und die Bedeutung der Aussage analysieren.

Ein Text wird vorgegeben und die Studierenden sollen diesen Text auf bestimmte Faktoren hin analysieren.

Sind in diesem Text stillschweigende Annahmen zu erkennen? Wenn ja, welche?

Welche Motive des Autors sind zu erkennen?

Welches sind die Hauptargumente für die These „XY“ des Autors?

Begeht der Autor in seiner Argumentation logische Fehler? Wenn ja, welche?

▶ **Kreieren/Synthese**

Aus einzelnen Elementen und **aus Teilen soll etwas Ganzes oder Neues** aufgebaut oder **gestaltet** werden, indem etwa Untersuchungsdesigns **abgewandelt** werden oder eine Werbestrategie für ein Unternehmen **entwickelt** wird.

Kreieren/Synthese:

Beispiele für die Formulierung von Studienzielen und Prüfungsfragen

Mögliches Studienziel

- ▶ Die Fähigkeit ein Gebäude zu entwerfen, das einem bestimmten Zweck dient (Entwerfen eines Plans für bestimmte Handlungen)

Mögliche Prüfungsfrage

Mögliches Studienziel: Studierende können eine Unterrichtseinheit zu einem bestimmten Thema planen.

Entwerfen Sie eine 45 Minuten dauernde Unterrichtseinheit zum Thema „Studieren an der Universität Wien“ als Vorbereitung für eine Exkursion an die Universität Wien mit einer Maturaklasse.

▶ **Bewertung**

Die **Güte eines vorgelegten Materials** kann beurteilt werden, die **Korrektheit** eines Materials kann aufgrund der vorhandenen Informationen beurteilt werden, **quantitative und qualitative Urteile** können getroffen werden, z.B. inwieweit Material und Methoden bestimmte Kriterien erfüllen.

Bewertung:

Beispiele für die Formulierung von Studienzielen und Prüfungsfragen

Mögliche Studienziele

- ▶ Die Fähigkeit, logische Fehler in einer Begründung nachzuweisen (Urteilen aufgrund innerer Evidenz).
- ▶ Die Fähigkeit, ein Werk mit den höchsten bekannten Standards in diesem Bereich zu vergleichen und zu beurteilen (Urteilen aufgrund äußerer Kriterien).

Mögliche Prüfungsfrage

Mögliches Studienziel: Studierende können anhand externer Normen eine Leistung beurteilen.

Den Studierenden wird ein wissenschaftlicher Artikel vorgegeben. Der Artikel soll begutachtet werden.

Ist die Begründung der Fragestellung durch die zitierte Literatur gegeben?

Ist die Herleitung der Hypothesen logisch und nachvollziehbar?

Ist die Wahl der Untersuchungsmethode passend?

Wurde die Untersuchung sorgfältig durchgeführt und dokumentiert?

Sind die Resultate angemessen aufbereitet?

Ist die Interpretation der Ergebnisse schlüssig und vollständig?

Gütekriterien von Prüfungen

Prüfungen sollten idealerweise bestimmte **Gütekriterien** erfüllen. Um die Qualität einer Prüfung zu sichern oder zu heben, sind bei der Konstruktion und Konzeption einer Prüfung insbesondere folgende Fragen zu berücksichtigen:

- ▶ **Misst die Prüfung, was sie messen soll, ist sie gültig (Validität)?** Deckt eine Prüfung den Umfang der Studienziele repräsentativ ab? Messen die Prüfungsaufgaben das in den Studienzielen definierte Wissen und Können?

Um diese Fragen in der **Praxis** zu klären, können sich **KollegInnen gegenseitig Feedback auf Prüfungsfragen** geben. Jede Prüfungsfrage wird dahingehend überprüft, ob sie „misst, was die Prüfung messen soll“, d.h. misst die Frage eines der Studienziele? Hierfür eignen sich KollegInnen, die mit dem Prüfungsstoff und den Studienzielen vertraut sind. Neben der Studienzielorientierung können die KollegInnen beispielsweise auch überprüfen, ob bei der Frage ungewollt Sprachverständnis oder Aufmerksamkeit mitgeprüft werden.

- ▶ **Wie genau misst die Prüfung (Reliabilität)?** Eine Prüfung ist dann reliabel, wenn das geprüfte Merkmal (Wissen und Können in einem bestimmten Fach) mit einem hohen Maß an Genauigkeit gemessen wird. **Eine höhere Anzahl mittelschwerer Prüfungsfragen führt in der Regel zu einer höheren Reliabilität.** Wenn die Umstände in der Praxis es zulassen, sollen **in Parallelgruppen dieselben Fragen** (vielleicht in unterschiedlicher Reihenfolge) vorgegeben werden.
- ▶ **Sind die Ergebnisse einer Prüfung unabhängig von der durchführenden Person (Objektivität)?** Sind die Ergebnisse einer Prüfung unabhängig davon, wer diese durchführt, auswertet oder interpretiert? In der Praxis ermöglichen im Vorfeld verfasste **Musterlösungen und/oder Kriterien** für die Beurteilung der Antworten eine standardisierte und faire Beurteilung der Leistungen durch verschiedene PrüferInnen.

Daneben spielen natürlich u.a. die notwendigen Ressourcen (Zeit, Budget, Material etc.) gemessen am Informationsgewinn einer Prüfung (Gütekriterium der **Ökonomie**) und die zeitliche, psychische und körperliche Belastung für die Studierenden (**Zumutbarkeit**) eine Rolle.

III. Beurteilen und Prüfungsmethoden

Bezugsnormen als Maßstab für die Beurteilung

Um eine Leistung fair beurteilen zu können, sind am besten die **vorab definierten Studienziele als Bezugsnorm** heranzuziehen. Die Studienziele sollten auch den Studierenden bekannt gemacht werden, ebenso gilt es ein **Kriterium** für „Ziel erreicht“ zu definieren (z.B.: „mehr als 50% der Punkte müssen für eine positive Note erreicht werden“ – siehe „Notenschlüssel“ weiter unten). Bei dieser **kriteriumsorientierten Bezugsnorm** gelten absolute **Leistungsstandards (in welchem Ausmaß werden die gesetzten Leistungsstandards bzw. Studienziele von der geprüften Person erfüllt?)**.

In der universitären **Praxis** kann erfahrungsgemäß der soziale Kontext in der Prüfungssituation eine gewisse Rolle spielen, besonders bei mündlichen Prüfungen. Allerdings ist eine **Orientierung an der sozialen Norm (die einzelne Leistung wird relativ zur Leistung aller anderen beurteilt) studienrechtlich problematisch** und gegenüber den Studierenden unfair. Bei Bewertungen nach der sozialen Norm geht man davon aus, dass die Leistungen der Studierenden einer Standardnormalverteilung (**Gauß'sche Glockenkurve**) folgen. Das führt dazu, dass die besten eines Prüfungstermins – unabhängig vom Erreichen der Studienziele – ein „Sehr Gut“ und die Schlechtesten ein „Nicht Genügend“ erhalten.

Wird nach der sozialen Norm bewertet, könnten PrüfungsteilnehmerInnen mit der gleichen Leistung bei verschiedenen Prüfungsterminen unterschiedlich benotet werden. **Die Beurteilung nach der sozialen Norm wäre mit den studienrechtlichen Vorgaben im Universitätsgesetz und der Satzung der Universität Wien nicht vereinbar** (vgl. Seite 7f.).

Notenschlüssel

Die Leistung der einzelnen Studierenden soll mit im Vorfeld definierten **Beurteilungskriterien** (Notenschlüssel) verglichen werden. Der Grad, in dem die auf Basis der Studienziele definierten Anforderungen erfüllt werden (kriteriumsorientierte Bezugsnorm), bestimmt die Note. Die Note der einzelnen Studierenden hängt also von der eigenen Leistung im Vergleich zu einem vorher festgelegten **Kriterium** und nicht von der Leistung der anderen Studierenden ab. **Im Normalfall wird eine Punkteanzahl bestimmt, für die die Studierenden gerade noch eine positive Note erhalten (Bestehensgrenze)**. Die Bestehensgrenze und die Aufteilung der restlichen Noten werden in der Regel von den Lehrenden festgelegt. **In der Praxis werden oft „mehr als 50%“ der Punkte für eine positive Note verlangt**. Zwischen der Bestehensgrenze und der maximalen Punktezahl können die restlichen Noten entweder linear bestimmt werden (der Abstand zwischen 4 und 3 ist gleich groß wie der Abstand zwischen 3 und 2 oder 2 und 1) oder man kann einer anderen Verteilung folgen (z.B.: der Abstand zwischen 4 und 3 ist größer als der Abstand zwischen 3 und 2, dieser Abstand ist wiederum größer als der Abstand zwischen 2 und 1).

Prüfungsmethoden

Je nach Art der Studienziele und der gegebenen Rahmenbedingungen können verschiedene Methoden der Leistungsüberprüfung zum Einsatz kommen. **Teilweise sind die Prüfungsmethoden bereits im Curriculum festgelegt.**

In der Folge werden aus prüfungsdidaktischer Perspektive behandelt:

- ▶ Multiple Choice-Fragen und schriftliche Prüfungen mit einfachen offenen Fragen
- ▶ schriftliche Prüfungen (Essay-Fragen)
- ▶ mündliche Prüfungen

Auf der [Website](#) des CTL sind **Infoblätter** für die **Vorbereitung, Durchführung und Beurteilung** der drei oben genannten und weiteren Methoden der Leistungsüberprüfung zu finden, die auch rechtliche und administrative Aspekte berücksichtigen.

Allen Methoden der Leistungsüberprüfung gemeinsam ist, dass **Musterlösungen** und/oder **Kriterien für das Beurteilen der Antworten** eine möglichst standardisierte und faire Beurteilung der Leistungen ermöglichen. Der größte **Praxisvorteil bei Musterlösungen** ist, dass die einzelnen Prüfungen (auch von unterschiedlichen PrüferInnen) nach dem gleichen Schema und den gleichen Maßstäben fair beurteilt werden können. Nicht zuletzt stellen Musterlösungen und Kriterien für das Beurteilen der Antworten eine nützliche Argumentationshilfe bei möglichen Prüfungseinsichten bzw. Einsprüchen dar (vgl. UG § 79 (5) und § 84). Je nach Art der Prüfung ist dieser Schritt unterschiedlich aufwändig und sollte im Vorfeld erfolgen. **Generell gilt:** Je größer der Gestaltungsraum für die Studierenden beim Erbringen des Leistungsnachweises ist, desto anspruchsvoller ist das Festlegen und Aufschlüsseln von Kriterien für das Beurteilen der Antworten.

Studierende können sich ein gutes Bild über die Anforderungen einer Prüfung machen, wenn sie einige Beispielfragen mit Musterlösungen zur Verfügung gestellt bekommen.

Multiple Choice-Fragen und schriftliche Prüfungen mit einfachen offenen Fragen

Hier ist das Festlegen von **Kriterien für das Beurteilen der Antworten** meist einfach, indem als Kriterium „richtig/falsch angekreuzt“ bzw. „richtiges/falsches Wort eingefügt“ verwendet werden kann. Für jede richtige Antwort wird beispielsweise ein Punkt vergeben (oder auch ein halber Punkt für eine teilweise richtige Antwort) und ab einer bestimmten Anzahl von Punkten wird die Leistung positiv beurteilt.

	MC-Fragen	schriftliche Kurzantwortfragen
Beispiele für Kriterien zur Beurteilung der Antworten	<ul style="list-style-type: none"> ▶ angekreuzt ▶ nicht angekreuzt ▶ teilweise richtig angekreuzt 	<ul style="list-style-type: none"> ▶ richtig ▶ falsch ▶ teilweise richtig

Schriftliche Prüfungen (Essay-Fragen) und mündliche Prüfungen

Gerade für die Korrektur/Beurteilung von offenen Fragen (schriftlichen und mündlichen) empfiehlt es sich, **im Vorfeld Musterlösungen und/oder Kriterien für das Beurteilen der Antworten** zu formulieren. In einer Musterlösung wird eine ideale Lösung formuliert, wobei alle wichtigen Aspekte und Begriffe genannt werden sollen.

Musterlösungen eignen sich hier vor allem für Fragen, auf die es nur eine sinnvolle richtige Antwort gibt. Entsprechend der oben genannten Klassifikation nach Bloom sind das Wissens-, Verständnis- und Anwendungsfragen.

Zusätzlich besteht die Möglichkeit, einzeln zu bewertende Kriterien für jede Frage zu erstellen, sodass **diese einzelnen Kriterien** am Ende ein **Beurteilungsschema** ergeben.

Offene schriftliche oder mündliche Fragen	
Beispiele für Kriterien zur Beurteilung der Antworten	<ul style="list-style-type: none"> ▶ alle Teilaspekte wurden genannt ▶ die geforderten Beispiele wurden angeführt ▶ die zentralen Fachausdrücke wurden richtig verwendet ▶ Zusammenhänge wurden erkannt und aufgezeigt ▶ Lösungen wurden präsentiert und begründet ▶ die Argumentation ist schlüssig ▶ usw....

Beispiel für eine schriftliche oder mündliche Frage mit Musterlösung sowie ein Bündel von Kriterien (Beurteilungsschema)
<p>Frage 1</p> <ul style="list-style-type: none"> ▶ Was sind Nocizeptoren, wo kommen diese vor und worin unterscheiden sie sich von anderen Sinnen? (3 Punkte möglich) <p>Musterlösung</p> <ul style="list-style-type: none"> ▶ Nocizeptoren sind Schmerzrezeptoren, die sich im gesamten Körper außer im Gehirn und in der Lunge befinden. Im Gegensatz zu anderen Sinnen tritt bei Wiederholung oder Andauern von Schmerzreizen keine Adaptation ein, sondern das Gegenteil, eine Sensibilisierung. <p>(Mögliche) Kriterien für das Beurteilen der Antworten</p> <ul style="list-style-type: none"> ▶ Es wird erklärt, was Nocizeptoren sind (1 Punkt) ▶ Es wird darauf hingewiesen, wo Nocizeptoren vorkommen (1 Punkt) ▶ Der Unterschied zu anderen Sinnen wird erklärt (1 Punkt; wenn nur Adaptation oder nur Sensibilisierung genannt wird ½ Punkt) ▶ Weitere nicht direkt erfragte, aber relevante Anmerkungen (1 Punkt) ▶ Insgesamt aber nur 3 Punkte möglich

IV. Beurteilungsfehler

Um eine faire Korrektur von Prüfungen zu gewährleisten, sollten mögliche Beurteilungstendenzen und Beurteilungsfehler bekannt sein, um - falls notwendig - gezielt gegensteuern zu können.

Erwartungseffekte: Positive oder negative Erwartungen von Lehrenden an die Studierenden können sich nach Art einer sich selbst erfüllenden Prophezeiung auf die Beurteilung der Leistungen auswirken. Erwartungen an die Notenverteilung können ebenfalls Auswirkungen haben. So widerstrebt es etwa vielen Lehrenden, fünf- oder zehnmal hintereinander ein „Sehr gut“ zu vergeben, da dies einer impliziten Annahme widerspricht, dass derartig viele „Sehr gut“ nicht hintereinander auftreten können.

Projektionsfehler und Halo-Effekt: Eigene Eigenschaften, Ansichten, Wünsche oder Fehler (zumeist unbewusst) werden auf die Studierenden übertragen und fließen in die Beurteilung

ein. Von einem Merkmal wird auf ein anderes, vielleicht völlig unabhängiges Merkmal geschlossen. Die Wahrnehmung einer Person „überstrahlt“ die zu messende Leistung (Halo-Effekt), wenn sich etwa Kleidung, Sprache, Schriftbild oder Attraktivität von Studierenden auf die Beurteilung ihrer Leistung auswirken. Das kann z. B. dazu führen, dass die wahrgenommene „Begabung“ oder Persönlichkeitseigenschaften und nicht die Leistung beurteilt wird.

Reihenfolgeeffekt: Die Reihenfolge bei der Beurteilung von Prüfungen kann das Ergebnis beeinflussen. So besteht oft die Neigung, bei der Durchsicht die ersten Arbeiten strenger zu beurteilen als die letzten. Auch das Leistungsniveau vorangegangener Prüfungen kann sich auf die Beurteilung auswirken, etwa indem eine mäßige Arbeit besser beurteilt wird, wenn vorher einige sehr schlechte Arbeiten gelesen wurden bzw. umgekehrt. Bei mündlichen Prüfungen merkt man sich die Leistung zu Beginn und am Ende einer Prüfung am besten, diese haben so einen größeren Einfluss auf die Beurteilung.

Strenge- und Mildefehler: Schon „geringe Mängel“ fließen übermäßig stark in die Beurteilung ein und „gute“ Leistungen werden kaum gewertet (Strengefehler). „Gute“ Leistungen werden besonders stark gewichtet, „schlechte“ Leistungen fließen hingegen kaum in die Bewertung ein (Mildefehler). Die eigene Leistungserwartung (z.B. von „jugendstrengen“ oder „altersmilden“ BeurteilerInnen) sollte die Beurteilung nicht beeinflussen.

Tendenz zu Extremurteilen: Es wird vor allem zwischen „guten“ und „schlechten“ Leistungen unterschieden. Ist die Schwelle „gut“ erreicht, wird ungerechtfertigt die beste Beurteilung vergeben, bei Nichterreichen die schlechteste Beurteilung. Urteile in der Mitte des Notenspektrums werden vermieden.

Tendenz zur Mitte: Im Gegensatz zur Tendenz zu Extremurteilen kann auch die Tendenz bestehen, eindeutige (extreme) Urteile zu vermeiden. Die Tendenz zur Mitte tritt vor allem dann auf, wenn die Beurteilenden unsicher sind.

Beurteilungsfehler vermeiden

Lehrende haben verschiedenste Möglichkeiten, Beurteilungsfehler zu minimieren. Beurteilungsfehler wie etwa der Halo-Effekt, Reihenfolgeeffekte, Erwartungseffekte oder Strenge- und Mildefehler können minimiert werden, indem **Prüfungen „quer“ beurteilt** werden. D.h. es wird jeweils eine Prüfungsaufgabe bei allen Prüfungsteilnehmenden kontrolliert, statt jeden Prüfungsbogen in einem Durchlauf zu bewerten. Der „Gesamteindruck“ über die Leistung sollte sich erst am Ende der Korrektur und nicht schon nach ein paar Fragen bilden. Ähnlich hilfreich sind **Musterlösungen und/oder Beurteilungsschemata**.

Bei längeren Korrekturarbeiten oder bei mündlichen Prüfungen mit vielen Studierenden sollten **Pausen** eingebaut werden. Schwankungen in der Strenge der Beurteilung, Kontrasteffekte zu vorherigen Prüfungen und eigene Ermüdungseffekte können so verringert werden.

Darüber hinaus kann es bereits helfen zu überlegen, inwieweit eine persönliche **Neigung zu bestimmten Beurteilungstendenzen** (Strenge oder Milde; Tendenz zu Extremurteilen etc.) besteht und diese bei der Beurteilung in Erinnerung zu behalten. Insbesondere bei mündlichen Prüfungen sollte versucht werden, hervorstechende Merkmale der Studierenden, die mit der eigentlich zu beurteilenden Leistung nichts zu tun haben, auszublenden.

Quellen und weiterführende Literatur

- Anderson, L. W., & D. R. Krathwohl, D. R. (Eds.) (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Bloom, B. S. (Hrsg.) (1956/1972). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim und Basel: Beltz.
- Eugster B., & Lutz, L. (2003/04). *Leitfaden für das Planen, Durchführen und Auswerten von Prüfungen an der ETHZ*. ETH Zürich: Didaktikzentrum. Verfügbar unter: http://www.let.ethz.ch/pruefungen/Leitfaden_PDA_Pruefungen_DiZ-2003.pdf [Zuletzt aufgerufen am 18.02.2015]
- Dubs, R. (o.J.). *Koordinierte Prüfungsverfahren und valide Ergebnisse – eine Selbstverständlichkeit?* Universität St. Gallen.
- Metzger.C., & Nüesch.C. (2004). *Fair prüfen. Ein Qualitätsleitfaden für Prüfende an Hochschulen*. St. Gallen: Institut für Wirtschaftspädagogik.
- Walzik, W. (2012). *Kompetenzorientiert prüfen. Leistungsbewertung an der Hochschule in Theorie und Praxis*. Opladen und Toronto: Verlag Barbara Budrich UTB.