

Handreichung der Prüfungswerkstatt

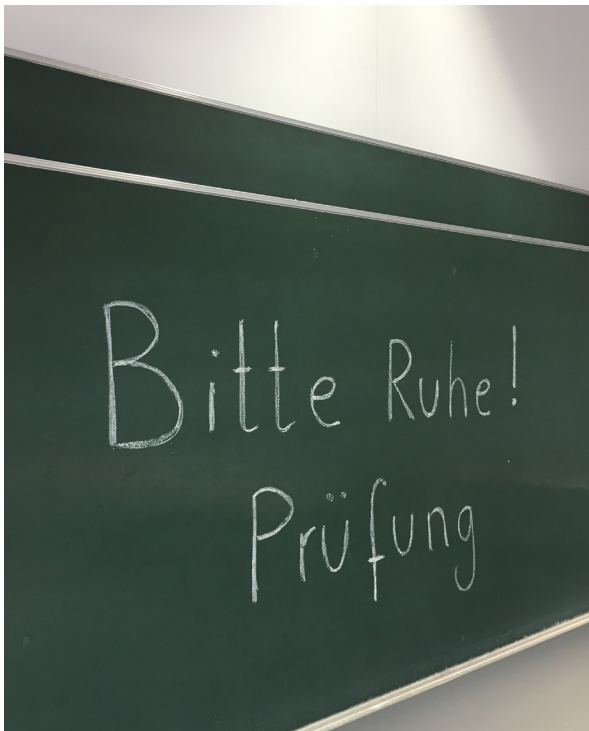
Kriterienraster

Dr. Pia Kreuzer

LOB LEHREN – ORGANISIEREN – BERATEN



Bundesministerium
für Bildung
und Forschung



Inhalt dieser Handreichung

Einleitung	3
Definition / Variation	3
Effekte des Einsatzes eines Kriterienrasters	4
Die Erstellung eines Kriterienrasters	5
Praxisbeispiele	6
Literaturverzeichnis	7

EINLEITUNG

Bereits während der Veranstaltungsplanung sollten Lernziele festgelegt, aber auch Prüfungsform und Prüfungsaufgaben durchdacht und, im Sinne des constructive alignments, aufeinander abgestimmt werden. Nach diesem Schritt müssen klare Maßstäbe (Bewertungskriterien) für die Beurteilung der Prüfungsleistungen definiert werden. Diese haben dabei folgende Anforderungen zu erfüllen.

Bewertungskriterien sollten:

- möglichst objektiv, handhabbar sowie transparent sein und
- nur Kompetenzen und Lernziele erfassen, die in den Lehrveranstaltungen erworben werden konnten.

Im optimalen Fall können diese Bewertungskriterien auch unter Mitwirkung der Studierenden entwickelt und in Form von Kriterienrastern festgehalten und angewendet werden. Sie sollten nicht nur als bloßes Bewertungsinstrument verstanden, sondern von Anfang an ins Kursdesign integriert werden (Haught et al. 2016).

DEFINITION / VARIATION

In einem Kriterienraster werden die für die Bewertung relevanten Kriterien festgehalten. Außerdem kann es gestaffelte Erwartungsniveaus für die einzelnen Kriterien enthalten (Reddy & Andrade 2010). Hierfür sollte für jedes Kriterium definiert werden, wann ein bestimmtes Erwartungsniveau erfüllt ist. Diese Qualitätsdefinitionen nennt man Deskriptoren.

Kriterienraster eignen sich für die Bewertung nahezu aller Formen studentischer Leistungen von der klassischen schriftlichen Arbeit über Referate, Portfolios, mündliche Prüfungen und praktische Arbeiten. Ihr Einsatz in der

Schul- und Hochschullehre wurde in zahlreichen Studien inzwischen gut untersucht. Allerdings variieren Aufbau und Einsatz der Raster stark und konkrete Definitionen oder Angaben zum Einsatz fehlen häufig in diesen Studien (Dawson 2017). Große Unterschiede betreffen zum Beispiel die Art der Bewertung, die durch das Kriterienraster verfolgt werden soll, z.B. holistische oder analytische Bewertung (Dawson 2017, Jonsson & Svingby 2007) (siehe Handreichung „Prüfungsbewertung“).

• Holistische Bewertung

Bei der holistischen Bewertung wird die Prüfungsleistung ganzheitlich beurteilt. Die Erwartungsniveaus umfassen jeweils alle relevanten Kriterien in einem Text. Somit wird der Gesamteindruck beurteilt und die Bewertung erfolgt oft schneller und einfacher als bei der analytischen Bewertung. Die Gewichtung der einzelnen Kriterien kann von Fall zu Fall unterschiedlich ausfallen und der Beurteiler hat eine größere Bewertungsfreiheit. Der Erfahrungsschatz und die Intuition des Beurteilers haben bei dieser Bewertungsmethode einen größeren Einfluss auf das Ergebnis.

• Analytische Bewertung

Bei der analytischen Bewertung werden dagegen einzelne Kriterien festgelegt, die alle zur Gesamtbewertung beitragen. Die Erwartungsniveaus werden hier für jedes Kriterium einzeln definiert und jedes Kriterium wird einzeln beurteilt. Die Einzelbewertungen werden schließlich zu einem Gesamtergebnis zusammengefasst. Hierbei ist es möglich die einzelnen Kriterien unterschiedlich zu gewichten, um so ihren Anteil an der Gesamtbewertung anzupassen. Die Gewichtung wird hier allerdings im Vorfeld festgelegt und ist für alle Bewertungen gleich. Die Gesamtleistung der Studierenden kann so objektiver, reliabler und valider bewertet werden (Schaper, Hilkenmeier & Bender 2013). In weiteren Studien konnte dieser

Effekt jedoch nicht eindeutig belegt werden (Rezaei & Lovron 2010). Analytische Kriterienraster ermöglichen ein detailliertes Feedback, da Studierende ihre Stär-

ken und Schwächen anhand der Bewertung der einzelnen Kriterien gespiegelt bekommen.

Tabelle 1: Beispiel Holistische Bewertung

Erwartungsniveau +	Erwartungsniveau 0	Erwartungsniveau -
Deskriptor umfasst alle Kriterien		

Tabelle 2: Beispiel Analytische Bewertung

	Erwartungsniveau +	Erwartungsniveau 0	Erwartungsniveau -
Kriterium 1	Deskriptor für Kriterium 1		
Kriterium 2	Deskriptor für Kriterium 2		
Kriterium 3	Deskriptor für Kriterium 3		

EFFEKTE DES EINSATZES EINES KRITERIENRASTERS

- **Kriterienraster erhöhen die Bewertungstransparenz**

Schon die bloße Bekanntgabe der Bewertungskriterien durch Aushändigung des Kriterienrasters kann dazu führen, dass den Studierenden bewusst wird, welche Punkte sie bei der Bearbeitung einer Aufgabe berücksichtigen müssen (bspw. inhaltliche und formale Anforderungen bei Hausarbeiten oder Präsentationen). Werden die entwickelten Kriterienraster zusätzlich bereits zu Beginn der Lehrveranstaltung mit den Studierenden besprochen, kann dies das Verständnis der Studierenden für die Bewertungskriterien und die Erwartungen an die Studierenden erhöhen (Andrade & Du 2005).

- **Kriterienraster als Strukturierungs- und Lernhilfe nutzen**

Für Studierende können Kriterienraster als Lernhilfe dienen, da sie die für die Bewertung relevanten Kriterien auflisten. Sie können von Studierenden zum Beispiel als Checklisten oder zur Strukturierung der Aufgaben genutzt werden (Haught, Ahern & Ruberg 2016) und so den Lernprozess unterstützen (Andrade & Du 2005). So können sowohl die Qualität studentischer Arbeiten verbessert, als auch Ängste bezüglich der Erfüllung der an sie gestellten Aufgaben abgebaut werden (Andrade & Du 2005).

- **Kriterienraster zur Selbsteinschätzung und als Feedbackinstrument nutzen**

Kriterienraster können von Studierenden zur Selbsteinschätzung der eignen Leistung genutzt werden (Andrade & Du 2005). Dies kann sowohl während des Arbeitsprozesses als auch zur abschließenden Einschätzung geschehen (Haught et al. 2016). Kriterienraster haben so das

Potenzial, selbstreguliertes Lernverhalten wie das Setzen von Zielen, die Selbsteinschätzung und die Überarbeitung eigener Arbeiten zu fördern (Andrade & Du 2005). Sie können aber auch Hilfestellung bei der Formulierung von Peer-Feedback geben (Andrade & Du 2005). Beim Einsatz bspw. im Rahmen eines Referats können die Studierenden mit Hilfe des Kriterienrasters also nicht nur ihr eigenes Referat an die Qualitätsanforderungen anpassen, sondern auch die Referate der anderen Kursteilnehmer/innen anhand dieser Kriterien bewerten. Diese Anwendung der Bewertungskriterien auf die Arbeiten anderer kann das Verständnis für die Lernziele zusätzlich verbessern (Timmerman & Strickland 2010).

- **Kriterienraster Lehrveranstaltungsübergreifend einsetzen**

Besonders für Kompetenzen, die in mehr als einer Lehrveranstaltung erworben und trainiert werden, z.B. das wissenschaftliche Schreiben oder Präsentieren, sind lehrveranstaltungsübergreifende Kriterienraster sinnvoll. Sie zeigen Studierenden schon zu Beginn des Studiums „wo die Reise hingehen soll“ und helfen den aktuellen Leistungsstand einzuordnen. Über die Auswahl des zu erreichenden Erwartungsniveaus können die Anforderungen an den jeweiligen Leistungsstand der Studierenden angepasst werden. Die Standardisierung der Bewertungskriterien und somit der Anforderungen an Studierende eines Fachs bzw. Studiengangs unterstützen die Entwicklung der Lernenden im Laufe des Studiums und können zudem eine Bewertung des gesamten Curriculums ermöglichen. Dies kann sich auch positiv auf die Leistung der Studierenden auswirken (Timmerman et al. 2010).

- **Orientierungshilfe für die Korrektur in Teams**

Kriterienraster sind besonders dann eine sinnvolle Orientierungshilfe, wenn Tuto-

ren/innen oder Gruppen von Lehrenden an der Korrektur und Bewertung von Arbeiten beteiligt sind. In diesem Fall empfiehlt sich, die Anwendung des Rasters an Hand von Beispielen in der Gruppe im Vorfeld zu besprechen, um eine gemeinsame Interpretation der Kriterien zu entwickeln (Timmerman et al. 2010). Dies reduziert nicht nur den Aufwand für Lehrende bei der späteren Bewertung, sondern verhindert auch eine inkonsistente Beurteilung und erhöht damit ihre Reliabilität (White & Ostheimer 2006; Bloxham 2015).

DIE ERSTELLUNG EINES KRITERIENRASTERS

Allgemeine Überlegungen

Vor der Erstellung des Kriterienrasters sollten zunächst einige Vorüberlegungen angestellt werden, denen folgende Fragen zugrunde gelegt werden können:

- Für welche Studien-/Prüfungsleistung soll ein Raster entwickelt werden?
- Soll ein holistisches oder ein analytisches Raster erstellt werden?
- Für welche Anwender wird das Kriterienraster erstellt (einzelne/r Lehrende/r, Korrekturteam, studentische Tutoren/innen, Studierende für Peer-Feedback)?
- Soll ein lehrveranstaltungsübergreifendes Raster entwickelt werden?
- Soll das Raster zur formativen oder summativen Bewertung genutzt werden?
- Soll es ausschließlich skalierte Einstufungen geben oder wird Platz für individuelle Bemerkungen benötigt?

Bewertungskriterien festlegen

Im nächsten Schritt müssen die Bewertungskriterien festgelegt werden. Hierbei sollte folgendes beachtet werden:

- Die Kriterien der Bewertung ergeben sich aus den Lernzielen der Veranstaltung/des Moduls/des Studiums.

- Die Kriterien sollten alle wesentlichen Bereiche der Lernziele abdecken und trotzdem in einem übersichtlichen Rahmen bleiben.
- Die Kriterien können auch in Zusammenarbeit mit den Studierenden entwickelt werden. Dies schafft ein höheres Empfinden von Fairness sowie eine vertiefte Auseinandersetzung mit den Bewertungsstandards.

Zu jedem Kriterium müssen im Anschluss Qualitätsdefinitionen (Deskriptoren) für die einzelnen Niveaustufen formuliert werden. Hierbei sollten folgende Punkte Beachtung finden:

- Welche Anzahl an Niveaustufen (siehe Handreichung „Kompetenzorientiertes Prüfen“) ist im jeweiligen Kontext sinnvoll? Die Anzahl sollte eine klare Differenzierung zwischen den Niveaustufen zulassen. Die Einordnung sollte eindeutig sein, aber auch Unterschiede zwischen den Studierendenleistungen abbilden. Drei bis sechs Stufen sind in der Regel sinnvoll.
- Eine geringere Anzahl an Stufen erhöht die Reliabilität, differenziertere Stufen (und Kriterien) ermöglichen eine genauere Rückspiegelung der erbrachten Leistung und eventuelle Ansätze der Verbesserung.
- Die Deskriptoren sollten leicht verständlich und eindeutig formuliert werden. Alle Anwender des Kriterienrasters sollten dieselbe Interpretation der Kriterien teilen. Interpretierungsbedürftige Formulierungen wie „gut“, „angemessen“ oder „sinnvoll“ sollten vermieden werden.
- Für die Formulierung empfiehlt es sich zunächst die Eigenschaften des Kriteriums zu bestimmen und diese dann je nach Level abzustufen.

Bewertungsschlüssel erarbeiten

Dient das Kriterienraster der anschließenden Benotung der studentischen Leistungen, sollte das Kriterienraster auch einen Bewer-

tungsschlüssel enthalten. Dieser übersetzt die Einstufungen in die unterschiedlichen Erwartungsniveaus in eine Gesamtnote. Hierfür können die Niveaustufen mit einer numerischen Skala (Punkte) oder Prozentwerten verknüpft werden, um zu einer einheitlichen Bewertung zu gelangen (Quellmelz & Ruschin 2013). Es sollte auch auf die Gewichtung der einzelnen Kriterien geachtet werden, um eine Konzentration auf einfach zu messende Kriterien, wie Rechtschreibung oder Layout, zu vermeiden (Driessen 2006). Zur Bewertung sollten nur Kriterien des Kriterienrasters herangezogen werden. Um Raum für Unerwartetes zu lassen, kann aber eine entsprechende Kategorie mit vorher festgelegter Gewichtung eingeplant werden (Bonus).

PRAXISBEISPIELE

http://www.fb05.uni-mainz.de/Dateien/Bewertungskriterien_FB_05_FBR130522.pdf

Leistungsmessung und -bewertung von Proseminararbeiten

http://www.geschichte.uni-mainz.de/Dateien/Bewertungskriterien_Proseminararbeiten.pdf

http://www.wipaed.uni-mainz.de/ls/719_DEU_HTML.php

Leistungsmessung und -bewertung von Bachelorarbeiten, Masterarbeiten und Hauptseminararbeiten

http://www.geschichte.uni-mainz.de/Dateien/Bewertungskriterien_HS-BA-BED-MA-MED.pdf

Leitfaden zur differenzierten Rückmeldung zu mündlichen Referaten, Präsentationen, Seminarbeiträgen

http://www.geschichte.uni-mainz.de/Dateien/Bewertungskriterien_muendliche_Vortraege.pdf

LITERATURVERZEICHNIS

- Andrade, H. L., Du, Y. (2005): Student perspectives on rubric-referenced assessment. In: *Practical Assessment, Research & Evaluation* 10 (3), S. 1–11.
- Bloxham, Sue; den-Outer, Birgit; Hudson, Jane; Price, Margaret (2015): Let's stop the pretence of consistent marking. Exploring the multiple limitations of assessment criteria. In: *Assessment & Evaluation in Higher Education* 41 (3), S. 466–481. DOI: 10.1080/02602938.2015.1024607.
- Briana Eileen Timmerman; Denise Strickland; Robert L. Johnson; and John R. Payne (2010): Development of a universal rubric for assessing undergraduate scientific reasoning skills using scientific writing. In: *Assessment and Evaluation in Higher Education*, S. 1–39.
- Dawson, Phillip (2015): Assessment rubrics. Towards clearer and more replicable design, research and practice. In: *Assessment & Evaluation in Higher Education* 42 (3), S. 347–360. DOI: 10.1080/02602938.2015.1111294.
- Driessen, Erik W.; Overeem, Karlijn; van Tartwijk, Jan; van der Vleuten, Cees P M; Muijtjens, Arno M. M. (2006): Validity of portfolio assessment: which qualities determine ratings? In: *Medical education* 40 (9), S. 862–866. DOI: 10.1111/j.1365-2929.2006.02550.x.
- Haught, P. A.; Ahern, Terence C.; Ruberg, Laurie F. (2016): Student Awareness and Use of Rubrics in Online Classes. In: *HES* 7 (1), S. 69. DOI: 10.5539/hes.v7n1p69.
- Jonsson, Anders; Svingby, Gunilla (2007): The use of scoring rubrics. Reliability, validity and educational consequences. In: *Educational Research Review* 2 (2), S. 130–144. DOI: 10.1016/j.edurev.2007.05.002.
- Ostheimer, Martha W.; White, Edward M. (2005): Portfolio assessment in an American Engineering College. In: *Assessing Writing* 10 (1), S. 61–73. DOI: 10.1016/j.asw.2005.02.003.
- Quellmelz, M., & Ruschin, S. (2013): Kompetenzorientiert prüfen mit Lernportfolios. In: *Journal Hochschuldidaktik*, S. 1–2.
- Reddy, Y. Malini; Andrade, Heidi (2010): A review of rubric use in higher education. In: *Assessment & Evaluation in Higher Education* 35 (4), S. 435–448. DOI: 10.1080/02602930902862859.
- Rezaei, Ali Reza; Lovorn, Michael (2010): Reliability and validity of rubrics for assessment through writing. In: *Assessing Writing* 15 (1), S. 18–39. DOI: 10.1016/j.asw.2010.01.003.
- Schaper, N., Hilkenmeier, F., & Bender, E. (2013). Umsetzungshilfen für kompetenzorientiertes Prüfen. Ausgearbeitet für die HRK, unter Mitarbeit von Bender, E. Projekt nexus, Konzepte und gute Praxis für Studium und Lehre, verfügbar unter: <http://www.hrk-nexus.de/fileadmin/redaktion/hrk-nexus/07-Downloads/07-03-Material/zusatzgutachten.pdf>.