pedocs
Open Access Erziehungswissenschaften

DIPF

Schaper, Niclas

Prüfungsgestaltung

Pasternack, Peer [Hrsg.]; Reinmann, Gabi [Hrsg.]; Schneijderberg, Christian [Hrsg.]: Hochschulforschung. Forschung über Hochschule und Wissenschaft. Baden-Baden : Nomos Verlag 2025, S. 263-273. - (NomosHandbuch | NomosHandbook)

Quellenangabe/ Reference:
Schaper, Niclas: Prüfungsgestaltung - In: Pasternack, Peer [Hrsg.]; Reinmann, Gabi [Hrsg.]; Schneijderberg, Christian [Hrsg.]: Hochschulforschung. Forschung über Hochschule und Wissenschaft. Baden-Baden : Nomos Verlag 2025, S. 263-273 - URN: urn:nbn:de:0111-pedocs-339136 - DOI: 10.25656/01:33913

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-339136
https://doi.org/10.25656/01:33913

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: http://creativecommons.org/licenses/by/4.0/deed.de - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: http://creativecommons.org/licenses/by/4.0/deed.en - You may copy, distribute and render this document accessible, make adaptations of this work or its contents accessible to the public as long as you attribute the work in the manner specified by the author or licensor.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der Leibniz-Gemeinschaft

II.1.6
Prüfungsgestaltung

Niclas Schaper

Abstract | Der Beitrag gibt einen Überblick über Konzepte und Ansätze der Prüfungsgestaltung aus didaktischer und diagnostischer Sicht. Nach einer Einleitung werden zunächst verbreitete Formate und Bewertungsansätze des Prüfens hinsichtlich zentraler Gestaltungsaspekte vorgestellt und analysiert. Daraufhin folgt die Darstellung zentraler konzeptioneller Ansätze der didaktischen und diagnostischen Prüfungsgestaltung sowie von Studien und Konzepten zur Evaluation und Qualitätssicherung von hochschulischen Prüfungen. Abschließend wird ein Fazit gezogen und ein Ausblick auf Forschungsbedarfe gegeben.

Stichworte | Prüfen in der Hochschullehre, Prüfungsdidaktik, kompetenzorientiertes Prüfen, lernorientierte Prüfungsgestaltung.

Einleitung

Prüfungen sind ein essenzieller Bestandteil der Lehrtätigkeit an Hochschulen. Sie erfüllen wichtige Funktionen sowohl im Rahmen des allgemeinen Studiums als auch speziell für die Lehrangebote innerhalb eines Studiengangs. Dabei sind Prüfungen nicht nur für die Anerkennung von individuellen Leistungen im Studium von zentraler Bedeutung, sondern haben auch wichtige Funktionen im Lernprozess (z. B. in Form von Rückmeldungen zum Lernstand). Darüber hinaus sind Prüfungen für Lehrende ein relevantes didaktisches Element bei der Planung, Umsetzung und Erfolgskontrolle von Lehrveranstaltungen. Vor allem kann anhand der Leistungen in den Prüfungen – wenn diese entsprechend gestaltet sind – erkannt werden, ob die Lernziele einer Lehrveranstaltung von den Studierenden erreicht wurden oder nicht.

Trotz seiner Bedeutung wird das Prüfen von Lernleistungen oftmals wenig professionell gehandhabt und eher als formales Element des Lehrprozesses betrachtet (siehe z. B. Metzger/Nüesch 2004; Schindler et al. 2015). Dies führt dazu, dass Prüfungen an Hochschulen oftmals weder diagnostisch (als Instrument zur Erfassung der erreichten Lernleistungen) noch didaktisch (als Teil des Lernprozesses zur Aneignung von bestimmten Kompetenzen) angemessen eingesetzt und gestaltet werden. Auch die theoriegeleitete Befassung mit hochschulischen Prüfungen – insbesondere unter (hochschul-)didaktischen Perspektiven – ist in vielerlei Hinsicht noch entwicklungsbedürftig, auch wenn erste Ansätze dazu vorliegen (z. B. Reinmann 2019).

Prüfungen erfüllen nicht nur verschiedene Funktionen im Rahmen eines Studiums, sondern müssen auch unterschiedlichsten Anforderungen gerecht werden. Bei ihrer Konzeption und Umsetzung in der Hochschullehre sind daher verschiedene Anforderungsbereiche zu berücksichtigen: curriculare (z. B. das Erreichen von Lernzielen und das Bestehen von Modulen überprüfen), didaktische (z. B. Informationen über den Kompetenzerwerb bzw. Lernzuwachs der Studierenden zur Verfügung stel-

len), diagnostische (z. B. Gütekriterien bei der Konzeption, Durchführung und Auswertung von Prüfungen beachten), rechtliche (z. B. analysieren, ob prüfungsrelevante Rechtsnormen verletzt werden) und administrativ-organisatorische Anforderungen (z. B. gewährleisten, dass die Abwicklung von Prüfungen rechtskonform, nutzerfreundlich und effizient erfolgt). Erst die Berücksichtigung aller genannten Bereiche führt zu einer tauglichen und anforderungsgerechten Prüfung – auch wenn die dabei zu berücksichtigenden Anforderungen in mehrfacher Hinsicht widersprüchlich und nur schwierig auszutarieren sind (Reinmann 2019). Aus Platzgründen muss die Darstellung im Folgenden auf didaktische und diagnostische Aspekte des Prüfens in der Hochschullehre fokussiert werden, die sich insbesondere auf Zugänge der pädagogischen Psychologie zum Prüfen beziehen.

Gegenstände der Prüfungsgestaltung aus didaktisch-diagnostischer Perspektive

Um einen Überblick zu zentralen Aspekten der Analyse und Gestaltung von Prüfungen zu gewinnen, kann auf ein Rahmenmodell von Metzger und Nüesch (2004) Bezug genommen werden. In diesem Modell werden die Funktionen und Anforderungen an das Prüfen sowie der Prüfungsprozess in einen systematischen Zusammenhang gestellt. Prüfen wird dabei als mehrphasiger Prozess verstanden, der Phasen der Planung, Durchführung und Auswertung umfasst. Integriert in diese Phasen müssen eine Vielzahl von Entscheidungen getroffen werden, die vor allem folgende Aspekte betreffen: Was soll geprüft werden (Inhalte)? Wie soll geprüft werden (Formate)? Wie sollen die Prüfungsergebnisse bewertet werden (Bewertung)? Die Gestaltung des Prüfungsprozesses hängt außerdem davon ab, wozu geprüft wird, d. h., welche Funktionen Prüfungen in einem bestimmten Ausbildungskontext erfüllen sollen. Außerdem ist zu klären, in welchem Maße bestimmten Qualitätsanforderungen an eine Prüfung (z. B. Validität oder Fairness) zu erfüllen sind. Schließlich ist der Prüfungsprozess eingebettet in den curricularen Rahmen eines Studiengangs sowie in das Lehr-Lernverständnis, das sich in einer Hochschule bzw. einer Lehreinheit entwickelt hat und diese prägt.

Im Folgenden wird auf Formate des Prüfens und Formen der Bewertung von Prüfungsleistungen eingegangen.

Formate des Prüfens

In Bezug auf Formate des Prüfens wird grundsätzlich zwischen schriftlichen, mündlichen und praktischen Prüfungsformen/-formaten unterschieden, die wiederum in jeweils vielfältiger Form umgesetzt und methodisch gestaltet werden können (z. B. Schaper/Soyka 2021).

Bei schriftlichen Prüfungen werden die Prüfungsleistungen durch die Beantwortung oder Bearbeitung schriftlich vorgegebener Aufgaben erbracht. Bei der Gestaltung schriftlicher Prüfungsaufgaben wird zwischen stark vorstrukturierten (z. B. Multiple-Choice-Aufgaben) und offeneren Formen der Bearbeitung (z. B. offene Fragen zu einer Themenstellung) unterschieden. Außerdem ist zwischen zeitpunktbezogenen (z. B. Klausuren) versus zeitraumbezogenen (z. B. Haus- oder Abschlussarbeiten) Formaten zu differenzieren. Grundsätzlich gilt, dass strukturiertere Formate sich eher zur Überprü-

fung von umgrenzten Wissensleistungen eignen, wohingegen offenere Formate in der Lage sind, auch komplexere Leistungen bzw. Kompetenzen abzuprüfen. Schriftliche Prüfungen kommen in fast allen Studiengängen und auf allen Ausbildungsstufen vor und sind auch das am häufigsten eingesetzte Prüfungsformat (z. B. Schindler et al, 2015).

Das nach schriftlichen Prüfungen verbreitetste Format sind mündliche Prüfungen. Hierbei werden die Prüfungsleistungen in mündlicher Form erbracht. Am häufigsten erfolgen mündliche Prüfungen in Form eines strukturierten Abfragens von Wissen. Die Fragen zielen dabei nicht nur auf Erinnerns- und Verstehensleistungen ab, sondern können sich auch auf Analyse-, Beurteilungs- oder Syntheseleistungen beziehen. Eine weitere, kompetenzorientiertere Gestaltungsvariante ist das Fachgespräch. Hier wird erwartet, dass der Prüfling mit dem Prüfenden in einen fachlichen Dialog eintritt, indem er z. B. eine Position in einem fachlichen Diskurs bezieht und verteidigt. Der Vorteil mündlicher Prüfungen ist, dass Nachfragen gestellt und Denkhilfen gegeben werden können, um individueller auf den Prüfling einzugehen. Der Nachteil ist jedoch die geringere Vereinheitlichung der Prüfungsbedingungen, weil sie die Objektivität und Reliabilität dieses Formats einschränkt.

Praktische Prüfungen zielen darauf ab, praxisbezogene bzw. professionsnahe Leistungen in Form z. B. eines praxisbezogenen Arbeitsauftrags oder im Rahmen von Simulationen zu überprüfen. Sie weisen damit einen ausgeprägteren Kompetenzbezug auf als andere Prüfungsformate. Um auch bei praktischen Prüfungen zu gewährleisten, dass sie vergleichbar, transparent und fair durchgeführt und ausgewertet werden, bedarf es einer genauen Vorbereitung (z. B. standardisierte Vorgaben) und Dokumentation der Durchführung (z. B. Protokollierung von Beobachtungen) sowie eindeutig definierter Bewertungskriterien und -regeln.

Zunehmend verbreitet sind auch digital gestützte Prüfungsformate, unter denen ganz allgemein Prüfungen mit digitalem Medieneinsatz zu verstehen sind. Neben der Umsetzung klassischer Formate in Form von Klausuren oder als mündliche Prüfungen per Videokonferenz finden sich hier auch E-Portfolios, computergestützte Simulationen, gemeinsames Schreiben von Wiki-Beiträgen, Video-basierte Präsentation von Aufgabenszenarios etc. Digitale Formate bieten nicht nur Effizienzvorteile, da sie die Ausführung von Routineaufgaben wie die Bewertung, Analyse und Archivierung der Prüfungen erleichtern, sie haben auch das Potenzial, einen qualitativen Mehrwert zu erzeugen (Franke/Handke 2012; Butler-Henderson/Crawford 2021): So erlaubt die multimediale Aufbereitung von Aufgabenszenarien bei digitalen Prüfungen bspw. eine realitätsnahe Darstellung von Praxissituationen (indem z. B. Fallstudienszenarien video-basiert präsentiert werden).

Bewertungsformen zur Leistungsbewertung

Zentral beim Prüfen ist die Frage, wie eine Prüfungsleistung bewertet werden kann. Grundsätzlich geht es bei dieser Phase darum, die gezeigten Prüfungsleistungen mit den Erwartungen zu vergleichen, die im Vorfeld der Prüfung in Bezug auf bestimmte Anforderungen bzw. Learning Outcomes formuliert wurden. Gemäß dem Zweck der Prüfung, der Art der erwarteten Leistung und des Prüfungsformats sollte hierzu ein Beurteilungsschema entwickelt werden. Ein solches Schema erfordert die Bestimmung beurteilungsrelevanter Kriterien, die für eine transparente und objektivere Beurteilung herangezogen werden können.

Um eine transparente und von subjektiven Einflüssen reduzierte Bewertung von Prüfungsleistungen zu gewährleisten, empfiehlt es sich, entweder ein detailliertes Bewertungsschema, bei dem alle bewertungsrelevanten Aspekte genau beschrieben bzw. vorgegeben werden, oder sog. Rubrics zu verwenden (Schaper 2021). Rubrics sind Bewertungsschemata, die eine Liste von kompetenzbezogenen Bewertungskriterien und für jedes Kriterium eine Beschreibung von Kompetenzstufen beinhalten. Rubrics werden nicht nur zur Bewertung komplexer Lern- bzw. Prüfungsleistungen (z. B. Projektarbeiten oder Abschlussarbeiten) herangezogen. Sie unterstützen auch den Lernprozess, indem sie Lernenden vermitteln, worauf es bei der Bewältigung kompetenzbezogener Anforderungen ankommt (Jonsson/Svingby 2007; Stevens/Levi 2012).

Darüber hinaus wird bei Beurteilungsverfahren zwischen Fremd- und Selbstbewertungen differenziert, wobei die Fremdbewertungen weiter untergliedert werden in solche von Lehrkräften oder von Peers, d. h. anderen Studierenden. In der Regel erfolgen Fremdbewertungen durch Lehrkräfte als Verantwortliche für alle Durchführungsaspekte eines Studienangebots. Bei Verfahren der Selbstbewertung beurteilen die Studierenden ihre eigene Leistung anhand vorgegebener Kriterien. Damit fördern sie bei Studierenden Fähigkeiten, ihre eigenen Stärken und Schwächen systematischer zu erkennen und eigenständig Verbesserungsstrategien zu entwickeln. Bei Peer-Assessment-Methoden sollen die Studierenden hingegen die Leistungen ihrer Mit-Studierenden anhand festgelegter Kriterien beurteilen (z. B. bei der Präsentation von Projektarbeiten). Peer- und Selbstbewertungen werden zunehmend in die Lehr- und Prüfungsgestaltung mit einbezogen, um die Auseinandersetzung mit den Kompetenzanforderungen in einem Lerngebiet und damit den Lernprozess zu fördern, was nachweislich gelingt (siehe z. B. Double et al. 2019), weniger um daraus verbindliche Bewertungen für das Bestehen einer Lehr-/Lerneinheit abzuleiten.

Didaktisch-diagnostische Ausrichtungen der Prüfungsgestaltung

Die Konzeption von hochschulischen Prüfungen sollte auch vor dem Hintergrund theoretisch fundierter Ansätze der Prüfungsgestaltung vorgenommen werden. Im Folgenden werden vier Ansätze vorgestellt, die grundlegende didaktische und diagnostische Konzepte der Prüfungsgestaltung in der Hochschullehre beschreiben.

Testtheoretisch orientierte Prüfungsgestaltung

Ansätze der testtheoretisch orientierten Prüfungsgestaltung verstehen den Prozess des Prüfens vor allem als einen diagnostischen Vorgang. Es geht darum, den Leistungs- bzw. Lernstand von Studierenden in einem bestimmten Studiumskontext möglichst genau und inhaltlich angemessen zu erfassen. Hochschulprüfungen können demgemäß als kriteriumsorientierte Tests aufgefasst werden, da ihnen die Aufgabe zukommt, das Erreichen von Lernzielen bzw. Learning Outcomes zu überprüfen (Klapproth 2023). Dies erfordert die Festlegung eines Kriteriums bzw. einer kritischen Leistung, die erbracht werden muss, damit ein Lernziel als erreicht bewertet werden kann. Das Kriterium bezieht sich dabei auf eine Menge von Aufgaben, die die Fähigkeit zur Lösung von Aufgaben in einer

umgrenzten inhaltlichen Domäne umfasst. Die jeweils ermittelten Testwerte sind als erwartungstreue Schätzer entsprechender Fähigkeiten zu interpretieren.

Prüfungen bzw. Testaufgaben sind daher unter der diagnostischen Perspektive so zu gestalten, dass sie diesen Anspruch an einen kriteriumsorientierten Test möglichst zuverlässig (reliabel) und valide im Sinne der inhaltlichen Validität erfüllen. In erster Linie kommen dafür schriftliche Aufgaben in strukturierter und offener Form in Frage. So wird den Prüfenden z. B. nahegelegt, nach der Durchführung der Prüfungen Item- und Testgütekriterien (z. B. Item-Schwierigkeit, Trennschärfe des Items, Kennwerte der Testreliabilität) numerisch zu bestimmen, um die Güte der Prüfung bzw. des kriteriumsorientierten Tests nachzuweisen und bei Bedarf zu verbessern (z. B. durch Selektion zu schwerer oder zu leichter Aufgaben). Allerdings sind solche Ansätze nur dann handhabbar, wenn die Prüfenden selbst testtheoretische Kenntnisse und Fähigkeiten haben oder sie auf entsprechende Expertise zurückgreifen können.

Kompetenzorientierte Prüfungsgestaltung

Eine kompetenzorientierte Gestaltung von Lehre ist darauf ausgerichtet, den Erwerb von Kompetenzen im Studium zu fördern, d. h. Fähigkeiten, die über Wissensreproduktion und -verständnis hinausgehen und insbesondere die Anwendung von fachlichem Wissen, Fähigkeiten zur Analyse und Beurteilung von komplexen Sachverhalten sowie zum Problemlösen und zur Zusammenarbeit bei komplexen Problemen beinhalten. Didaktisch wird dies anhand von kompetenzorientierten Learning Outcomes sowie einer konsequenten Ausrichtung der Lerninhalte und -aufgaben sowie der Prüfungsanforderungen auf diese Outcomes realisiert. Die Gestaltung kompetenzorientierter Prüfungen ist ein wesentliches didaktisches Element bei diesem Ansatz.

Zur inhaltlichen Konzeption von kompetenzorientierten Prüfungsanforderungen sollte auf das Constructive Alignment Konzept Bezug genommen werden, auch wenn dies kontrovers diskutiert wird (z. B. Reinmann 2018). Hierbei geht es um den Abgleich der Learning Outcomes mit den Lerninhalten und -aufgaben sowie Prüfungsanforderungen eines Lehrangebots und die Überprüfung und Anpassung der Kompatibilität dieser drei Elemente (Biggs/Tang 2011). Bei der Konzeption von kompetenzorientierten Prüfungen unter Bezugnahme auf das Constructive Alignment Konzept sind auch die Bewertungskriterien und -methoden im Hinblick auf ihre Passung zu den zu überprüfenden Learning Outcomes zu analysieren (Schaper 2024). Hierdurch wird gewährleistet, dass die in einem Lehrkontext anzueignenden Kompetenzen auf einem angemessenen Niveau abgeprüft werden und die didaktische Kohärenz von Lernzielen, Lehrangebot und Prüfung sichergestellt wird. Außerdem wird eine wichtige Grundlage für die Transparenz der Lern- und Prüfungsanforderungen hergestellt, und es werden eindeutige Anreize, sich mit den Lernanforderungen im Kurs ernsthaft auseinanderzusetzen, geschaffen (Biggs/Tang 2011). Die Abstimmung der verschiedenen Elemente wirkt sich auch auf die Qualität des Lernprozesses aus und fördert nachgewiesenermaßen die Lernmotivation, Lernstrategien für ein vertiefteres Lernen und das Erreichen anspruchsvollerer Lernziele (Biggs 2014; Hailikari et al. 2022).

Lernorientierte Prüfungsgestaltung

Die Durchführung von Prüfungen im Studium wird bei Ansätzen der lernorientierten Prüfungsgestaltung als Teil des Lern- bzw. Kompetenzerwerbprozesses verstanden. Prüfungen sind in diesem Zusammenhang nicht nur Ziel des Lernens im Sinne des Bestehens der Prüfung bzw. dem Erhalt einer möglichst guten Bewertung. Vielmehr sind sie auch als Teil des Kompetenzerwerbs zu gestalten, d. h., dass bei der Bewältigung der Prüfungsanforderungen noch gelernt wird und man seine Kompetenzen entwickelt. Dies gelingt vor allem dann, wenn die Lernenden ein zeitnahes und informatives Feedback zu den Prüfungsleistungen erhalten (Narciss 2014; Wisniewski et al. 2020) und formative Assessment-Elemente z. B. in Form von Testaten oder Rückmeldungen zu Übungsaufgaben im Lernprozess eingesetzt werden (Evans 2013; Morris et al. 2021).

Partizipationsorientierte Prüfungsgestaltung

Bei dem Ansatz der partizipationsorientierten Prüfungsgestaltung wird davon ausgegangen, dass es förderlich für den Lernprozess der Studierenden ist, wenn sie in den Bewertungsprozess mit eingebunden werden. Traditionell nehmen Studierende in Prüfungs- und Bewertungsprozessen eine wenig selbstbestimmte Rolle ein, da ihre Aufgabe auf das Absolvieren der Prüfungen beschränkt ist. Es wird daher vorgeschlagen, eine aktivere Einbindung in den Bewertungsprozess zu fördern, indem die Studierenden aufgefordert werden, sich selbst oder ihre Mitstudierenden anhand entsprechender Selbst- und Peer-Assessment-Methoden zu beurteilen (Falchikov 2003). Evaluationsstudien hierzu zeigen, dass die Studierenden durch die aktive Beteiligung am Bewertungsprozess lernen, sich kritisch mit ihren eigenen Leistungen und denen ihrer Peers auseinanderzusetzen, was in eine intensivere und tiefere Bearbeitung der Lerninhalte und -anforderungen resultiert (Topping 2010; Pereira et al. 2016). Außerdem führt die Einbindung in den Bewertungsprozess zu einem besseren Verständnis und einer höheren Akzeptanz der Bewertungen, da die Bewertungsmaßstäbe und -kriterien durch die aktive Beteiligung am Prozess als transparenter und nachvollziehbarer wahrgenommen werden. Dies setzt allerdings voraus, dass die Studierenden und Lehrenden sorgfältig auf die Handhabung der Methoden vorbereitet bzw. geschult werden (Flores et al. 2014).

Evaluation und Qualitätssicherung bei der Prüfungsgestaltung

Nicht nur die Hochschullehre, sondern auch Prüfungen sollten Gegenstand von Evaluation und Qualitätsmanagement sein. Im Folgenden werden Evaluationsstudien zu hochschulischen Prüfungen sowie Kriterien und Ansätze der Qualitätssicherung bei Prüfungen vorgestellt.

Wahrnehmung und Einstellungen zu Prüfungen

Im Unterschied zur Evaluation von Lehre (→ II.2.5 Lehrevaluation und Qualitätsentwicklung), werden Gestaltungsansätze zu Prüfungen in der Hochschullehre nur wenig evaluiert. In der Regel finden sich dazu Befragungen von Studierenden, teilweise auch Lehrenden oder auch beide Zielgruppen. Im Fokus solcher Befragungen stehen seit den 2000er Jahren Assessment-Ansätze, die im Kontext der Bologna-Reform entstanden sind. Einen Überblick zu solchen Studien, die sich mit Bewertungsmethoden und -modi und deren Zusammenhängen mit veränderten Lehr-/Lernkonzepten (insbesondere studierendenzentrierten Ansätzen) und den damit verknüpften Lehr-/Lernmethoden befassen, geben z. B. Pereira et al. (2016). Die Auswertung der Studien zeigt, dass weiterhin auch traditionelle Prüfungsformate wie schriftliche Klausuren und mündliche Prüfungen im Vordergrund stehen, aber vermehrt alternative Assessment-Methoden wie Portfolios, Gruppenbewertungen bei Präsentationen oder digitale Tagebuch-Methoden eingesetzt werden. Die Bewertungen von Studierenden verdeutlichen insgesamt, dass Portfolios und Bewertungen von Gruppenleistungen nicht nur als effektiver und fairer wahrgenommen werden im Vergleich zu traditionellen Prüfungsformaten, sondern zudem ein vertiefteres und kritisch-analytisches Denken fördern (s.a. Flores et al. 2014).

Auch im deutschen Kontext wurden Studien zur evaluativen Bewertung von hochschulischen Prüfungsformen durchgeführt. Beispielsweise haben Lindner et al. (2019) Einstellungen und Präferenzen von Psychologie-Studierenden zu geschlossenen (Multiple-Choice-Aufgaben) und offenen (Freitextaufgaben) Prüfungsaufgaben erhoben und miteinander verglichen. Die Ergebnisse zeigen, dass Studierende Multiple-Choice-Aufgaben als weniger lernaufwendig wahrnehmen, die Erfolgschancen höher einschätzen und als objektiver und fairer beurteilen. Freitextaufgaben wird hingegen ein höheres Potenzial zugeschrieben, um die eigene Leistung und das Wissen unter Beweis zu stellen. Die Auswertungen zu den individuellen Präferenzen für die jeweiligen Prüfungsformate zeigen zudem, dass Studierende, die eher oberflächliche Lernstrategien verfolgen, Multiple Choice Aufgaben bevorzugen, während jene, die ein tieferes Verständnis anstreben, Freitextaufgaben präferieren.

Qualitätskriterien und Qualitätsmanagement

Prüfungsformate und -aufgaben in der Hochschullehre sowie deren Auswertung sollten bestimmten Anforderungen bzw. Qualitätsansprüchen gerecht werden. In der einschlägigen Literatur werden hierzu in erster Linie die psychometrischen Gütekriterien der pädagogisch-psychologischen Diagnostik genannt (Objektivität, Reliabilität und Validität) (Metzger/Nüesch 2004). Unter dem Konzept der Objektivität werden Maßnahmen zusammengefasst, die dazu beitragen, Ursachen für Messfehler zu reduzieren, indem verschiedene Aspekte im Prüfungsprozess standardisiert werden. Dabei wird zwischen Durchführungsobjektivität (stellt sicher, dass alle Prüflinge unter denselben Bedingungen geprüft werden), Auswertungsobjektivität (ist erfüllt, wenn unterschiedliche Bewertende bei der Klassifizierung von Prüfungsantworten zum gleichen Ergebnis kommen) und Interpretationsobjektivität (gewährleistet, dass unterschiedliche Bewertende die gleichen Schlüsse aus dem Testergebnis ziehen) unterschieden (Bühner 2011).

Bei dem Konzept der Reliabilität steht die Zuverlässigkeit einer Messung im Vordergrund (ebd.). Eine hohe Reliabilität liegt z. B. vor, wenn bei mehreren Messungen eines bestimmten Merkmals bei einer Person keine Unterschiede in den Messwerten auftreten, die durch die Messung bedingt sind. Sollten dennoch Unterschiede auftreten, sind diese auf Messfehler bzw. Messungenauigkeiten zurückzuführen. Das Konzept der Validität bezieht sich schließlich darauf, ob mit den Prüfungsaufgaben gültige Rückschlüsse auf das Vorhandensein von Kompetenzen und kompetenzrelevantem Wissen bei Studierenden gezogen werden können (Schaper 2014). Hierzu werden verschiedene Validitätsaspekte unterschieden: Die inhaltliche Validität einer Prüfung ist z. B. gewährleistet, wenn in den Prüfungsaufgaben ausschließlich die für einen Lernbereich relevanten Inhalte und Anforderungssituationen enthalten sind. In Bezug auf die kognitive Validität einer Prüfung gilt es zu analysieren, ob die Studierenden zur Lösung der Prüfungsaufgaben tatsächlich auf die für eine Kompetenz relevanten kognitiven Prozesse zurückgreifen müssen.

Bei der Konzeption und Umsetzung von Prüfungen in der Hochschullehre sind neben den klassischen psychometrischen Gütekriterien weitere Aspekte zu berücksichtigen: Von Bedeutung ist beim Prüfen insbesondere, dass die Prüfungsanforderungen transparent für die Studierenden sind. Hierzu sollten ausführliche, nachvollziehbare und wiederholte Information über Rahmenbedingungen, Formate, Aufgaben und Anforderungen einer Prüfung vermittelt werden. Darüber hinaus sind Prüfungen fair zu gestalten, d. h., bestimmte Teilgruppen der Prüflinge dürfen durch die Art der Prüfung nicht benachteiligt werden (z. B. in Bezug auf sprachliche Voraussetzungen). Im Hinblick auf die Akzeptanz des Prüfungsformats sollten diese von den Studierenden als angemessen in Bezug auf Inhalte, Umfang und Bewertungsverfahren wahrgenommen werden.

Zur Sicherung von Qualität und Angemessenheit einer Prüfung sind aber nicht nur die genannten Ergebniskriterien, sondern auch verschiedene prozessuale Aspekte zu berücksichtigen. Hierzu gehört z. B., dass Prüflinge sich angemessen auf die Prüfung vorbereiten können (etwa indem ausreichende Informationen über Prüfungsanforderungen oder angemessenes Prüfungsverhalten gegeben werden) und Prüfende eine angemessene Vorbereitung (und Schulung) erhalten (z. B. zum Gesprächsverhalten in mündlichen Prüfungen).

Fazit und Ausblick

Es sollte deutlich geworden sein, dass Prüfungen nicht nur unterschiedliche Funktionen im Kontext von Studium und Lehre erfüllen, sondern auch unterschiedlichsten Anforderungen gerecht werden müssen. Anhand eines Rahmenmodells ließ sich der Prüfungsprozess zudem als mehrphasig beschreiben, in dem eine Vielzahl von gestaltungs- und qualitätsrelevanten Entscheidungen getroffen werden muss. Allerdings wird die Vielfalt der zur Verfügung stehenden Prüfungsformate und Bewertungsformen oftmals nur unzureichend genutzt, was nicht nur zu Monokulturen der Prüfungsgestaltung führt, sondern auch zur Folge hat, dass das Potenzial zur diagnostischen und didaktischen Gestaltung der Prüfungsaufgaben nicht ausgeschöpft wird. Dies verdeutlicht, dass Prüfungen in vielen Fällen immer noch nicht oder nur unzureichend in die didaktische Gesamtplanung des Lehr-/Lernsettings integriert werden.

Bei der Prüfungsgestaltung gilt es außerdem, unterschiedliche konzeptionelle Ansätze im Hinblick auf die didaktische und diagnostische Gestaltung von Prüfungen zu berücksichtigen. Anhand dieser Ansätze wird deutlich, dass die Gestaltung von Prüfungen nicht nur mit diagnostisch-methodischen Fragen verknüpft ist, sondern maßgeblich auch die Berücksichtigung inhaltlich didaktischer Konzepte (z. B. kompetenzorientierte Prüfungsgestaltung) sowie lerntheoretischer Prinzipien erfordert, und auch die Einbettung von Prüfungen in die jeweilige Lehr-/Lernphilosophie (z. B. partizipationsorientierte Prüfungsgestaltung) eine wesentliche Rolle spielt. Studien zur Evaluation von hochschulischen Prüfungen zeigen zudem, dass die Wahrnehmung und Beurteilung von Prüfungen nicht nur vor dem Hintergrund von diagnostischen Qualitätsmerkmalen vorgenommen werden (Flores et al. 2014; Pereira et al. 2016; Meier et al. 2023). Daher sind neben testtheoretischen Qualitätsmerkmalen auch Transparenz, Fairness, Handhabbarkeit und Akzeptanzaspekte sowie prozessbezogene Aspekte bei der Prüfungsgestaltung zu berücksichtigen.

Zum Stand der Prüfungsforschung in der Hochschuldidaktik ist jedoch insgesamt festzustellen, dass dieser in vieler Hinsicht noch wenig entwickelt ist. Dies gilt insbesondere in Bezug auf die Überprüfung von – nicht nur messtheoretischen – Qualitätskriterien von Prüfungsaufgaben sowie verbreiteter Auswertungsansätze, aber auch in Bezug auf Maßnahmen für eine lernförderliche Gestaltung von Prüfungen bzw. Ansätze zur wirkungsvollen Einbettung von Prüfungen in den Lernprozess. Nicht zuletzt ist auch auf den theoriebezogenen Entwicklungsbedarf, z. B. einer hochschuldidaktischen Theorie des Prüfens, hinzuweisen.

Literaturempfehlungen

Schaper, Niclas (2024): Didaktik und Theorie kompetenzorientierter Prüfungen. In: Vöing, Nerea (Hg.), Praxishandbuch Hochschullehre I. Grundlagen, Methoden und Lehrformate. Stuttgart: transcript/utb. *Der Handbuchbeitrag vermittelt einen Überblick zu zentralen theoretischen und praxisorientierten Konzepten des kompetenzorientierten Prüfens in der Hochschullehre.*

Morris, Rebecca/Perry, Thomas/Wardle, Lindsey (2021): Formative assessment and feedback for learning in higher education: A systematic review. In: Review of Education 9, e3292. DOI: 10.1002/rev3.3292. *Das narrative Literaturreview von Morris et al. (2021) untersucht die Auswirkungen von formativen Bewertungs- und Feedbackansätze auf Lernen und Leistungen Studierenden in der Hochschulbildung und liefert eine systematische Übersicht zur Wirksamkeit solcher Ansätze.*

Reinmann, Gabi (2019): Forschendes Lernen prüfen: Hochschuldidaktische Gedanken zu einer Theorie des Prüfens. In: Zeitschrift für Pädagogik 65(4), 608–626. *Der Aufsatz beschreibt zentrale Widersprüche und Schwierigkeiten beim hochschulischen Prüfen, und versucht exemplarisch anhand des Prüfens in Kontexten des forschenden Lernens Impulse für eine hochschuldidaktische Theorie des Prüfens zu entwickeln.*

Literaturverzeichnis

Biggs, John B. (2014): Constructive alignment in university teaching. In: HERDSA Review of Higher Education 1, 5–22, www.herdsa.org.au/herdsa-review-higher-education-vol-1/5-22 (07.04.2025).

Biggs, John B./Tang, Catherine (2011): Teaching for quality learning at University, 4. Aufl. Oxford: Oxford University Press.

Bühner, Markus (2011): Einführung in die Test-und Fragebogenkonstruktion. Upper Saddle River, NJ: Pearson Studium.

Butler-Henderson, Kerryn/Crawford, Joseph (2021): A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. In: Computers & Education 159, 104024. DOI: 10.1016/j.compedu.2020.104024.

Double, Kit S./McGrane, Joshua A./Hopfenbeck, Therese N. (2020): The impact of peer assessment on academic performance: A Meta-analysis of control group studies. In: Educational Psychology Review 32, 481–509. /doi.org/10.1007/s10648-019-09510-3.

Evans, Carol (2013): Making sense of assessment feedback in higher education. In: Review of Educational Research 83(1), 70–120. journals.sagepub.com/doi/abs/10.3102/0034654312474350.

Falchikov, Nancy (2003): Involving Students in Assessment. In: Psychology Learning and Teaching 3(2), 102–108.

Flores, Maria A./Veiga Simão, Ana M./Barros, Alexandra/Pereira, Diana (2014): Perceptions of effectiveness, fairness and feedback of assessment methods: a study in higher education. In: Studies in Higher Education. DOI: 10.1080/03075079.2014.881348.

Franke, Peter/Handke, Jürgen (2012): E-Assessment. In: Handke, Jürgen/Schäfer, Anna Maria (Hg.), E-Learning, E-Teaching und E-Assessment in der Hochschullehre. Eine Anleitung. München: Oldenbourg, 147–208. doi.org/10.1524/9783486716849.147.

Hailikari, Telle/Virtanen, Viivi/Vesalainen, Marjo/Postareff, Liisa (2022): Student perspectives on how different elements of constructive alignment support active learning. In: Active Learning in Higher Education 23(3), 217–231. doi.org/10.1177/1469787421989160.

Hattie, John/Beywl, Wolfgang/Zierer, Klaus (2013): Lernen sichtbar machen. Überarb. deutsche Ausgabe von ‚Visible Learning'. Baltmannsweiler: Schneider.

Jonsson, Anders/Svingby, Gunilla (2007): The use of scoring rubrics: Reliability, validity, and educational consequences. In: Educational Research Review 2(2), 130–144.

Klapproth, Florian (2023): Von den Lehrzielen zur schriftlichen Prüfung. Ein Leitfaden für Lehrende der Psychologie. Göttingen: Hogrefe.

Lindner, Marlit A./Mayntz, Sonja/Schult, Johannes (2019): Studentische Bewertung und Präferenz von Hochschulprüfungen mit Aufgaben im offenen und geschlossenen Antwortformat. In: Zeitschrift für Pädagogische Psychologie 32(4), 239–248. DOI: 10.1024/1010-0652/a000229.

Meier, Jana/Janzen, Thomas/Wotschel, Philipp/Vogelsang, Christoph (2023): Rollenspielbasierte Simulationen als Übungs- und Prüfungsformate im Lehramtsstudium. Eine explorative Studie zu Erfahrungen und Einschätzungen aus Studierendensicht. In: die hochschullehre 9(7), 85–100, https://doi.org/10.3278/HSL2307W.

Metzger, Christoph/Nüesch, Charlotte (2004): Fair prüfen. Ein Qualitätsleitfaden für Prüfende an Hochschulen. St. Gallen: Institut für Wirtschaftspädagogik.

Morris, Rebecca/Perry, Thomas/Wardle, Lindsey (2021): Formative assessment and feedback for learning in higher education: A systematic review. In: Review of Education 9, e3292. DOI: 10.1002/rev3.3292.

Narciss, Susanne (2014): Modelle zu den Bedingungen und Wirkungen von Feedback in Lehr-/Lernsituationen. In: Müller, Andreas/Ditton, Hartmut (Hg.), Rückmeldungen und Feedback: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder. Münster: Waxmann, 43–82.

Pereira, Diana/Flores, Maria A./Niklasson, Laila (2016): Assessment revisited: a review of research in Assessment and Evaluation in Higher Education. In: Assessment & Evaluation in Higher Education 41(7), 1008–1032. DOI: 10.1080/02602938.2015.1055233.

Reinmann, Gabi (2018): Shift from Teaching to Learning und Constructive Alignment: Zwei hochschuldidaktische Prinzipien auf dem Prüfstand. Impact Free 14. Hamburg: Universität Hamburg.

Reinmann, Gabi (2019): Forschendes Lernen prüfen: Hochschuldidaktische Gedanken zu einer Theorie des Prüfens. In: Zeitschrift für Pädagogik 65(4), 608–626.

Schaper, Niclas (2014): Validitätsaspekte von Kompetenzmodellen und -tests für hochschulische Kompetenzdomänen. In Musekamp, Frank/Spöttl, Georg (Hg.), Kompetenz im Studium und in der Arbeitswelt. Nationale und internationale Ansätze zur Erfassung von Ingenieurkompetenzen. Frankfurt a. M.: Peter Lang, 178–203.

Schaper, Niclas (2021): Prüfen in der Hochschullehre. In: Kordts-Freudinger, Robert/Schaper, Niclas/Scholkmann, Antonia/Szczyrba, Birgit (Hg.), Handbuch Hochschuldidaktik. Bielefeld: wbv/utb, 87–102.

Schaper, Niclas (2024): Didaktik und Theorie kompetenzorientierter Prüfungen. In: Vöing, Nerea (Hg.), Praxishandbuch Hochschullehre I. Grundlagen, Methoden und Lehrformate. Stuttgart: transcript/utb.

Schaper, Niclas/Soyka, Chantal (2021): Kompetenzorientiertes Prüfen. Grundlagen, präsenz- und onlinegestützte Formate, Bewertung und Rückmeldung kompetenzorientierter Prüfungsleistungen. In: Berendt, Brigitte/Fleischmann, Andreas/Salmhofer, Gudrun/Schaper, Niclas/Sczcyrba, Birgit/Wildt, Johannes (Hg.), Neues Handbuch Hochschullehre, H 6.8. Stuttgart: Steiner.

Schindler, Christoph/Bauer, Johannes/Strasser, Alexandra/Schlomske-Bodenstein, Nadine/Seidel, Tina/Prenzel, Manfred (2015): Prüfungen als Indikatoren für Studienerfolg. In Berthold, Christian/Jorzik, Bettina/Meyer-Guckel, Volker (Hg.), Handbuch Studienerfolg. Essen: Stifterverband für die Deutsche Wissenschaft, 62–79.

Stevens, Dannelle D./Levi, Antonia J. (2013): Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning. Sterling: Stylus Publishing.

Topping, Keith J. (2010): Methodological quandaries in studying process and outcomes in peer assessment. In: Learning and Instruction 20(4), 339–343. DOI: 10.1016/j.learninstruc.2009.08.003.

Wisniewski, Benedikt/Zierer, Klaus/Hattie, John (2020): The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. In: Frontiers in Psychology 10, 1–14. doi.org/10.3389/fpsyg.2019.03087.