



**Jg. 17 / Nr. 1 (März 2022)**

**Karin Sonnleitner & Martin Gartmeier (Hrsg.)**

**Prüfen im Kontext kompetenzorientierter  
Hochschulbildung**

**Prüfen im Kontext  
kompetenzorientierter Hochschulbildung**

**Karin Sonnleitner & Martin Gartmeier (Hrsg.)**

**Prüfen im Kontext  
kompetenzorientierter Hochschulbildung**

**Zeitschrift für Hochschulentwicklung  
Jg. 17 / Nr. 1 (März 2022)**

# **Impressum**

**Zeitschrift für Hochschulentwicklung**

**Jg. 17 / Nr. 1 (März 2022)**

**Prüfen im Kontext kompetenzorientierter Hochschulbildung**

herausgegeben vom Verein Forum Neue Medien in der Lehre Austria  
Graz, 2022

**Herausgeber**

**Karin Sonnleitner & Martin Gartmeier**

**ISBN**

9783755740636

**Druck und Verlag**

Books on Demand GmbH, Norderstedt

# Inhalt

Vorwort .....	7
Editorial: Prüfen im Kontext kompetenzorientierter Hochschulbildung .....	9
<i>Karin Sonnleitner &amp; Martin Gartmeier</i>	
Was Prüfungen leisten sollen: Prüfungen für akademische Kompetenz .....	17
<i>Sabrina Gallner</i>	
Prüfungsmisserfolg auf die Agenda! Ansätze an der Universität Wien .....	35
<i>Heidi Niederkofler</i>	
Kompetenzorientiertes Lehren und Prüfen in den Ingenieurwissenschaften: Lehrkonzept der Veranstaltung Materialfluss in Logistiksystemen .....	47
<i>Uta Mohring, Christoph Jacobi &amp; Kai Furmans</i>	
Die Prüfung von interdisziplinären Kompetenzen im Rahmen des Prinzips des Constructive Alignments .....	69
<i>Mirjam Braßler</i>	
Multimodal take-home exams in online teaching and beyond: constructive and professional alignment in teacher education .....	87
<i>Silke Braselmann, Jolene Mathieson &amp; Oliver Moisch</i>	
Online Two-stage Exams als Prüfungsmethode in Statistik-Lehrveranstaltungen .....	103
<i>Bernhard Spangl, Dóra Kértész &amp; Christian F. Freisleben-Teutscher</i>	
Kompetenzorientierung und Fairness bei individualisierten E-Klausuren .....	121
<i>Christian Spoden, Aron Fink, Andreas Frey, Hanna Köhler &amp; Patrick Naumann</i>	
Online-Midterms zwischen formativer und summativer Bewertung .....	141
<i>Miriam Clincy, Karin Melzer, Gunther Schaaf, Achim Eichhorn &amp; Nathalie Verné</i>	

Ich packe meinen Koffer: Unterlagen und Performanz bei Open-Book-Klausuren .....	157
<i>Sandra Niedermeier, Michael Sailer, Lukas Remböck &amp; Matthias Stadler</i>	
Designing a case study approach for competence-oriented examination in large courses .....	179
<i>Eva Osteresch</i>	
Konstruktion und Durchführung eines Video-Portfolios. Eine innovative schulpraktische Prüfung .....	201
<i>David Bisang, Kerstin Bäuerlein &amp; Corinne Wyss</i>	
Simulationen in virtuellen Lernumgebungen: Welche Vor- und Nachteile haben avatarbasierte und videokamerabasierte Formate bei der Durchführung von Rollenspielen? .....	215
<i>Julia Fecke &amp; Lars Müller</i>	
Open Educational Resources (OER) in der wissenschaftlichen Weiterbildung .	233
<i>Olaf Zawacki-Richter, Wolfgang Müskens, Nadine Dembski &amp; Sonja Lübben</i>	

# Vorwort

Als wissenschaftliches Publikationsorgan des Vereins Forum Neue Medien in der Lehre Austria kommt der Zeitschrift für Hochschulentwicklung besondere Bedeutung zu. Zum einen, weil sie aktuelle Themen der Hochschulentwicklung in den Bereichen Studien und Lehre aufgreift und somit als deutschsprachige, vor allem aber auch österreichische Plattform zum Austausch für Wissenschaftler:innen, Praktiker:innen, Hochschulentwickler:innen und Hochschuldidaktiker:innen dient. Zum anderen, weil die ZFHE als Open-Access-Zeitschrift konzipiert und daher für alle Interessierten als elektronische Publikation frei und kostenlos verfügbar ist.

Ca. 3.000 Besucher:innen schauen sich im Monat die Inhalte der Zeitschrift an. Das zeigt die hohe Beliebtheit und Qualität der Zeitschrift sowie auch die große Reichweite im deutschsprachigen Raum. Gleichzeitig hat sich die Zeitschrift mittlerweile einen fixen Platz unter den gern gelesenen deutschsprachigen Wissenschaftspublikationen gesichert.

Dieser Erfolg ist einerseits dem international besetzten Editorial Board sowie den wechselnden Herausgeberinnen und Herausgebern zu verdanken, die mit viel Engagement dafür sorgen, dass jährlich mindestens vier Ausgaben erscheinen. Andererseits gewährleistet das österreichische Bundesministerium für Wissenschaft, Forschung und Wirtschaft durch seine kontinuierliche Förderung das langfristige Bestehen der Zeitschrift. Im Wissen, dass es die Zeitschrift ohne diese finanzielle Unterstützung nicht gäbe, möchten wir uns dafür besonders herzlich bedanken.

Zur Ausgabe:

Prüfungen sind zentraler Bestandteil des Alltags von Lehrenden und Studierenden an Hochschulen und Universitäten. Folglich müssen bei der Diskussion rund um die Qualität des Lehrens und Lernens auch Prüfungen berücksichtigt werden. Kompetenzorientiertes Prüfen an Universitäten bedeutet, die Frage zu adressieren, ob Studierende Kompetenzen in einer Art und Weise entwickelt haben, die sie zur Lösung relevanter Probleme befähigt. Durch die COVID-19-Pandemie waren viele Lehrende gezwungen, ihre Lehre umzustrukturieren und Online-Formate ebenfalls für Prüfungen zu nutzen. Dadurch haben schon länger virulente wissenschaftliche Diskurse zu Online-Lehre und zu Online-Prüfungen ein neues Momentum und neue Brisanz erhalten. In 13 Beiträgen beschäftigen sich die Autor:innen mit theoretischen Grundlagen unterschiedlicher Prüfungsformate und nehmen Bezug auf praktische Beispiele unterschiedlicher Fachbereiche.

Seit der Ausgabe 9/3 ist die ZFHE auch in gedruckter Form erhältlich und beispielsweise über Amazon beziehbar. Als Verein Forum Neue Medien in der Lehre Austria freuen wir uns, das Thema „Hochschulentwicklung“ durch diese gelungene Ergänzung zur elektronischen Publikation noch breiter in der wissenschaftlichen Community verankern zu können.

In diesem Sinn wünschen wir Ihnen viel Freude bei der Lektüre der vorliegenden Ausgabe!

Martin Ebner und Hans-Peter Steinbacher  
Präsidenten des Vereins Forum Neue Medien in der Lehre Austria

## Editorial: Prüfen im Kontext kompetenzorientierter Hochschulbildung

Prüfungen sind elementare Bestandteile des Lehrens und Lernens an Hochschulen (WISSENSCHAFTSRAT, 2012). Für Absolvent\*innen sind Noten und Leistungspunkte essentiell, da Arbeitgeber\*innen diese als wichtige Indikatoren für die Leistungsfähigkeit und die Kompetenz nutzen. Gleichzeitig sind Prüfungen in enger Verschränkung mit den Lehrveranstaltungen zu betrachten, deren Inhalte während des Semesters (formativ) oder zu Semesterende (summativ) thematisiert bzw. abgefragt werden: So determinieren Prüfungen u. a., auf welche Lehrinhalte sich Studierende in ihrer Vorbereitung konzentrieren und auf welche Art und Weise sie sich vorbereiten (SCHINDLER, 2015). Der letztere Zusammenhang ist besonders im Kontext aktueller Entwicklungen in Richtung einer stärkeren Kompetenzorientierung in der Hochschulbildung relevant.

Dieser Trend bedeutet eine Abkehr der Sichtweise von Universitäten als Instanzen der (reinen) Vermittlung von Wissen, das die Absolvent\*innen in einer zeitlich nachgelagerten Phase nutzen und anwenden. Neuere bildungstheoretische Konzeptionen des Arbeitens, der beruflichen und persönlichen Weiterentwicklung sowie der gesellschaftlichen Teilhabe in einer immer stärker globalisierten und digitalisierten Welt gehen davon aus, dass universitäre Bildung mehr umfasst als die Vermittlung von Wissen, das sodann in der Praxis anwendungsbezogen transformiert wird. Aktuelle, kompetenzorientierte Konzeptionen von universitärer Bildung nehmen stattdessen an, dass die Vermittlung von wissenschaftlichem Wissen und von Methoden zur Generierung dieses Wissens mit Fragen nach der Anwendung zur Lösung fachspezifischer sowie auch fachübergreifend relevanter Problemstellungen verschränkt

---

<sup>1</sup> E-Mail: [karin.sonnleitner@uni-graz.at](mailto:karin.sonnleitner@uni-graz.at)

werden sollten. Diese Sichtweise hat Konsequenzen für die Konzeption und Ausgestaltung sowohl von Lehrveranstaltungen als auch von Prüfungen.

Mit dem Anspruch, *kompetenzorientiert* zu prüfen, ist die Zielsetzung verbunden, Studierende in Prüfungen mit Aufgabenstellungen zu konfrontieren, die inhaltliche und/oder strukturelle Ähnlichkeiten mit komplexen Problemstellungen aufweisen, die in einschlägigen beruflichen Kontexten auftauchen. Einen konzisen Überblick über verschiedene Grundformen erlaubt die Unterscheidung kompetenzorientierter Prüfungsformen von SCHINDLER (2015): Bei *Kompetenztests* werden Verhaltensweisen von Prüflingen in relevanten Situationen unter Realbedingungen beobachtet – z. B. bei einer schulischen Lehrprobe, in der eine junge Lehrperson eine Schulklasse unterrichtet. Demgegenüber werden bei *kompetenzorientierten Tests* Situationen gestaltet, die Annäherungen an eine berufliche Praxis beinhalten, z. B. wenn ein Medizinstudent an einem anatomischen Modell eine neurologische Untersuchungstechnik demonstriert. Eine dritte Form kompetenzorientierter Prüfungen sind *kompetenzorientierte Wissenstests*, bei denen Wissenseinheiten oder kognitive Fähigkeiten kontextbezogen angewendet werden sollen, z. B. wenn eine Studentin im Studium der sozialen Arbeit ein theoriebasiertes Beratungskonzept für einen vorgegebenen Fall entwickelt.

Bereits anhand dieser grundlegenden Unterscheidung wird deutlich, dass kompetenzorientierte Prüfungen meist substanziell höheren Entwicklungsaufwand darstellen und hohe Ansprüche an Prüfende sowohl bei der Entwicklung als auch bei der Durchführung beinhalten. Weitere relevante Herausforderungen für die Hochschule betreffen die Nutzung moderner, digitaler Technologien und deren Stellenwert für formative und summative Prüfungsszenarien sowie das *Constructive Alignment* (BIGGS & TANG, 2011; BIGGS, 2014) kompetenzorientierter Lehrveranstaltungen und Prüfungen.

Mit dem Begriff *Constructive Alignment* ist auch schon die Brücke zum kompetenzorientierten Lehren geschlagen. Denn kompetenzorientiertes Lehren und Prüfen sollten sinnvoll miteinander kombiniert werden, um den Ansätzen des Bologna-Prozesses folgend die Kompetenzen Studierender umfassend zu fördern. Weiterhin wird durch die Idee des *Constructive Alignment* ein starker Bezug zu konstruktivistischen Vorstellungen von Lehren und Lernen hergestellt, die TYLER (1949, S. 63) folgendermaßen zusammenfasst: „Lernen findet durch das aktive Verhalten des Lernenden statt; es ist das, was er tut, was er lernt, nicht das, was der Lehrer

tut.“ Die Lehrperson agiert als *guide by the side*, gibt Feedback und unterstützt die Studierenden (SONNLEITNER, 2016). Konstruktivistische Lernumgebungen sind durch komplexe und lebensnahe Problemstellungen gekennzeichnet, die zusätzlich mit unterschiedlichen Medien vermittelt werden. Die Initiierung von Lernprozessen erfordert einen aktiven, selbstgesteuerten, konstruktiven, situativen und sozialen Prozess (SHUELL, 1986; ERTMER & NEWBY, 2013, S. 55; GRÄSEL & GNIEWOSZ, 2015, S. 22). Der Begriff ‚Alignment‘ bedeutet, dass sowohl der Unterricht als auch die Bewertung auf die beabsichtigten Lernergebnisse ausgerichtet sein müssen (BIGGS, 2014, S. 9).

Die kursorisch beschriebenen Herausforderungen für und Anforderungen an eine kompetenzorientierte Prüfungsgestaltung greifen die Beiträge der vorliegenden Ausgabe der Zeitschrift für Hochschulentwicklung auf, über die wir nachfolgend einen Überblick geben.

In diesem Heft bildet der Beitrag von *Sabrina Gallner* den Ausgangspunkt für die Verortung von gesellschaftlichen Erwartungen an Prüfungen im Hochschulkontext in der heutigen Zeit und die Analyse der Funktionen einer Prüfung. Während Letztere sich im historischen Verlauf wenig verändert haben, lässt sich feststellen, dass sich die Erwartungen an Prüfungen in Richtung einer kompetenzorientierten Perspektive entwickelt haben. Unter diesem Aspekt werden die Bedeutung und die Prüfung der akademischen Kompetenz näher betrachtet.

Daran anschließend klärt *Heidi Niederkofler* in ihrem Artikel den Umgang mit Prüfungserfolgen bzw. primär mit Misserfolgen. Beides gehört zum studentischen Erfahrungsschatz, doch wurde gerade den Prüfungsmisserfolgen bisher wenig Beachtung geschenkt. Didaktisch endet der Prüfungsprozess mit der Notenbekanntgabe. Misserfolge werden fast gar nicht mit den Studierenden aufgearbeitet. In letzter Konsequenz kann er bis hin zum Studienabbruch führen. Dabei könnte mittels Selbstreflexion und Hilfestellung durch die Universität das Misserfolgserlebnis umgewandelt und gewinnbringend in den Lernprozess eingebaut werden.

*Uta Mohring, Christoph Jacobi und Kai Furmans* stellen ein innovatives Lehrkonzept vor, das am *Inverted Classroom*-Konzept angelehnt ist und das Ziel verfolgt, die Kompetenzen von Studierenden der Ingenieurwissenschaften zur selbstständigen Bearbeitung offener und zukünftig relevanter Fragen zu fördern. Die Ausgestaltung

der Lehrmethoden und der Prüfungen orientiert sich an den Zielen des Constructive Alignment. Im Beitrag wird das Potenzial des entwickelten Lehrkonzepts aufgezeigt. Eckpunkte dieses Potenzials sind die selbstständige Erarbeitung von Inhalten, die anschließende gemeinsame Vertiefung und die Bearbeitung von Fallbeispielen. Auch durch die Anerkennung von Kolloquien als Prüfungsleistungen wird ein praxisnahes Lernen fokussiert.

*Miriam Braßler* bezieht in ihrem Beitrag ebenfalls die Perspektive des Constructive Alignment im Zusammenhang mit interdisziplinärem Arbeiten ein. Die Sicherstellung und Prüfung jener Kompetenzen erfordert zunehmend interdisziplinäres Teamteaching. Dabei wird vor allem auf interdisziplinäre Lernziele, auf daraus resultierende Lehr-Lern-Methoden und auf die dafür passenden Prüfungsmethoden eingegangen. Die Autorin gibt Gestaltungsempfehlungen in Form von Leitfragen und Leitlinien für die gemeinsame Entwicklung von Lernzielen und verdeutlicht den Einsatz anhand von zwei Praxisbeispielen.

Durch die Covid19-Pandemie wurde der Trend verstärkt, Lehrveranstaltungen in die Online-Welt zu übertragen. Dadurch wurde die Entwicklung neuer Formen von Wissens- und Kompetenzüberprüfungen gefördert. Der Beitrag von *Silke Braselmann, Jolene Mathieson und Oliver Moisch* über multimodale take-home exams in der Online-Lehre zeigt die Vorteile und Herausforderungen solcher Prüfungen in der Lehrausbildung auf, indem ein Prüfungsdesign eines take-home exams beschrieben und evaluiert wird.

Zwei Lehrveranstaltungen und deren Prüfungsdesign, nämlich einerseits der Bachelor-LV „Angewandte Statistik“ und andererseits der Master-LV „Statistische Versuchsplanung“ beschreiben *Bernhard Spangl, Dóra Kértész und Christian F. Freisleben-Teutscher*. In diesen Lehrveranstaltungen wurde pandemiebedingt das online two-stage exam als kollaboratives Prüfungsformat aufgenommen. Ein Mehrwert für die Studierenden ergibt sich durch einen intensiven Austausch in der Gruppe und die Möglichkeit, während der Prüfung gemachte Fehler noch in der Prüfungssituation korrigieren zu können.

*Christian Spoden, Aron Fink, Andreas Frey, Hanna Köhler und Patrick Naumann* stellen ein Konzept zu kompetenzorientierten adaptiven E-Klausuren vor, die beispielsweise die Vergleichbarkeit von Studierendenkohorten und die Fairness einzelner Prüfungen erhöhen. Solche Klausuren stellen daher eine neue zeitökonomische

Prüfungsform mit individueller Feedbackmöglichkeit für Studierende dar. Anhand der Beschreibung der Konzeption kompetenzorientierter adaptiver E-Klausuren und auf der Basis eines Anwendungsbeispiels beschreiben Spoden et al. Potenziale und Herausforderungen dieser Prüfungsform.

In ihrem Beitrag konzentrieren sich *Miriam Clincy, Karin Melzer, Gunter Schaaf, Achim Eichhorn und Nathalie Verné* ebenfalls auf online-basierte Tests und beziehen sich auf die Mathematik und Physik. Jene Tests spornen die Studierenden zum Mitlernen an und eröffnen semesterbegleitende Feedbackmöglichkeiten. Die „Mid-terms“ bilden das Ende von Lehreinheiten, sind unabhängig von den jeweiligen Lehrkonzepten und durch unterschiedliche Umsetzungsvarianten je nach Rahmenbedingungen angepasst einsetzbar.

*Sandra Niedermeier, Michael Sailer, Lukas Remböck und Matthias Stadler* beschäftigen sich in ihrem Artikel mit dem Prüfungsformat Open-Book, welches das zuvor dominante Close-Book-Prüfungsformat in den Hintergrund drängt. Das Open-Book-Prüfungsformat zeichnet sich durch einen problemlösungsorientierten und realitätsnahen Charakter aus. Studierende erleben weniger Angst im Vergleich zu Close-Book-Prüfungen. Im Raum steht nun die Frage, ob und in welchem Umfang mitgebrachte Unterlagen die Performanz in einer Open-Book-Klausur beeinflussen. Open-Book-Prüfungsformate werden laut Niedermeier et al. zukünftig zunehmen und erfordern vonseiten der Lehrenden höhere Ansprüche bei der Frageformulierung.

Der Einsatz von Fallstudien eignet sich als kompetenzüberprüfende Methode, wird aber sehr selten in großen Kursen eingesetzt. *Eva Osteresch* zeigt in ihrem Beitrag auf, wie diese Vorgehensweise entworfen und umgesetzt werden kann. Das entwickelte Konzept führt zu einer hohen Zufriedenheit, fördert den Lernerfolg bzw. die Motivation und unterstreicht das Potenzial von Fallstudien als kompetenzorientierte Prüfung sowie den sinnvollen Einsatz in großen Kursen.

Im Werkstattbericht von *David Bisang, Kerstin Bäuerlein und Corinne Wyss* geht es um eine alternative schulpraktische Abschlussprüfung in Form eines Video-Portfolios für Lehramtsstudierende zur Überprüfung der Erreichung von Kompetenzzielen. Das Video-Portfolio überprüft dabei die Planungs-, Unterrichts- und Analysekompetenz von angehenden Lehrpersonen. Ein solches Prüfungsformat bringt

zwar Herausforderungen mit sich, bietet jedoch sowohl für Studierende als auch für eine realitätsnahe Beurteilung Vorteile.

Durch Simulationen in Form von Rollenspielen sollen Lehramtsstudierende in der Entwicklung ihrer kommunikativen Kompetenz gefördert werden. In ihrem Beitrag untersuchen *Julia Fecke und Lars Müller* Vor- und Nachteile von Simulationen im virtuellen Raum in avatar- und videokamerabasierten Formaten. Den Ergebnissen zufolge sind vor allem avatarbasierte Simulationen vielversprechend. Die Testpersonen beginnen im Vergleich zu videobasierten Formaten Gespräche schneller und ungehemmter, während gleichzeitig Sorgen um die richtige Deutung von Mimik und Gestik des Gegenübers wegfallen.

*Olaf Zawacki-Richter, Wolfgang Müskens, Nadine Dembski und Sonja Lübben* widmen sich in ihrem freien Beitrag Open Educational Resources (OER) in der wissenschaftlichen Weiterbildung. Denn eine Herausforderung für Weiterbildungsanbieter besteht in der Auswahl qualitativ hochwertiger Materialien. Zudem stellen sich Fragen zu den didaktischen und technischen Qualitätsdimensionen für digitale Lernmaterialien sowie zu der Qualitätssicherung im Hinblick auf die praktische Implementierung. Entlang der Kriterien opportunity, innovation, benefits, challenges und potentials analysieren die Autoren und Autorinnen die Chancen und Herausforderungen von OER in der wissenschaftlichen Weiterbildung

Neugierig geworden? Folgen Sie den facettenreichen Beiträgen, die zahlreiche Konzepte von Prüfungen im Kontext kompetenzorientierter Hochschulbildung beinhalten. Wir sind jedenfalls davon überzeugt, dass die Beiträge als wohlgesetzte thematische Spots Lehr- und Prüfungsformate theoretisch und praktisch fassbar machen.

## Literatur

**Biggs, J., & Tang, C.** (2011). *Teaching for quality learning at university*: McGraw-hill education (UK).

**Biggs, J.** (2014). Constructive alignment in university teaching. *Hersda Review of Higher Education*, 1, 5–22.

**Ertmer, P. A., & Newby, T. J.** (2013). Behaviorism, Cognitivism, Constructivism: Comparing critical features from an instructional design perspective. *Performance improvement quarterly*, 26(2), 43–71.

**Gräsel, C., & Gniewosz, B.** (2015). Überblick Lehr-Lernforschung. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Gegenstandsbereiche* (2. Aufl.) (S. 19–24). Wiesbaden: Springer VS.

**Schindler, C.** (2015). *Herausforderung Prüfen*. (Dissertation). Technische Universität. <http://d-nb.info/1079001778/34>

**Shuell, T. J.** (1986). Cognitive conceptions of learning. *Review of Educational Research*, 56, 411–436. DOI: <https://doi.org/10.2307/1170340>

**Sonnleitner, K.** (2016). From Sage on the Stage to Guide on the Side. Einsatz kooperativer und selbstgesteuerter Lernszenarien zur Förderung von fachlichen und sozialen Kompetenzen. *ZDRW*, 4, 288–302.

**Tyler, R.W.** (1949). *Basic principles of curriculum and instruction*. Chicago: University of Chicago Press.

**Wissenschaftsrat, W.** (2012). Empfehlungen zur Akkreditierung als Instrument der Qualitätssicherung. *Bremen*, 25.

## Autor\*in



DDr.in Karin SONNLEITNER, M.A. | Karl-Franzens-Universität Graz,  
Zentrum für Soziale Kompetenz | Merangasse 12, A-8010 Graz

<https://soziale-kompetenz.uni-graz.at/de/>

[karin.sonnleitner@uni-graz.at](mailto:karin.sonnleitner@uni-graz.at)



PD Dr. Martin GARTMEIER | TUM Medical Education Center, Lehrstuhl  
für Medizindidaktik, medizinische Lehrentwicklung und Bildungsforschung  
| Ismaninger Straße 22, D-81675 München

<http://www.mec.med.tum.de/de/pd-dr-phil-habil-martin-gartmeier>

[martin.gartmeier@tum.de](mailto:martin.gartmeier@tum.de)

Sabrina GALLNER<sup>1</sup> (Luzern)

# Was Prüfungen leisten sollen: Prüfungen für akademische Kompetenz

## Zusammenfassung

Die Prüfungskulturen an Hochschulen werden mitunter als problematisch und wenig reflektiert beschrieben. Dabei steht heute die Forderung nach kompetenzorientierten Prüfungen im Mittelpunkt der Diskussionen. Im Rahmen der europäischen Studienreform rückt im Zuge des kompetenzorientierten Prüfens die didaktische Funktion von Prüfungen in den Vordergrund. Die Selektionsfunktion scheint weit- aus stärker tradiert, aber im Widerspruch zur Kompetenzorientierung gesehen zu werden. Prüfungen gelten als traditionelles Element hochschulischer Lehre, die sich an gesellschaftlichen Erwartungen ausrichten. Dieser Beitrag skizziert anhand eines bildungswissenschaftlich historischen Rückblicks auf das Prüfungswesen entlang der Prüfungsfunktionen, wie sich die gesellschaftlichen Erwartungen auf das heutige Prüfungswesen auswirken und welche Bedeutung Prüfungen heute haben. Es kann gezeigt werden, dass die Prüfungsfunktion(en) sich im Grunde wenig geändert haben, sehr wohl jedoch die Erwartung an die Prüfungsleistung. Dies führt zu Missverständnissen über Prüfungen und deren Funktion, die wiederum zu ‚problematischen Prüfungskulturen‘ beitragen können.

## Schlüsselwörter

Prüfungsfunktionen, Prüfung, Kompetenzorientierung, Akademische Kompetenz

---

1 E-Mail: [sabrina.gallner@phlu.ch](mailto:sabrina.gallner@phlu.ch)



## **What exams should accomplish: Exams for academic competencies**

### **Abstract**

Examination cultures at universities are sometimes described as problematic not very well thought out. Today, there is much discussion about competence-oriented examinations. The various documents that accompanied the Bologna Process focus on the educational function of examinations for competence-oriented testing. The function of selection seems to be traditionally more in the foreground and seen as a contradiction to competence-based teaching. Examinations are seen as a traditional element of higher education teaching, which is oriented around societal expectations. This paper explains how societal expectations affect today's examination system by means of an historical review of the examination system and the functions of examinations. It can be shown that the examination function(s) have not changed significantly, but expectations for examination performance have changed. This leads to misunderstandings about exams and their function, which in turn can contribute to ‚problematic exam cultures‘.

### **Keywords**

examination, competence-oriented exam, academic competence

## **1 Einleitung**

Im Call zum Themenheft „Prüfen im Kontext kompetenzorientierter Hochschulbildung“ Jg. 17/Nr. 1 der ZFHE heißt es: „Gleichzeitig haben sich in unterschiedlichen Institutionen der formalen Bildung sowie in verschiedenen Fächern spezifische, manchmal wenig reflektierte oder gar problematische Prüfungskulturen herausgebildet, die das Lehren und Lernen innerhalb der Organisationen stark prägen (DÖBLER, 2019).“

DÖBLER (2019) weist darauf hin, dass weniger die Strukturvorgaben an sich problematisch seien als vielmehr deren Implementation „unter lernkulturellen Rahmenbedingungen, die noch im Modus klassischer Lehr- und Prüfungsverständnisse“ der

---

jeweiligen Fachkultur operieren würden. Wird von klassischen Lehr- und Prüfungsverständnissen gesprochen, so liegt es auf der Hand, dass es ein ‚zeitgemäßeres‘, ‚aktuelleres‘ Verständnis geben muss. Im Zuge der europäischen Studienreform werden kompetenzorientierte Lehr- und Prüfungsformen gefordert. Diese Forderung ist bildungspolitisch und somit auch von wirtschaftlichen Interessen geprägt. In Zusammenhang mit der Bologna-Reform und kompetenzorientierten Prüfungsformen rückt die didaktische Funktion von Prüfungen in den Fokus.

Eine Fokussierung rein auf die didaktische Funktion kann jedoch zu einem Verschwimmen des Prüfungscharakters führen, sodass die Prüfungsfunktionen der Linearisierung und Systemreproduktion zu wenig berücksichtigt werden (vgl. REIS & RUSCHIN, 2008).

Um die Auseinandersetzung mit Prüfungsfunktionen fruchtbar zu machen, werden die beiden Begriffe ‚Prüfung‘ und ‚Funktion‘ zunächst kurz eingeordnet.

Unter Prüfungen kann die „systematisch ausgeführte Ermittlung des Standes der Kenntnisse und Fertigkeiten eines Prüflings durch Sachverständige verstanden“ (PAULSEN, 1902) werden. Prüfungen sind als Instrument der Leistungsdifferenzierung (REIS & RUSCHIN, 2008) bzw. Leistungsmessung (EUGSTER & LUTZ, 2004) zu verstehen. Prüfung bedeutet im Wortursprung auch ‚Erprobung an der Realität‘ (vgl. PRAHL, 1976). Ohne nun näher auf den Begriff der ‚Realität‘ einzugehen, wird hier doch auch die Verbindung zu heutigen Kompetenzdefinitionen deutlich. Kompetenz zeigt sich in der ‚Performance‘, im Handeln in herausfordernden neuartigen Situationen (vgl. HOFER et al., 2016; GNAHS, 2010; CHOMSKY 1973; WEINERT, 2001; KLIEME, 2004). Kompetenzen können in ‚künstlichen‘ Prüfungssituationen in Bildungseinrichtungen daher immer nur exemplarisch dargestellt werden.

Unter Funktion wird eine „[klar umrissene] Tätigkeit, Aufgabe innerhalb eines größeren Zusammenhanges; Rolle“ (DUDEN.DE, o. J.) oder von einem Programm o. Ä. zu leistende Aufgabe bzw. ein zu lieferndes Resultat verstanden. Prüfungsfunktionen definieren, warum es Prüfungen überhaupt geben soll.

PRAHL (1995) weist darauf hin, dass (entsprechend) zu unterscheiden sei zwischen gesellschaftlichen, wirtschaftlichen und politischen Anforderungen von Prüfungen und den pädagogischen, fachlichen und hochschulorganisatorischen Zwecksetzungen, die in Wechselwirkung miteinander stehen. Die Auseinandersetzung mit gängi-

gen Systematisierungen zu Prüfungsfunktionen zeigt, dass ein Konsens hin zu einer Dreiteilung besteht. Zunächst lieferte PAULSEN (1902) eine erste Unterscheidung. Er führt in seinem Werk „Die deutschen Universitäten“ verschiedene ‚Arten‘ von Prüfungen auf. Er unterscheidet den Nutzen von Prüfungen für den Unterricht und dem Unterricht fremden Zweck:

- a) Prüfungen „wachsen aus den Zwecken des Unterrichts hervor und wollen ihm dienen, sie zeigen dem Lehrer den Stand des Schülers und geben ihm damit den Maßstab zur Beurteilung des Fortschreitens [...] und Aufsteigens auf eine höhere Stufe.“ (PAULSEN, 1902, S. 426).
- b) Sie geben „zugleich dem Schüler einen Antrieb, durch die Vorbereitung seine Kenntnisse zu befestigen, und Gelegenheit, seines Besitzes inne zu werden.“ (PAULSEN, 1902, S. 426)
- c) Eine weitere wesentliche ‚Art von Prüfung‘ als nach außen gerichtete Prüfungsfunktion sind bei ihm die „Staats- oder Amtsprüfungen [...]: sie dienen als Mittel der Auslese unter den Bewerbern um Aemter oder bestimmte Berufe. Darum werden sie nicht von den Lehrern (wenigstens nicht von ihnen als solchen), sondern von eigens dazu bestellten Examinatoren nach einer amtlich festgestellten Norm (Prüfungsordnung) abgehalten.“ (PAULSEN, 1902, S. 426f).

Diese Dreiteilung zeigt die Bedeutung der Prüfung a) für die Selektion und Rekrutierung (intern), b) die didaktische Wirkung sowie c) die Legitimation nach außen. Ähnlich ist diese Systematisierung auch bei späteren Autor\*innen zu finden (KVALE, 1972, sowie SCHEER & ZENS, 1973; FLECHSIG, 1976; PRAHL, 1995, und REIS & RUSCHIN, 2008), wobei es jeweils leichte Variationen im Sprach- und Bedeutungsgebrauch gibt.

Im Folgenden wird näher auf die historische Entwicklung der Prüfungsfunktionen bis in die Gegenwart eingegangen. Die Ausführungen beziehen sich mehrheitlich auf den deutschsprachigen Hochschulraum.

## 2 Prüfungsfunktionen in ihren historischen und aktuellen gesellschaftlichen Bedeutungen

### 2.1 Prüfung als Ausweis für ein Qualitätsminimum

PAULSEN (1902) beschreibt das ‚Aufsteigen auf eine höhere Stufe‘ noch als einen rein internen Zweck. Spätestens bei FLECHSIG (1976) werden der Rekrutierungsfunktion weitere Aspekte zugeordnet und die rein interne Perspektive löst sich auf. Dies ist auch historisch begründbar. Bereits die ersten Universitäten um 1200 verfügten über eine rechtlich autonome Selbstverwaltung und konnten selbstständig akademische Grade verleihen (vgl. SCHELSKY, 1963, sowie FISCH, 2015, und PRAHL, 1976). Die Prüfung verfolgte hier den Nutzen, die Qualität des Lehrpersonals für die interne Rekrutierung über eigene nationale und territoriale Grenzen hinweg sicherzustellen (vgl. FISCH, 2015).

PRAHL (1995) leitet daraus zwei miteinander einhergehende Funktionen von Prüfungen ab: Die *Initiation* bezieht sich auf das mit der Prüfung und dem Grad gegebene Versprechen, dass eine Person für die Lehre kompetent sei. Damit ist auch immer eine *Selektion* verbunden, die sich zunächst jedoch auf die Auswahl der Studierenden für das Weiterkommen innerhalb der Universität bezog.

Bei FLECHSIG (1976) werden der Qualifikationsnachweis (Herstellen eines Vertrauensminimums, dass die angegebenen Kenntnisse, Fertigkeiten, Haltungen tatsächlich vorhanden sind), die Platzierung (Rangfolge innerhalb der Teilnehmenden des gleichen Prüfungssystems) und die Auslese (Wirksamkeit der Rangfolge für das Weiterkommen) unter der Rekrutierungsfunktion subsumiert. Dies aber nun nicht mehr nur, um die interne Rekrutierung sicherzustellen:

„Das mittelalterliche Versprechen für die Lehre kompetent zu sein, hat sich bereits im späten Mittelalter auf andere Tätigkeitsbereiche übertragen. Der mit der Prüfung erbrachte Kompetenznachweis wurde zum Selektionskriterium für die Besetzung von Berufspositionen (so in Verwaltung, Justiz, später auch im Bildungswesen und in der Medizin).“ (PRAHL, 1995, S. 444)

Bei REIS & RUSCHIN (2008) sind diese drei Aspekte umbenannt (Kompetenznachweis für extern, Platzierung in Kohorte und Auslese aus Kohorte) und die Funktion als ‚Linearisierung (Selektions- und Allokationsfunktion)‘ bezeichnet. Deutlich bleibt aber erhalten, dass es um das Herstellen eines Minimalkonsenses über die Befähigung von Personen geht, die über die Institution hinaus gültig bleibt.

Diese selektierende Funktion von Prüfungen hat auch heute noch eine hohe Bedeutung, regeln Prüfungen die Zulassung zu Hochschulprogrammen, das Weiterkommen innerhalb des Studiums und dienen der Klassifizierung der Absolvent\*innen in ihren Leistungen.

## 2.2 Status durch Prüfung

PAULSEN (1902) beschreibt eine weitere Funktion von Prüfung als nach außen gerichtet und verweist auf Staats- oder Amtsprüfungen. Diese Funktion hat sich erst mit der Entwicklung spezifischer Berufsfelder entwickelt, denn die mittelalterlichen Universitäten waren zunächst nicht darauf ausgerichtet, für einen Beruf bzw. für praktisches Handeln auszubilden (vgl. KINTZINGER, 2007). Tatsächlich erlangte nur rund ein Viertel der Studierenden einen akademischen Grad (vgl. EULENBERG, 1904, nach PRAHL, 1976). Personen, die die mittelalterliche Universität besuchten, konnten auch ohne Abschluss Anerkennung außerhalb der Universität finden. Sie waren im landesherrlichen Dienst eingesetzt, konnten als Ärzte praktizieren, Ämter in Domstiften oder Vikariaten ausüben sowie als Notare, Stadtschreiber oder Schulmeister arbeiten. Dies auch deswegen, weil das Studium generell (auch in der Medizin oder den Rechtswissenschaften) zunächst nicht mit einer praktischen Tätigkeit in Bezug stand. Das eigene Beziehungsnetz, entsprechend dem Stand beeinflusste stark die Position, die jemand erhielt. Jedoch konnten Studenten aus den unteren Ständen prinzipiell durch Studium Zugang zu entsprechenden höherwertigen Tätigkeiten erhalten (vgl. KINTZINGER 2007; RÜEGG 1993), so dass Prüfungen auch „sozialen Aufstieg durch Leistung“ ermöglichten (vgl. PRAHL, 1995).

Im 16. Jh. wurde die Pflege der Wissenschaften in Deutschland vermehrt zur staatlichen Aufgabe erklärt und die Universitäten wurden der staatlichen Kontrolle unterstellt, die Professoren beim Staat angestellt. Durch die Vereinheitlichung der Abschlüsse durch institutionalisierte Examen wurden zunehmend die universitären Abschlüsse bedeutsam für den Zugang zu höheren Berufspositionen. Stand und

Netzwerke verloren somit an Bedeutung (vgl. SCHELKY, 1963). Für die Arbeitgebenden kann so scheinbar eine Kontinuität in der Stellenbesetzung sichergestellt werden, da durch die Zertifikate den Absolventen bestimmte Merkmale zugeordnet werden. Der Zugang zu Amtsstellen unterlag nun vollständig der staatlichen Kontrolle. Daraus erwuchs ein umfassendes Berechtigungswesen mit einer hohen Bürokratisierung, denn die „zertifizierten Absolventen der Universität beanspruchen zugleich eine Alleinzuständigkeit für die entsprechenden Berufsfelder und ein Alleinzugangsrecht auf die entsprechenden Stellen.“ (MATTHIES & STOCK, 2020, S. 227). Anfang des 19. Jahrhunderts wurde die berufsvorbereitende Funktion des Studiums immer stärker (vgl. PRAHL, 1976). Es entstand ein akademischer Arbeitsmarkt, die Berufe wurden spezialisierter, was sich auch in Prüfungen bemerkbar machte. Weniger der Stand, Gilde oder Zunft war nun für die Identifikation maßgeblich, sondern mehr und mehr der Beruf (vgl. PAULSEN, 1902).

Die universitären Grade versicherten somit auch eine Legitimation außerhalb der Hochschule. Bei KVALE (1972) wird dies als ‚Machtfunktion‘ ausgewiesen und betont den sanktionierenden Charakter von Prüfungen sowie die sozialisierende Funktion. So würde ein Stück weit auch immer die Anpassungsfähigkeit der Studierenden an die Fachkultur geprüft. Bei SCHEER und ZENZ (1973) ist diese Funktion als ‚Statusverleihung‘ benannt, denn die Prüfung vergibt den Status als Akademiker\*in und eröffnet somit exklusiven Zugang zu bestimmten beruflichen Positionen. FLECHSIG (1976) orientiert sich an KVALE (1972) mit der ‚Herrschafts- und Sozialisationsfunktion‘, die bei REIS und RUSCHIN (2008) wiederum in ‚Systemreproduktion‘ benannt ist, jedoch die gleichen Unterkategorien enthält: Initiation, Statusverleihung und Legitimation. Bei PRAHL (1995) ist die Initiation eine der beiden Hauptfunktionen: Personen werden in bestimmte Wissensbestände eingeweiht und gelten mit der Prüfung als berechtigt. Mit dieser Berechtigung, die anderen Personen vorenthalten ist, geht Sozialprestige einher und Akademiker\*innen rücken an den Adelsstatus heran. Nicht mehr das Geburtsrecht ist für den gesellschaftlichen Stand ausschlaggebend, sondern die Leistung bringt das Ansehen, das sich in Graden niederschlägt.

Auch in der Gegenwart ist eine Hauptfunktion von hochschulischen Prüfungen, die Legitimierung von Personen für bestimmte Positionen innerhalb und außerhalb der Hochschule.

## 2.3 Kompetenzen zertifizieren

Der Trend in Europa ging in Richtung Bürokratisierung und Preußen war durch die Niederlage gegen Frankreich sehr vom napoleonischen Vorbild beeinflusst. Es gab napoleonisch ausgerichtete Universitäten, die ihre Ausbildung auf den Arbeitsmarkt fokussierten und die humboldtsch ausgerichteten Universitäten, die unabhängig vom Beruf ausbilden und die Neugier an der Wissenschaft wecken wollten (vgl. FRAENKL-HAEBERLE, 2014).

MATTHIES & STOCK (2020) weisen darauf hin, dass die grundlegenden Umstellungen zu einer modernen Forschungsuniversität einen enormen Einfluss auf die Stellung und Aufgaben der Professoren hatten. In der mittelalterlichen Universität stand die Tradierung von Erkenntnissen im Zentrum. Nun rückte der Erwerb neuer Kenntnisse in den Vordergrund und Forschungsschwerpunkte wurden nach eigenen Kriterien ausgewählt. Die Publikationserwartungen, die damit einhergehen, veränderten das Aufgabenfeld, da Reputation nun verstärkt mit Forschung und Publikation in Verbindung steht. Die Disziplin versteht sich als Forschergemeinschaft „die akademische Lehre ruht jetzt auf der Autonomie der Wissenschaft und deren Entwicklungslogik.“ (MATTHIES & STOCK, 2020, S. 232)

Eine forschungsnahe Lehre bringt andere Prüfungsformen, wie die studentische Hausarbeit, hervor. Grenzen dieser Verknüpfung von Forschung und Lehre wurden im ersten Drittel des 19. Jahrhunderts deutlich, da die Expansion der Studierendenzahlen (durch das Berechtigungswesen über das Abitur) einen Anstieg der Lehrbelastung zur Folge hatte, was die Zeit für Forschungsaktivitäten eingrenzte (vgl. MATTHIES & STOCK, 2020). Mit dem Anstieg der Studierendenzahlen nahm auch die Bedeutung der Platzierung innerhalb der Kohorte (Linearisierungsfunktion) an Bedeutung zu, da nun viele um eine begrenzte Anzahl an Posten wetteiferten. Die Prüfung wurde stärker für die Selektion eingesetzt, um den Zugang zum Studium einzuschränken (Allokation) und die besten Studierenden zu rekrutieren.

Durch den Aufschwung der experimentellen Wissenschaften rückten wiederum Prüfungsformen in den Fokus, die eine stärker standardisierte Messung erlaubten. So sollten subjektive Einflüsse in den Prüfungen vermieden und auch Kapazitätsproblemen entgegengewirkt werden (vgl. PRAHL, 1995). Durch die verstärkte Selektion kam es zu Veränderungen der Prüfungsformen respektive Aufgabenstellungen hin zu Formen und Aufgaben, die gut messbar sind (vgl. HAMMERSTEIN,

---

2007). Ein Grenzstreit zwischen Prüfungsformen und Fächern entsteht, der auch heute noch relevant erscheint.

In den 1970er-Jahren kam es zu einer weiteren Expansion des Hochschulwesens, in dieser Zeit wurde vor allem die Prüfungsdichte bemängelt (vgl. SCHÜTZ et al., 1969). Die seit den 1990er-Jahren schließlich umgesetzte europäische Studienreform verfolgt das Ziel, die Europäische Union zum „wettbewerbsfähigsten und dynamischsten wissensbasierten Wirtschaftsraum der Welt“ zu machen (vgl. KONFERENZ DER EUROPÄISCHEN HOCHSCHULMINISTERINNEN UND -MINISTER, 2003). Dies führte zu weiteren starken Veränderungen des Prüfungswesens. So soll es überschaubare Studien- und Prüfungsordnungen geben, kontinuierliche Leistungskontrollen sollen zur Orientierung der Studierenden dienen und zu erbringende Leistungsnachweise sollen für die Studierenden transparent sein (HRK, 1997).

Bereits PAULSEN (1902) erwähnt die Prüfung als ‚Antreiber‘ für das Lernen. Diese Funktion ist bei KVALE (1972) als studiensteuernde Funktion zu finden. Auch bei SCHEER und ZENZ (1973) ist die ‚Steuerung‘ eine pädagogische Funktion, die das Verhalten der Studierenden lenkt, normiert und sanktioniert. FLECHSIG (1976) bezeichnet diese Funktion als ‚Didaktische Funktion‘. Die Prüfung steuert durch zeitliche und inhaltliche Gliederung den Studiengang, orientiert die Studierenden über die Studienziele, gibt eine Rückmeldung über die Lehr- und Lernerfolg an Lernende und Lehrende, motiviert für das Lernen und diagnostiziert zukünftige Lernprozesse. REIS und RUSCHIN (2008) übernehmen diese Definition etwas abgewandelt als ‚Systemoperation‘ mit den Aspekten der Prüfung als strukturierendes Element, Rückmeldung des Lehr-Lern-Erfolgs, extrinsisches Motivationsinstrument und Diagnoseinstrument.

In der Gegenwart wird nach wie vor bemängelt, dass zu wenig kompetenzorientiert geprüft würde. So wird im ‚Fachgutachten zur Kompetenzorientierung in Studium und Lehre‘ (vgl. SCHAPER, 2012) festgestellt, dass eine didaktisch kohärente Planung von Lehr-Lern-Settings im Sinne des Constructive Alignment und Lernergebnisse in den Prüfungen daher kaum eine Rolle spielten.<sup>2</sup>

---

2 Eine aktuellere Studie liegt m. E. zufolge nicht vor.

## 3. Was Prüfungen leisten sollen

### 3.1 Zusammenfassung

In Abbildung 1 sind die Prüfungsfunktionen zusammengefasst nach Autor\*innen dargestellt.

Autor*innen	pädagogische, fachliche und hochschulorganisatorische Zwecksetzungen	gesellschaftliche, wirtschaftliche und politische Anforderungen	
Paulsen 1902	<b>Interner Schulvorgang: für den Schüler</b> - geben dem Schüler einen Antrieb - Befestigung der Kenntnisse	<b>Interner Schulvorgang: Nutzen für den Unterricht</b> - Stand des Schülers zeigen - der Möglichkeiten des Aufsteigens auf eine höhere Stufe	<b>Staats- oder Amtsprüfungen</b> Mittel der Auslese unter den Bewerbern um Aemter oder bestimmte Berufe
Kvale 1972	<b>Studiensteuernde Funktion</b>	<b>Rekrutierende Funktion</b>	<b>Machtfunktion</b>
Scheer und Zenz 1973	<b>Steuerung</b> ... ist eine pädagogische Funktion und steuert das Verhalten der Studierenden, normiert aber ebenfalls stark durch Sanktionen.	<b>Rekrutierende Funktion</b> Die Rekrutierung dient der Beschaffung von Nachwuchs und ist somit ein diagnostischer Vorgang.	<b>Statusverleihung</b> ... stellt den Zugang zur Ausübung einer Tätigkeit sicher und vergibt den Status als Akademikerin bzw. Akademiker.
Flechsig 1976	<b>Didaktische Funktion</b> - Zeitliche und inhaltliche Gliederung des Studiengangs - Orientierung der Lehrenden und Lernenden über die Studienziele - Rückmeldung des Lehr-Lernerfolgs an Lehrende und Lernende - Motivation - Diagnose Lernvoraussetzungen	<b>Rekrutierungsfunktion</b> - Qualifikationsnachweis - Platzierung - Auslese	<b>Herrschafts- und Sozialisationsfunktion</b> - Initiation - Statusverleihung - Legitimation
Prahl 1995	<b>Didaktische Funktionen</b> - Abgrenzung zwischen Hochschultypen, Fächern und Studiengängen - Studiensteuernde Effekte für das Subjekt - Evaluation und Rückmeldefunktion	<b>Selektionsfunktion</b> - Selbstrekrutierung des hochschul. Nachwuchses	<b>Initiationsfunktion</b> - Statuszuweisung
Reis & Ruschin 2008a	<b>Systemoperation (Didaktische Funktion)</b> - Strukturierendes Element - Rückmeldung des Lehr-Lernerfolgs - Extrinsisches Motivationsinstrument - Diagnoseinstrument	<b>Linearisierung (Selektions- und Allokationsfunktion)</b> - Kompetenznachweis für extern - Platzierung in Kohorte - Auslese aus Kohorte	<b>Systemreproduktion (Herrschafts- und Sozialisierungsfunktion)</b> - Initiation - Statusverleihung - Legitimation

Abb. 1: Systematisierung von Prüfungsfunktionen verschiedener Autor\*innen (eigene Darstellung)

Die Verleihung von (akademischen) Graden durch Prüfungen ist ein Alleinstellungsmerkmal von Universitäten, das es bei deren Vorläufern so nicht gab. PRAHL (1976) stellt dar, dass die Initiationsriten, die der Kirche entlehnt sind, die Legitimation des Prüfungswesens verstärkten, da eine Legitimation durch Tradition nicht möglich war. So wurde auch die Nähe zu den tonangebenden Gruppen hergestellt (Adel und Klerus). Später wurde durch die Prüfungskultur das Leistungsprinzip verstärkt, denn auch untere Schichten konnten durch Leistung den Aufstieg in einen

höheren Stand schaffen. Insbesondere durch die Einführung von Staatsprüfungen, der Entwicklung konkreter Berufe und der Einführung des Abiturs als Zugangsbe-  
rechtigung veränderte das Hochschul- und damit auch das Prüfungswesen und die  
Selektion rückte in den Vordergrund. Seit den Gründungszeiten bis heute ist eine  
entscheidende Aufgabe von Prüfungen die Selektion und die Legitimation.

Es zeigt sich das Spannungsfeld dessen, was Prüfungen leisten sollen zwischen den  
von PRAHL (1995) definierten Dimensionen der pädagogisch, fachlich und hoch-  
schulorganisatorischen und den gesellschaftlich, wirtschaftlich und politischen An-  
forderungen. Dass jedoch Prüfungen als Elemente von Studiengängen diesen struk-  
turieren, eine Orientierung bieten und studiensteuernde Effekte haben, ist weniger  
eine Funktion, im Sinne einer originären Aufgabe, als vielmehr eine Wirkung. Die  
eigentliche Aufgabe und Rolle der Prüfung ist die Messung der Leistung, die erwar-  
tet wird, um die Prüflinge anhand von Schwellenwerten zu unterteilen.

Dies wird deutlicher, wenn die Lehre mithilfe eines Wirkmodells angeschaut wird,  
wie das Constructive Alignment eines darstellt. Die Messung des Outcomes ist die  
Aufgabe der Prüfung. Die dort gemessenen Outcomes sind in Form von Lernergeb-  
nissen (Kompetenzziele) formuliert. Interaktionen in den Lehr-Lern-Settings för-  
dern den Kompetenzerwerb der Studierenden. Prüfungen sind hier kein Teil des  
Lehr-Lern-Settings, sie sind der Zielpunkt.

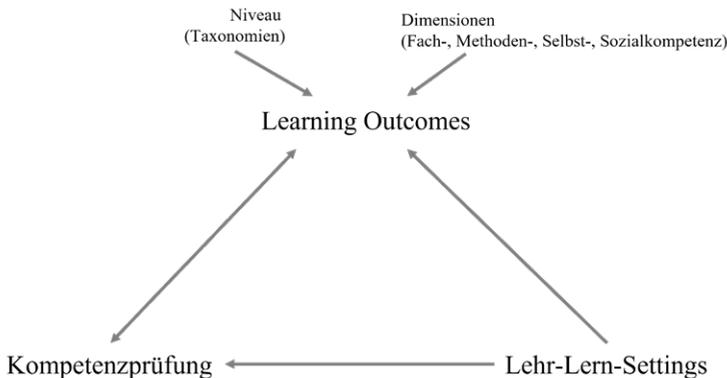


Abb. 2: Didaktische Kohärenz nach dem Modell des constructive alignments angelehnt an BIGGS, 2003 (eigene Darstellung)

Lehr-Lern-Interventionen haben die Absicht, das Lernen entsprechend der formulierten Ziele zu befördern. Wird die Prüfung zu einer Lehr-Lern-Intervention – häufig als formative Prüfung bezeichnet – soll sie das Lernen anregen (durch Feedback, Eigeneinsicht, Motivation, etc.). Werden Prüfungen nicht klar in ihrem Zweck adressiert, geraten sie in eine Zerreihsprobe zwischen Selektionserwartung und pädagogischer Wirkung. Dies hat erst einmal nichts damit zu tun, ob die Prüfung kompetenzorientiert ist – sich also an den formulierten Lernergebnissen ausrichtet.

Die Funktion der Prüfung hat sich, wie in den vorhergegangenen Betrachtungen deutlich wurde, nicht wesentlich verändert, sondern besteht nach wie vor vor allem in der Selektion und Legitimation. Der Wirkkreis hat sich aber erweitert.

Zu Beginn hieß die Anforderung an die Geprüften, fähig zu sein, innerhalb der Universität zu lehren. Die Außenwelt der Universität bemächtigte sich jedoch mit der Entwicklung der Staatsämter und Berufe dieses Prinzips. Nun sollen die Geprüften gewisse grundlegende Berufsanforderungen mitbringen. Mit dem humboldtschen Bildungsideal und der damit postulierten Verbindung von Lehre und Forschung kam nun noch der Anspruch hinzu, dass Forschung nicht nur Grundlage von Lehrinhalten sei, sondern in Prüfungen auch die wissenschaftliche Kompetenz nachgewiesen werden solle. Das Versprechen, welches das hochschulische Prüfungswesen geben soll, wurde also inhaltlich immer wieder erweitert. Dies entwickelte sich weiter bis in die heutige Zeit, in der diese Berufsanforderungen immer spezifischer werden und nun sehr genau definiert ist, was Absolvent\*innen am Ende des Studiums können sollen. Dies ist wiederum ein Versprechen, welches in den Studienplänen und -ordnungen niedergeschrieben wird (Qualifikationsprofil). Das Versprechen, dass eine Person bestimmte Leistungen zu erbringen in der Lage ist (über bestimmte Kompetenzen verfügt), ist schon auf der Ebene des Moduls gegeben (Modulprüfungen) und befördert theoretisch die Mobilität zwischen Hochschulen. Bereits nach einer bestandenen Modulprüfung besteht also eine Legitimation nicht erst mit dem Erwerb des akademischen Grads.

Mit der Kompetenzorientierung geht der Fokus deutlich in Richtung überfachliche Kompetenzen, sodass neben Fach- und Methodenwissen also auch soziale und personale Dimensionen eine Rolle spielen sollen. Zudem rücken die höheren Taxonomiestufen ins Zentrum mit Schlagworten wie Praxistransfer und Anwendungsorientierung. Über das Theorie-Praxis-Problem schrieben MATTHIES und STOCK (2020) ausführlich. In den vergangenen circa 200 Jahren wurde der Kompetenz-

erwerb in den Bildungsbiografien immer weiter vorverlagert. Durch immer engere Spezifizierungen, was ein Beruf ist, wurden auch die Studienprogramme immer spezialisierter. Es herrscht heute die Erwartung vor, dass Absolvent\*innen kompetent in den Beruf starten. Eine breite Ausbildung, wie sie im Mittelalter üblich war, und vielfältige Möglichkeiten, im ‚Berufsleben‘ auch ‚quer‘ einzusteigen, wurden abgelöst von konkreteren Berufen und einer Vorverlegung des Kompetenzerwerbs von ‚on the job‘ zu ‚in the university‘. Wenn der Kompetenzerwerb nun bereits in der Hochschule stattfinden soll, dann müssen Prüfungen Kompetenzen messen – was mit einigen Herausforderungen verbunden ist, die hier aber aus Platzgründen nicht näher diskutiert werden sollen.

Prüfungen mit Selektions- und Legitimationsfunktion in gut messbaren Formaten (in Folge eher Wissensabfragen) abzubilden und kompetenzorientierte Prüfungen – weil sie schwerer bewertbar sind – nur als formative Formate durchzuführen, kann nicht die Lösung sein. Dies widerspricht dem aktuellen Forschungsstand, denn ‚Assessment drives learning‘ (PROSSER & TRIGWELL, 1999, u. A.).

### **3.2 Prüfung akademischer Kompetenz**

Häufig wird argumentiert, Prüfungen würden lediglich auf eine Berufsbefähigung abzielen, die umfassende Bildung der Studierenden gehe dabei verloren. Das kann durchaus zutreffen, wenn, wie bei DÖBLER (2019) erwähnt, weniger die Strukturvorgaben an sich problematisch sind, als vielmehr deren Implementation, wie auch in Fachkreisen bereits mehrfach festgestellt wurde (vgl. u. a. BÜLOW-SCHRAMM, 2008; HUBER, 2008).

Employability ist ein wichtiges und breit diskutiertes Stichwort und wird häufig missverständlich ausgelegt: „Employability meint nicht – wie mitunter unterstellt – die Determinierung der Hochschulbildung durch den Arbeitsmarkt, sondern vor allem die Thematisierung und Reflexion des Zusammenhangs von Hochschule und Arbeitsmarkt in einem wissenschaftsbasierten Studium.“ (SCHUBARTH & SPECK, 2014, S. 9)

Die Umstellung der klassischen Prüfungsformate, in denen oftmals vermehrt die Wissensabfrage im Vordergrund stand (eben auch weil dies messtheoretisch leichter ist), zu kompetenzorientierten Prüfungen geht dann einher mit einer Operationali-

sierung von Berufsanforderungen. Verloren geht dabei die eigentliche akademische Kompetenz. Nach WICK (2011) sind akademische Kompetenzen:

- reflexiv (fachspezifisches Handeln ist der Reflexion und wissenschaftlichen Begründung zugänglich),
- explikationsfähig (Situation und Aktivitätsmöglichkeiten werden sorgfältig und deutlich herausgearbeitet – explizite Elaboration–, sodass die Situation reflexiv neu strukturiert werden kann),
- erkenntnisorientiert (methodisch kritisches, theoriegeleitetes Herantreten an Situationen im Bewusstsein der Vorläufigkeit der Ergebnisse/Erkenntnisse),
- disziplinar (Blick aus der Wissenschaft in angrenzende praktische Anwendungsfelder),
- auf neuartige Situationen bezogen (was Kompetenz in Routinesituationen sowie die Bildung von Routinen in Form von grundlegenden Herangehensweisen und den systematischen Einsatz wissenschaftlicher Methoden nicht ausschließt, aber darauf hinweist, dass der akademische Kontext besondere Anforderungen mit sich bringt, wie komplexe und häufig neuartige Situationen) sowie
- tätigkeitsfeldbezogen (Kompetenzen kommen in konkreten – mitunter wissenschaftsbezogenen – Situationen zum Ausdruck).

Wird dieser Definition gefolgt, so wird deutlich, dass es in der hochschulischen Grundbildung nicht darum gehen kann, konkrete Berufsanforderungen in Handlungskompetenzen zu formulieren und diese zu überprüfen (und das mit Employability auch nie gemeint war). Es muss vielmehr darum gehen, berufliche Anforderungen exemplarisch zu nutzen, um die oben genannten akademischen Kompetenzen zu erwerben und diese wiederum anhand von Anwendungsbeispielen aus (wissenschaftlichen) Berufsfeldern zu überprüfen.

## 4 Literaturverzeichnis

**Berendt, B.** (1998). How to Support and to Bring about the Shift from Teaching to Learning through Academic Staff Development Programmes – Examples and Perspectives. In UNESCO – CEPES (Eds.): *Higher Education in Europe, XXIII*, 317–329.

**Bibliographisches Institut GmbH (Duden.de)** (o.J.). *Funktion, die*; <https://www.duden.de/rechtschreibung/Funktion>

**Biggs, J. B.** (2003). *Teaching for quality learning at university: what the student does* (2. Aufl.). Buckingham; Philadelphia, PA: Society for Research into Higher Education. Open University Press.

**Bülow-Schramm, M.** (2008). Hochschuldidaktische Prüfungskritik revisited unter Bologna-Bedingungen. In S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (S. 27–44). Bielefeld: Bertelsmann.

**Chomsky, N.** (1973). Conditions on Transformations. In S. Anderson & P. Kiparsky (Hrsg.), *A Festschrift for Morris Halle* (S 232–286). New York: Holt, Rinehart & Winston.

**Döbler, J.** (2019). *Prüfungsregime und Prüfungskulturen. Soziologische Beobachtungen zur internen Organisation von Hochschule*. Wiesbaden: Springer Verlag.

**Eugster, B., & Lutz, L.** (2004). *Leitfaden für das Planen, Durchführen und Auswerten von Prüfungen an der ETHZ*. Zürich: ETH, Didaktikzentrum.

**Fisch, S.** (2015). *Geschichte der europäischen Universität: von Bologna nach Bologna* (Originalausgabe). München: Verlag C.H.Beck.

**Flehsig, K.-H.** (1976). Prüfungen und Evaluation. *Blickpunkt Hochschuldidaktik*, 40, 303–336.

**Fraenkel-Haeberle, C.** (2014). *Die Universität im Mehrebenensystem: Modernisierungsansätze in Deutschland, Italien und Österreich*. Tübingen: Mohr Siebeck.

**Gnahn, D.** (2010). *Kompetenzen-Erwerb, Erfassung, Instrumente* (2., aktualisierte und überarb. Aufl.). Bielefeld: Bertelsmann.

**Hochschulrektorenkonferenz (HRK)** (1997). *Entschließung des 182. Plenums vom 7.7.1997: Zu Kredit-Punkte-Systemen und Modularisierung*. <https://www.hrk.de/positionen/beschluss/detail/zu-kredit-punkte-systemen-und-modularisierung/>

**Hammerstein, N.** (2007). Vom Interesse des Staates. In R. C. Schwinges & M.-C. Schöpfer Pfaffen (Hrsg.), *Examen, Titel, Promotionen: Akademisches und staatliches Qualifikationswesen vom 13. bis zum 21. Jahrhundert* (S. 169–194). Basel: Schwabe.

**Hofer, M., Ledermüller, K., Lothaller, H., Mitterauer, L., Salmhofer, G., & Vectori, O.** (Hrsg.). (2016). *Kompetenzorientierung und Qualitätssicherung – Konfliktlinien, Interessenslagen, Anwendung*. Bielefeld: UVW Universitätsverlag Webler.

**Huber, L.** (2008). Kompetenzen prüfen. In S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf die Agenda! hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (S. 12–26). Bielefeld: Bertelsmann.

**Kintzinger, M.** (2007). Licentia. Institutionalität „Akademische Grade“ an der mittelalterlichen Universität. In R. C. Schwinges & M.-C. Schöpfer Pfaffen (Hrsg.), *Examen, Titel, Promotionen: Akademisches und staatliches Qualifikationswesen vom 13. bis zum 21. Jahrhundert* (S. 55–88). Basel: Schwabe.

**Klieme, E.** (2004). Was sind Kompetenzen und wie lassen sie sich messen? Auszug aus *Pädagogik*, 6, 10–13.

**Konferenz der europäischen Hochschulministerinnen und -minister** (2003, September 19). *Berlin Communiqué 2003 „Den Europäischen Hochschulraum verwirklichen“*.

**Kvale, S.** (1972). *Prüfung und Herrschaft: Hochschulprüfungen zwischen Ritual und Rationalisierung*. Weinheim: Beltz.

**Matthies, A., & Stock, M.** (2020). *Universitätsstudium und berufliches Handeln. Eine historisch-soziologische Skizze zur Entstehung des „Theorie-Praxis-Problems“*. In C. Scheid & T. Wenzl (Hrsg.), *Wieviel Wissenschaft braucht die Lehrerbildung?* (S. 215–253). Wiesbaden: Springer Verlag.

**Paulsen, F.** (1902). *Die deutschen Universitäten und das Universitätsstudium*. Berlin: Asher.

**Prahl, H.-W.** (1976). *Hochschulprüfungen, Sinn oder Unsinn? Sozialgeschichte und Ideologiekritik der akademischen Initiationskultur*. München: Kösel.

**Prahl, H.-W.** (1995). Prüfungen. In L. Huber (Hrsg.), *Ausbildung und Sozialisation in der Hochschule. Enzyklopädie Erziehungswissenschaft*. Band 10 (2. Aufl., S. 438–450) Stuttgart: Klett-Cotta.

**Prosser, M., & Trigwell, K.** (1999). *Understanding learning and teaching: the ex-*

*perience in higher education*. Buckingham [England]; Philadelphia, PA: Society for Research into Higher Education and Open University Press.

**Reis, O., & Ruschin, S.** (2008). Kompetenzorientiert Prüfen – Baustein eines gelungenen Paradigmenwechsels. In S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (S. 6–9). Bielefeld: Bertelsmann.

**Rüegg, W.** (Hrsg.) (1993). *Geschichte der Universität in Europa: Mittelalter*. München: Beck.

**Schaper, N.** (2012). *Fachgutachten Kompetenzorientierung* (Fachgutachten). Köln: HRK Hochschulrektorenkonferenz.

**Scheer, J. W., & Zenz, H.** (1973). *Studenten in der Prüfung: eine Untersuchung zur akademischen Initiationskultur*. Frankfurt am Main: Aspekte-Verlag.

**Schelsky, H.** (1963). *Einsamkeit und Freiheit. Idee und Gestalt der deutschen Universität und ihrer Reformen*. Reinbeck: Rowohlt Verlag.

**Schubarth, W., & Speck, K.** (2014). *Employability und Praxisbezüge im wissenschaftlichen Studium*. HRK-Fachgutachten ausgearbeitet für die HRK unter Mitarbeit von Juliane Ulbricht, Ines Dudziak und Brigitta Zylla.

**Schütz, M., Skowronek, H., & Thieme, W.** (Hrsg.) (1969). Prüfungen als hochschuldidaktisches Problem. Ergebnisse und Materialien eines Expertenseminars in Hamburg-Rissen vom 31.1.–2.2.1969. *Blickpunkt Hochschuldidaktik, 1*, 96 S.

**Weinert, F. E.** (Hrsg.) (2001). *Leistungsmessungen in Schulen*. Weinheim: Beltz.

**Wick, A.** (2011). Akademisch geprägte Kompetenzentwicklung: Kompetenzorientierung in Hochschulstudiengängen. Heidelberg: HeiDOK. [http://archiv.ub.uni-heidelberg.de/volltextserver/12001/1/Wick\\_Akademisch\\_gepraegte\\_Kompetenzen.pdf](http://archiv.ub.uni-heidelberg.de/volltextserver/12001/1/Wick_Akademisch_gepraegte_Kompetenzen.pdf)

## Autorin



Sabrina GALLNER || Pädagogische Hochschule Luzern, Zentrum für Hochschuldidaktik || Sentimatt 1, CH-6003 Luzern

<https://www.phlu.ch/sabrina-gallner.html>

[sabrina.gallner@phlu.ch](mailto:sabrina.gallner@phlu.ch)



Heidi NIEDERKOFLER<sup>1</sup> (Wien)

# Prüfungsmisserfolg auf die Agenda! Ansätze an der Universität Wien

## Zusammenfassung

Studentische Misserfolgserfahrungen in Folge nicht bestandener Prüfungen beinhalten neben der individuellen Enttäuschung auch ein Lernpotenzial, das didaktisch oft ungenutzt bleibt. In Hinblick auf ein kompetenzorientiertes Prüfungsverständnis wäre es sinnvoll, Studierende mit Wissen und Strategien zu unterstützen, um gezielt an ihren Lernkompetenzen zu arbeiten. Der Beitrag diskutiert den Umgang mit Misserfolg in Zusammenhang mit den ambivalenten Anforderungen des Prüfens und den widersprüchlichen Perspektiven auf Misserfolg selbst, um schließlich konkrete Fallbeispiele eines produktiven Umgangs mit Misserfolgserfahrungen vorzustellen.

## Schlüsselwörter

Misserfolg, Prüfung, Lernen, Kompetenzorientierung

---

<sup>1</sup> E-Mail: [heidi.niederkofler@univie.ac.at](mailto:heidi.niederkofler@univie.ac.at)



## **Put exam failure on the agenda! Approaches at the University of Vienna**

### **Abstract**

Failing an exam brings individual feelings of disappointment, but this experience can also offer a rich learning potential that often goes unrealised. In accordance with a competence-oriented approach to assessment, students should be provided with the resources and strategies they need to work on their learning abilities. This paper discusses the way we deal with failure within the ambivalent demands of assessment and the contradictory perspectives on failure itself. It concludes with specific case studies showing productive ways of dealing with failure experiences.

### **Keywords**

failure, assessment, learning, competence orientation

## **1 Einleitung<sup>2</sup>**

Prüfungen sind aufgrund ihrer Kontrollfunktion von Lernprozessen und der nachgängigen Auswahlfunktion eng mit studentischen Misserfolgserfahrungen verknüpft. Der Umgang mit diesen Erfahrungen ist weitgehend individualisiert: Von Studierenden wird vielfach erwartet, dass sie damit allein zurechtkommen, vonseiten der Lehrenden wie auch der Hochschulen sind Unterstützungsangebote in der Regel selten vorhanden. Das Lernpotenzial, das sich aus Fehlern und Misserfolgen ergeben kann, wie verschiedene Lerntheorien implizit annehmen (OSER, HASCHER & SPYCHIGER, 1999, S. 12), wird jedenfalls hochschuldidaktisch nicht ausgeschöpft. Im Sinne eines kompetenzorientierten Prüfungsverständnisses wäre es jedoch sinnvoll, Studierenden Wissen und Strategien an die Hand zu geben, um gezielt an ihren Lernkompetenzen zu arbeiten. Ausgehend von diesem Befund be-

<sup>2</sup> Für das produktive Feedback von Gisela Kriegler-Kastelic, Reinhard Schott und der anonymen Gutachten möchte ich mich herzlich bedanken. Für die Umsetzung der differenzierten Kommentare bin ich freilich selbst verantwortlich.

---

schäftigt sich das Center for Teaching and Learning (CTL) der Universität Wien aus einer prüfungsdidaktischen Perspektive mit dem Thema Misserfolg.

Ansatzpunkt dieses Beitrags ist Prüfungsmisserfolg als studentische Erfahrung nicht positiv bestandener Prüfungen. Hierzu wird anfangs die Bedeutung des Themas in Hinblick auf eine kompetenzorientierte und studierendenzentrierte Lehr-/Lernkultur thematisiert, anschließend wird die didaktische Vernachlässigung von Misserfolgserfahrungen in Zusammenhang mit den ambivalenten Handlungsanforderungen des Prüfens diskutiert. Misserfolgserfahrungen sind durch die Gleichzeitigkeit von Lerngelegenheiten wie auch von schambehafteten Erlebnissen charakterisiert – FEIGENBAUM (2021) spricht in diesem Zusammenhang von „generative failure“ und von „stigmatized failure“: Ein reflektierter Umgang damit wird schließlich anhand von Ansätzen an der Universität Wien vorgestellt.

## 2 Didaktische Relevanz von Prüfungsmisserfolgserfahrungen

Prüfungsmisserfolgserfahrungen werden didaktisch in der Regel selten strukturiert in den Lernprozess eingebunden, was in mehrfacher Hinsicht erstaunlich ist: Erstens in Hinblick auf den hochschulischen Wandel, der ausgehend vom Bologna-Prozess als Kompetenz- und Lernzielorientierung stattgefunden hat. Das Lernen selbst wurde dabei ins Zentrum gerückt (BARR & TAGG, 1995; BARGEL, 2008). Diese Tendenz wurde in den letzten Jahren noch verstärkt, u. a. durch die sogenannte „21<sup>st</sup> century skills“-Bewegung und jüngst durch den Future Skills-Ansatz (EHLERS, 2020): Lernen und insbesondere das selbstgesteuerte Lernen werden dabei als zentrale Selbstkompetenzen betrachtet. Prüfungen stellen Gelegenheiten dar, um „Reflexionsprozesse zum ‚Lernen lernen‘ grundständig [zu] initiieren und [zu] spiegeln“ (BARGEL 2008, S. 87) und Entwicklungsmöglichkeiten bewusst zu machen.

Obwohl den meisten Lehrenden die Wichtigkeit von Umwegen für das Lernen bewusst ist, werden – und damit komme ich zum zweiten Punkt – studentische Misserfolgserfahrungen selten didaktisch genutzt. Möglicherweise hat das Paradigma des Lernens den Blick noch stärker für das Gelingen und die Gelingensbedingungen der Lernprozesse geschärft. In den Hintergrund getreten ist in diesem Zuge die Einsicht,

dass Bildungsprozesse inhärent ambivalent sind: Sie zielen auf Ergebnisse, funktionieren aber nicht auf Basis von Kausalitätsbeziehungen. Sie entziehen sich vielmehr einer präzise kalkulierten Planbarkeit und verdanken sich hingegen oft Konstellationen, die sich nicht planmäßig erzeugen lassen (RIEGER-LADICH, 2014).

Und schließlich ist geringe Aufmerksamkeit, die Prüfungsmisserfolg erhält, auch hinsichtlich des Prüfprozesses zu beobachten: Dieser endet aus didaktischer Sicht nicht mit der Leistungsfeststellung, sondern umfasst auch die Rückmeldung der Ergebnisse an die Studierenden, die neben der inhaltlichen Rahmung mittels Erwartungshorizont auch die weitere Förderung beinhaltet. Der summative und der formative Charakter von Prüfungen ist hiermit angesprochen. In der Praxis erhalten Studierende kaum Rückmeldungen auf ihre Prüfungsleistung, die über die Bekanntgabe der Note hinausgehen (MÜLLER, 2012): Diese Leerstelle dürfte bei nicht bestandenen Prüfungen insofern schwerwiegender sein, da Studierende ihre Prüfungsleistung vielfach nicht mit den Leistungsanforderungen in Bezug setzen können und schlicht nicht wissen, wo ihre Fehler liegen und wie eine „richtige“ Beantwortung bzw. eine positive Leistung aussehen würde. Die oftmals fehlende Orientierung erhöht zudem die Prüfungsunsicherheit bei einem nächsten Antritt. Die Priorisierung der Leistungsmessung gegenüber der Rückmeldung an die Studierenden (MÜLLER, 2012, S. 128) hat viele Gründe: Ein Faktor sind jedenfalls die widersprüchlichen Anforderungen an das Prüfen selbst, die sowohl das Messen und Beurteilen, aber auch das Fördern beinhalten (REH & RICKEN, 2017). Lehrende sind mit diesen nicht aufzulösenden Widersprüchen konfrontiert, die besonders bei Rollenwechseln zwischen Wissensvermittlung und Beurteilung als Belastung wahrgenommen werden (HAINSCHINCK & ZAHRA-ECKER, 2018). Rückmeldungen zu negativen Prüfergebnissen, die über die Bekanntgabe der Note hinausgehen und Studierenden einen lernförderlichen Umgang damit ermöglichen, sind aufgrund der engen Verknüpfung der beiden genannten Anforderungen an das Prüfen ein Paradebeispiel für die damit verbundene Widersprüchlichkeit – und unterbleiben demzufolge oft.<sup>3</sup>

---

3 Um die Kommunikation zum Prüfprozess insgesamt zu stärken, wäre eine ausführliche Rückmeldung der Studierenden zur Qualität von Prüfungen (ähnlich der Befragungen zur Qualität der Lehre) ein weiterer Schritt. Damit wäre auch auf die mit Prüfungen verbundenen Lernprozesse der Lehrenden verwiesen.

### 3 Misserfolg – ein ambivalentes hochschuldidaktisches Thema

Das Nutzen von Fehlern im Lernprozess geht weit zurück: Oft wird Piagets kognitive Entwicklungstheorie genannt, die bereits implizit Hinweise auf die Bedeutung von Fehlversuchen für das Finden von Lösungen und das damit verbundene Entwicklungs- und Lernpotenzial enthält (WEINGARDT, 2004, S. 65f.). Spätestens mit konstruktivistischen Lernmodellen werden Fehler als Verbündete im individuellen Lernprozess betrachtet, zudem hat die in den letzten Jahrzehnten proklamierte Fehleroffenheit in hochschulischen wie unternehmerischen Forschungs- und Entwicklungsabteilungen den Weg in Richtung „Fehlerneubewertung“ geebnet (WEINGARDT, 2004, S. 275f.). Vor diesem Hintergrund haben sich verschiedene Hochschullehrende mit Fehlern und Misserfolg als Voraussetzung für das Lernen und die Entstehung neuer Erkenntnisse beschäftigt (z. B. BURGER, 2012; FEIGENBAUM, 2021; FERRANDINO, 2016; CARR, 2013; KAPUR & BIELACZYK, 2012).

Ein fehleroffener Zugang zu Fehlern und Misserfolg, der vordergründig das Lernen und weniger die Beurteilung ins Zentrum stellt, stößt an Herausforderungen: Zum einen werten Hochschulen Fehler und Misserfolge ambivalent, nicht zuletzt aufgrund der Selektionsfunktion von Prüfungen (FEIGENBAUM, 2021). Selbst wenn der „Mehrwert“ von Fehlversuchen für den Lernprozess genutzt wird, zählt im Ernstfall der Prüfung die Fehlervermeidung und es wird fehlerorientiert zwischen „Richtig“ und „Falsch“ unterschieden. ERICHSON (1987) hat diese ambivalente Vorgehensweise das „toll-aber-Ritual“ genannt hat: Fehleroffenheit hat demnach für Lernende den Charakter einer „netten Übung“, letztlich aber auch aufgrund der institutionell geförderten Prüfungsorientierung wenig Relevanz.<sup>4</sup>

Zum anderen können Misserfolgserfahrungen auf einer individuellen Ebene sehr unterschiedlich erlebt werden: Sie können das Lernen erschweren, darüber hinaus können sie auch weitergehende Auswirkungen haben, die von der Infragestellung der Studien- und Bildungswahl bis hin zum Studienabbruch reichen. HALLMARK

---

4 Um die Orientierung am Lernprozess gegenüber der Ausrichtung auf die Prüfung zu stärken, werden verschiedene Wege aus diesem Dilemma überlegt: Nennen möchte ich die Entkoppelung von Lehre und Prüfungen bzw. Beurteilungen (z. B. HUBER, 2008; EHLERS, 2020), oder die Reduktion von Prüfungen mit Rechtsfolgen (REINMANN, 2022).

(2018) weist darauf hin, dass dies vor allem für weniger privilegierte Studierende zutreffend ist, insbesondere first-generation-students sowie Personen aus marginalisierten gesellschaftlichen Gruppen. Ansätze, die einen offenen Umgang mit Misserfolg proklamieren, ohne dabei eine privilegienkritische Perspektive einzunehmen, „ignore the fact that failure affects people differently, and that privilege plays an important role in who is allowed to fail – and who isn’t.“ (HALLMARK, 2018, S. 1).

Misserfolg kann auf der einen Seite als beschämend und stigmatisierend erlebt werden, auf der anderen Seite gelten Fehler und Misserfolge als integrale Teile eines iterativen Lernprozesses. FEIGENBAUM (2021) spricht in diesem Zusammenhang von „dissonant paradigms of failure“ und stellt fest, dass es besonders für weniger privilegierte und marginalisierte Studierende herausfordernd ist, mit dieser Ambivalenz umzugehen. Die gleichzeitige Präsenz einer potenziell stigmatisierenden Erfahrung („stigmatized failure“) und einer Lerngelegenheit („generative failure“) in der Misserfolgserfahrung ist kognitiv wie emotional irritierend und anspruchsvoll und bedarf im Umgang verschiedener Reflexions- und Handlungsebenen. Um letztlich produktiv mit Misserfolg umgehen zu können, so FEIGENBAUM, müssen die Ambivalenz aufgegriffen und die stigmatisierenden Effekte von Misserfolg ernst genommen und (idealerweise individuell) berücksichtigt werden.

## 4 Produktive Ansätze zum Umgang mit studentischen Misserfolgserfahrungen

Auf Basis der vorgestellten Überlegungen entwickelte das CTL der Universität Wien bislang erste, direkt an Studierende und Studienbewerber/innen adressierte Maßnahmen.<sup>5</sup> Ein systematisches Aufgreifen des Themas bedarf längerer Zeiträume, ist aber nicht zuletzt vor dem Hintergrund, dass Misserfolg institutionell betrachtet teilweise ein (ökonomisches) Problem darstellt, das von der Erfüllung von Kennzahlen bis zu den gesellschaftlichen Kosten von hohen Drop-out-Raten reicht, sinnvoll.

---

<sup>5</sup> Impulsgebend dafür war letztlich der Vortrag von Ulrich FRISCHKNECHT (2021) zum Thema „Misserfolgsmanagement“.

## 4.1 Blogbeitrag „Prüfungsmisserfolg – Aus Erfahrungen lernen“

Ein wichtiger Faktor beim Umgang mit Misserfolg ist, diesen wahrzunehmen und zu analysieren. OSER & SPYCHIGER (2005, S. 72 ff.) gehen darüber hinaus und behaupten, dass ohne die Beteiligung von Emotionen kein Wissen um und Lernen aus Fehlern erfolgen kann.<sup>6</sup> Wichtig ist jedenfalls, dass eine ernsthafte Beschäftigung mit dem Misserfolg passiert, damit aus den gemachten Erfahrungen gelernt werden kann. Misserfolg ernst zu nehmen bedeutet auch anzuerkennen, dass Lernen anstrengend, schwierig und herausfordernd ist. Das CTL hat für den Blog der Universität Wien<sup>7</sup> und die damit verbundenen Social-Media-Kanäle einen Beitrag verfasst mit dem Ziel, Studierende beim strukturierten Umgang mit und der Reflexion der misslungenen Prüfung zu unterstützen.

Der Beitrag ist so aufgebaut, dass er Misserfolg nicht als etwas Banales behandelt, sondern auch auf mögliche weitreichende Folgen hinweist. Zugleich wird betont, dass die Misserfolgserfahrung ein Lernpotenzial hat, das es zu ergründen gilt. Verschiedene Schritte werden genannt, die Studierende dabei unterstützen, mit der Erfahrung umzugehen und aus ihr zu lernen: von der Wahrnehmung der Gefühle über die Situationsanalyse und Prüfungseinsicht bzw. Rückmeldung bis hin zu nächsten Schritten. Wichtig sind dabei auch Hinweise auf die strukturelle Individualisierung von Misserfolgserfahrungen, die nicht zu einem offenen Umgang beiträgt, sowie Verweise auf bereits vorhandene Unterstützungsstrukturen.

---

6 OSER & SPYCHIGER (2005) sprechen nicht explizit von Fehlern, sondern von „negativem Wissen“, das aufzeigt, was nicht ist, was nicht funktioniert oder warum etwas nicht stimmt. Da sie in ihrer Theorie des Negativen Wissens Fehler(wissen) und negatives Wissen zwar nicht explizit gleich definieren, aber faktisch nahezu gleich verwenden, scheint mir obige Darstellung gerechtfertigt.

7 <https://blog.univie.ac.at/studium/pruefungsmisserfolg-aus-erfahrungen-lernen/>, Stand vom 15. Februar 2022.

## 4.2 Merkblatt Prüfungseinsicht

Studierende haben in Österreich bei schriftlichen Prüfungen die rechtlich abgesicherte Möglichkeit zur Prüfungseinsicht innerhalb von sechs Monaten ab Bekanntgabe der Beurteilung. Die an der Universität Wien seit 2017 bestehende Feedback-Policy wirkt dahingehend, Studierende mittels Rückmeldungen in ihren Lernprozessen zu begleiten. Gerade bei Prüfungsmisserfolg unterstützt die Einsicht in die Beurteilungsgrundlagen die Situationsanalyse: Erstens hilft das Wissen darum, ob ein knappes oder ein deutliches Scheitern für den Misserfolg verantwortlich ist, dabei, die eigene Leistung besser einschätzen zu können; zweitens trägt das Erkennen der Fehler dazu bei, die Inhalte besser zu verstehen und Schlüsse für das eigene Lernen daraus zu ziehen; und schließlich ist dieses gesamte Wissen notwendig, um auf einer validen Basis die eigene Situation zu klären und Entscheidungen treffen zu können. Kurz gesagt, bei Misserfolg ist Feedback essentiell, um die Erfahrung einordnen und kompetent damit umgehen zu können. Da die Möglichkeit zur Prüfungseinsicht von Studierenden relativ selten in Anspruch genommen wird und demzufolge auch wenig Praxiswissen besteht, hat das CTL das „Merkblatt Prüfungseinsicht“<sup>8</sup> erarbeitet mit dem Ziel, Tipps und Hinweise zu geben, wie eine Prüfungseinsicht möglichst gut genutzt werden kann.

Das Merkblatt ist folgendermaßen ausgerichtet: Ausgehend von der Feststellung, dass Prüfungseinsicht trotz des damit verbundenen zeitlichen Aufwands und der Auseinandersetzung mit einem womöglich emotional schwierigen Thema eine Chance darstellt, werden einzelne Schritte für die Vorbereitung, die Prüfungseinsicht selbst und die Nachbereitung vorgestellt. Studierende werden dadurch angeleitet, die misslungene Prüfung detailliert zu betrachten, im Zuge dessen sowohl inhaltliches Wissen wie auch Prüfungswissen zu sammeln und in der Reflexion hilfreiche Erkenntnisse und weitere Vorgehensweisen festzuhalten.

---

<sup>8</sup> [https://ctl.univie.ac.at/fileadmin/user\\_upload/z\\_ctl\\_neu/Angebote\\_fuer\\_Lehrende/Wissenschaftliches\\_Schreiben\\_und\\_Arbeiten/Materialien\\_fuer\\_die\\_Lehre/Team\\_Schreiben/Handouts\\_Studierende/211207\\_Merkblatt\\_Pruefungseinsicht.pdf](https://ctl.univie.ac.at/fileadmin/user_upload/z_ctl_neu/Angebote_fuer_Lehrende/Wissenschaftliches_Schreiben_und_Arbeiten/Materialien_fuer_die_Lehre/Team_Schreiben/Handouts_Studierende/211207_Merkblatt_Pruefungseinsicht.pdf), Stand vom 15. Februar 2022.

### 4.3 Tipps für abgelehnte Studienbewerber/innen

Während sich die zwei erstgenannten Ansätze an Personen wenden, die bereits Studierende sind, richtet sich die folgende Maßnahme an Studienbewerber/innen, die bei zugangsbeschränkten Fächern mit Aufnahme- und Eignungsverfahren keinen Studienplatz erhalten haben.

Studienbewerber/innen erhalten nach den absolvierten Aufnahmeverfahren Informationen darüber, wie sie bei den Tests abgeschnitten haben und was ihre Leistung in Hinblick auf die zur Verfügung stehenden Studienplätze bedeutet. Bei den Verfahren im Jahr 2021 wurde erstmals versucht, die nicht aufgenommenen Bewerber/innen beim Umgang mit dieser Misserfolgserfahrung zu unterstützen. In standardisierten Schreiben wurden Sie auf diese Seite verwiesen, Tipps für abgelehnte Bewerber\*innen (univie.ac.at), die drei verschiedene Ziele verfolgt: Erstens wird von institutioneller Seite Empathie für die betreffenden Personen ausgedrückt und damit gezeigt, dass die mögliche Tragweite der Ablehnung bewusst ist. Zweitens werden häufige Gründe für ein nicht ausreichendes Testergebnis genannt und dabei konkrete verbesserungsorientierte Hinweise und Tipps gegeben. Und drittens wird schließlich dazu eingeladen, auf Basis der vorhandenen Informationen eine persönliche Analyse vorzunehmen, um davon ausgehend weitere Entscheidungen treffen zu können.

Ein wichtiger Punkt in diesem Prozess ist – und hier stütze ich mich auf FRISCH-KNECHT – die kommunikative Präsenz der Universität. Die Universität Wien kommuniziert den abgelehnten Studienbewerber/innen nicht nur Informationen zur Nicht-/Aufnahme, sondern unterstützt bei der Reflexion und übernimmt somit Verantwortung für den gesamten Prüfprozess.

## 5 Ausblick

Lernen ist ein komplexer Prozess, der nicht immer linear erfolgt und Umwege beinhaltet, Fehler und Misserfolge bei Prüfungen gehören dazu. Bemühungen um eine fehleroffene Lernkultur finden allerdings in einem hochschulischen Kontext statt, wo die letztlich bestimmenden Prüfungen grundsätzlich fehlerfeindlich angelegt sind – und es ist nicht davon auszugehen, dass sich die Prüfkultur in absehbarer Zeit grundlegend ändern wird. Vor diesem Hintergrund ist es in einer kompetenzorientierten Lehr-/Lernkultur zum einen wichtig, Studierende auf diese Dissonanz vorzubereiten und ihre Ambiguitäts-, Selbst- und Reflexionskompetenz dahingehend zu stärken, mit den zwei Aspekten von Misserfolg (stigmatized und generative failure) kompetent und situationsadäquat umgehen zu können. Zum anderen geht es darum, Studierende dabei zu unterstützen, ihre Lernkompetenzen und Fähigkeiten nicht zuletzt durch Misserfolgserfahrungen einschätzen zu lernen und überlegte Entscheidungen zu treffen. Denn: Misserfolg bietet zwar Lernpotenzial, wird aber existentieller und bedrohlicher, je weiter fortgeschritten Studierende im Studium sind.

## 6 Literaturverzeichnis

**Bargel, S.** (2008). Prüfungen in Transformation – Kleiner Erfahrungsbericht in wandelndem-gewandeltem Selbst- bzw. Rollenverständnis. In: S. Dany, B. Szczyrba & J. Wildt (Hrsg.), *Prüfungen auf die Agenda!* Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen (S. 74–90). Bielefeld: Bertelsmann Verlag.

**Barr, R. B. & Tagg, J.** (1995). From Teaching to Learning – A New Paradigm for Undergraduate Education. *Change* 27(6), 13-23.

**Burger, E.** (2012). Essay on the importance of teaching failure. *Inside Higher Ed* <https://www.insidehighered.com/views/2012/08/21/essay-importance-teaching-failure>

**Carr, A.** (2013). In support of failure. *Compositum Forum* 27. <https://compositionforum.com/issue/27/failure.php>

**Ehlers, U.-D.** (2020). *Future Skills. Lernen der Zukunft – Hochschule der Zukunft*. Karlsruhe: Springer Verlag. <https://link.springer.com/book/10.1007/978-3-658-29297-3>

- Erichson, C.** (1987). Aus Fehlern soll man klug werden? Überlegungen zum Lernen aus Fehlern. In H. Balhorn & H. Bragelmann (Hrsg.), *Welten der Schrift in der Erfahrung der Kinder* (S. 148–157). Konstanz: Faude.
- Feigenbaum, P.** (2021). Telling Students it's O.K. to Fail, but Showing Them it isn't. Dissonant Paradigms of Failure in Higher Education. *Teaching & Learning Inquiry*, 9(1), 13–26.
- Ferrandino, J. A.** (2016). Student Achievement in Undergraduate Statistics: The Potential Value of Allowing Failure. *Journal of the Scholarship of Teaching and Learning*, 16(6), 1–18.
- Frischknecht, U.** (2021). Misserfolgsmanagement. Workshop im Rahmen der Tagung „Studieneingangsphase“ am 19. Mai 2021 in Zürich. [https://studieneingangsphase.ch/wp-content/uploads/2021/05/misserfolgsmanagement\\_20210519.pdf](https://studieneingangsphase.ch/wp-content/uploads/2021/05/misserfolgsmanagement_20210519.pdf)
- Hainschink, V., & Zahra-Ecker, R. A.** (2018). Leben in Antinomien. Bewältigungsdispositionen aus arbeitsbezogenen Verhaltens- und Erlebensmustern. *Pädagogische Horizonte*, 2(2), 179–194.
- Hallmark, T.** (2018). When ‚Failure Is OK‘ Is Not OK. *Chronicle of Higher Education*, February, 11.
- Huber, L.** (2008). ‚Kompetenzen prüfen‘. In S. Dany, B. Szczyrba & J. Wild (Hrsg.), *Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (S. 12–26). Bielefeld: Bertelsmann.
- Kapur, M., & Bielaczyc, K.** (2012). Designing for Productive Failure. *The Journal of the Learning Sciences*, 21(1), 45–83.
- Müller, F.** (2012). Prüfen an Universitäten. Wie Prüfungen das Lernen steuern. In B. Kossek & C. Zwiauer (Hrsg.), *Universität in Zeiten von Bologna. Zur Theorie und Praxis von Lehr- und Lernkulturen* (S. 121–132). Göttingen: V&R unipress, Vienna University Press.
- Oser, F., Hascher, T., & Spychiger, M.** (1999). Lernen aus Fehlern. Zur Psychologie des „negativen“ Wissens. In W. Althof (Hrsg.), *Fehlerwelten. Vom Fehlermachen und Lernen aus Fehlern* (S. 11–41). Opladen: Leske + Budrich.
- Oser, F., & Spychiger, M.** (2005). *Lernen ist schmerzhaft. Zur Theorie des Negativen Wissens und zur Praxis der Fehlerkultur*. Weinheim: Beltz Verlag.
- Reh, S., & Ricken, N.** (2017). Prüfungen – Systematische Perspektiven der Geschichte einer pädagogischen Praxis. *Zeitschrift für Pädagogik*, 63(3), 247–258.

**Reinmann, G.** (2022). Prüfung oder Assessment an Hochschulen? Thesen für einen Wandel der Prüfungskultur. In J. Gerick, A. Sommer & G. Zimmermann (Hrsg.), *Kompetent Prüfungen gestalten. 53 Prüfungsformate für die Hochschullehre* (2. überarb. u. erw. Aufl.). Stuttgart: UTB. (erscheint im April 2022).

**Rieger-Ladich, M.** (2014). Auffälliges Vermeidungsverhalten: Scheitern als Gegenstand des pädagogischen Diskurses. In R. John & A. Langhoff (Hrsg.), *Scheitern – Ein Desiderat der Moderne? Innovation und Gesellschaft* (S. 279–299). Wiesbaden: Springer Verlag.

**Weingardt, M.** (2004). *Fehler zeichnen uns aus. Transdisziplinäre Grundlagen zur Theorie und Produktivität des Fehlers in Schule und Arbeitswelt*. Bad Heilbrunn: Klinkhardt.

## Autorin



Dr. Heidi NIEDERKOFLER || Universität Wien, Center for Teaching and Learning || Augasse 2-6, A-1090 Wien

<https://ctl.univie.ac.at/>

[heidi.niederkofler@univie.ac.at](mailto:heidi.niederkofler@univie.ac.at)

Uta MOHRING<sup>1</sup>, Christoph JACOBI & Kai FURMANS (Karlsruhe)

# Kompetenzorientiertes Lehren und Prüfen in den Ingenieurwissenschaften: Lehrkonzept der Veranstaltung Materialfluss in Logistiksystemen

## Zusammenfassung

Der Beitrag stellt das Lehrkonzept einer Wahlpflichtveranstaltung der Bachelor- und Masterstudiengänge Maschinenbau, Mechatronik und Wirtschaftsingenieurwesen mit ca. 30 Studierenden pro Jahrgang vor. Das Lehrkonzept hat zum Ziel, dass Studierende Kompetenzen zur Lösung ingenieurwissenschaftlicher Probleme durch die selbstständige Bearbeitung offener Fragestellungen nach wissenschaftlichen Standards erwerben. Hierzu werden Elemente des *Constructive Alignment*, des *Inverted Classroom* und des kompetenzorientierten Prüfens kombiniert. Lehr-evaluationen und studentisches Feedback bestätigen, dass die Ziele des Lehrkonzepts erreicht werden und die Studierenden berufliche Handlungskompetenz erwerben.

## Schlüsselwörter

Constructive Alignment, Inverted Classroom, Kompetenzorientiertes Prüfen, Offene Fragestellungen, Kolloquium

---

<sup>1</sup> E-Mail: [mohring@kit.edu](mailto:mohring@kit.edu)



## **Competence-oriented teaching and examination in engineering: Teaching concept of the course Material Flow in Logistics Systems**

### **Abstract**

The paper presents the teaching concept of an elective course in the bachelor's and master's programs of Mechanical Engineering, Mechatronics, and Industrial Engineering, which enrolls approx. 30 students per year. The vision of the teaching concept is that the students should acquire competencies for solving engineering problems by working autonomously on open-ended questions according to scientific standards. To this end, the teaching concept combines elements of constructive alignment, inverted classroom and competence-oriented examinations. Teaching evaluations and student feedback confirm that the teaching concept achieves its aim and that the students gain employability.

### **Keywords**

constructive alignment, inverted classroom, competence-oriented examination, open-ended questions, colloquium

## **1 Einleitung**

Der Beitrag stellt ein Lehrkonzept vor, dessen Ziel es ist, dass Studierende der Ingenieurwissenschaften des Karlsruher Instituts für Technologie (KIT) Kompetenzen für die effektive Planung und den effizienten Betrieb der zukünftigen Logistiksysteme erwerben. Den Kern des Lehrkonzepts bildet die selbstständige Bearbeitung offener Fragestellungen sowie die Prüfung der hierdurch erworbenen Kompetenzen, also die „kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (WEINERT, 2014, S.27). Mit dem Lehrkonzept wird die These vertreten, dass Studierende durch die selbstständige Bearbeitung offener Fragestellungen nach wissenschaftlichen Standards Kompeten-

zen zur Lösung ingenieurwissenschaftlicher Probleme erwerben. Offene Fragestellungen zeichnen sich dadurch aus, dass keine eindeutige bzw. eindeutig richtige Lösung existiert. Folglich umfasst eine tragfähige Lösung einer offenen Fragestellung die begründete Auswahl und Anwendung eines geeigneten Lösungsansatzes sowie die kritische Diskussion anderer denkbarer Lösungsansätze. Unsere Erfahrung aus der industriellen und wissenschaftlichen Praxis zeigt, dass die Fähigkeit, tragfähige Lösungen offener Fragestellungen erarbeiten und diskutieren zu können, ein wesentlicher Erfolgsfaktor der *employability*, der beruflichen Handlungskompetenz (WILDT & WILDT, 2011, S.5), darstellt. Die Umgestaltung der Lehrveranstaltung ist insbesondere durch die anhaltende Diskussion um den Paradigmenwechsel *shift from teaching to learning* (BARR & TAGG, 1995) sowie die Forderung der aktiven Einbindung und partizipativen Mitgestaltung der Studierenden an Lehrveranstaltungen (SPANNAGEL & FREISLEBEN-TEUTSCHER, 2016) getrieben.

Im vorgestellten Lehrkonzept sind die Lehr- und Lernmethoden sowie die Prüfungen nach dem *Constructive Alignment* (BIGGS & TANG, 2007, S. 52f.) auf die angestrebten Lernziele abgestimmt (vgl. Abb. 1). Daher werden im Folgenden zunächst die Lernziele definiert (Abschnitt 2) und hieraus die Lehrmethode (Abschnitt 3) und die kompetenzorientierten Prüfungen (Abschnitt 4) abgeleitet. In Abschnitt 5 werden die Evaluationsergebnisse der Lehrveranstaltung vorgestellt und eingeordnet. Abschnitt 6 fasst die Erkenntnisse zusammen.

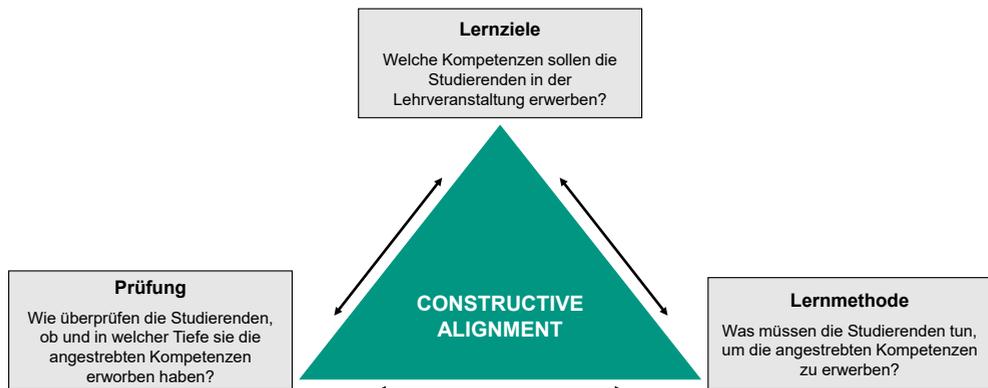


Abb. 1: Constructive Alignment (in Anlehnung an BIGGS & TANG, 2007, eigene Darstellung)

## 2 Lehrveranstaltung und Lernziele

Die Lehrveranstaltung „Materialfluss in Logistiksystemen“ des Instituts für Fördertechnik und Logistiksysteme (IFL) ist eine Wahlpflichtveranstaltung im Vertiefungsstudium der Bachelor- und Masterstudiengänge Maschinenbau, Mechatronik und Wirtschaftsingenieurwesen des KIT, die in jedem Jahrgang etwa 30 Studierende belegen.

Materialflusssysteme finden sich in nahezu allen Bereichen von Produktions- und Logistiksystemen, z. B. in der Automobilindustrie, einem Distributionszentrum für E-Commerce oder dem Lagerbereich eines Supermarktes. Allen Anwendungen ist gemein, dass konzeptionelle Fragestellungen nach Ablauf und Gestaltung der Prozesse und deren Kostenwirksamkeit mit Forderungen nach kurzen Durchlaufzeiten und niedrigen Beständen einhergehen. In der Materialflusslehre werden die real oft komplexen Materialflusssysteme in vereinfachte Modelle überführt und auf diese Weise Lösungen für praktische Probleme entwickelt (ARNOLD & FURMANS, 2019, S. 2f.). Hierzu zählen z. B. die Auslegung von Fördersystemen und Pufferplätzen in einer Produktion, die Auswahl passender technischer Systeme zur Lagerung und Transport von Gütern sowie die Berechnung der Leistungsfähigkeit dieser Systeme.

Die Lehrveranstaltung verfolgt das Ziel, dass Studierende Kompetenzen erwerben, um Materialflusssysteme planen, analysieren, simulieren und (eingeschränkt) bauen und betreiben zu können. Hierzu werden folgende Lernziele definiert: „Nach erfolgreichem Abschluss der Lehrveranstaltung können Sie alleine und im Team:

- Fachkundigen ein Materialflusssystem zutreffend beschreiben,
- die Systemlast und die typischen Materialflusselemente modellieren und parametrieren,
- hieraus ein Materialflusssystem für eine Aufgabe konzipieren,
- die Leistungsfähigkeit einer Anlage in Bezug auf die Anforderungen qualifiziert beurteilen,
- die wichtigsten Stellhebel zur Beeinflussung der Leistungsfähigkeit gezielt verändern,
- die Grenzen der heutigen Methoden und Systemkomponenten konzeptionell bei Bedarf erweitern.“

Die Definition der Lernziele orientiert sich an der Lerntaxonomie nach Bloom (BLOOM et al., 1972, und ANDERSON et al., 2001) und adressiert unterschiedliche Stufen der erworbenen Kompetenzen. Die Grundlage bildet der sachgemäße Gebrauch der Fachsprache, damit die Studierenden in der Lage sind, mit anderen Experten über Materialflusssysteme zu diskutieren. Weiterhin sollen die Studierenden in der Lage sein, einzelne technische Bestandteile von Materialflusssystemen so zu parametrieren, dass diese die geplante Systemlast bewerkstelligen können, um hieraus eine gesamte Anlage für eine Materialflussaufgabe zu konzipieren. Umgekehrt muss eine gegebene Anlage in Bezug auf ihre Leistungsfähigkeit analysiert und beurteilt werden können. Sollte diese Beurteilung nicht zufriedenstellend ausfallen, sollen die Studierenden in der Lage sein, die Leistungsfähigkeit gezielt zu beeinflussen. Da sich sowohl die Anforderungen als auch die technischen Möglichkeiten stetig weiterentwickeln, sollen die Studierenden zuletzt in der Lage sein, die Grenzen der heutigen Systeme und Methoden zu erkennen und selbstständig zu erweitern.

### 3 Lehrmethode

Die Lehrmethode orientiert sich am Konzept des *Inverted Classroom*, dessen Kern es ist, die Rollen und Lernaktivitäten inner- und außerhalb des Hörsaals im Vergleich zur konventionellen Lehre umzudrehen: In der Lernphase eignen sich die Studierenden zunächst selbstständig die Lerninhalte an, die in der anschließenden Vertiefungsphase gemeinsam mit dem Lehrenden vertieft und angewendet werden. Der Fokus liegt hierbei auf einer aktiven Beteiligung und einer partizipativen Mitgestaltung der Studierenden an der Präsenzveranstaltung (LAGE et al., 2000, und SPANNAGEL, 2013).

Die Lerninhalte der Veranstaltung werden in fünf thematische Blöcke unterteilt, die aufeinander aufbauend nacheinander behandelt werden. Jeder Themenblock ist in eine Lern- und eine Vertiefungsphase untergliedert. In der Lernphase erarbeiten sich die Studierenden die Lerninhalte im *Selbststudium* asynchron, ortsunabhängig, individuell, selbstgesteuert, in eigenem Lerntempo und mit individueller Lernstrategie anhand der bereitgestellten Lehrmaterialien. In der anschließenden Vertiefungsphase wird das erlernte Wissen im Rahmen einer *Präsenzveranstaltung* gemeinsam mit den Lehrenden interaktiv anhand von Übungsaufgaben vertieft und erstmalig

auf realitätsnahe offene Fragestellungen angewandt. Die Präsenzveranstaltung bereitet die Studierenden auf die sich anschließende selbstständige *Bearbeitung einer Fallstudie* in Gruppen vor. Abb. 2 gibt einen Überblick über die Lehrmethode, deren didaktische Bausteine im Folgenden im Einzelnen vorgestellt werden.

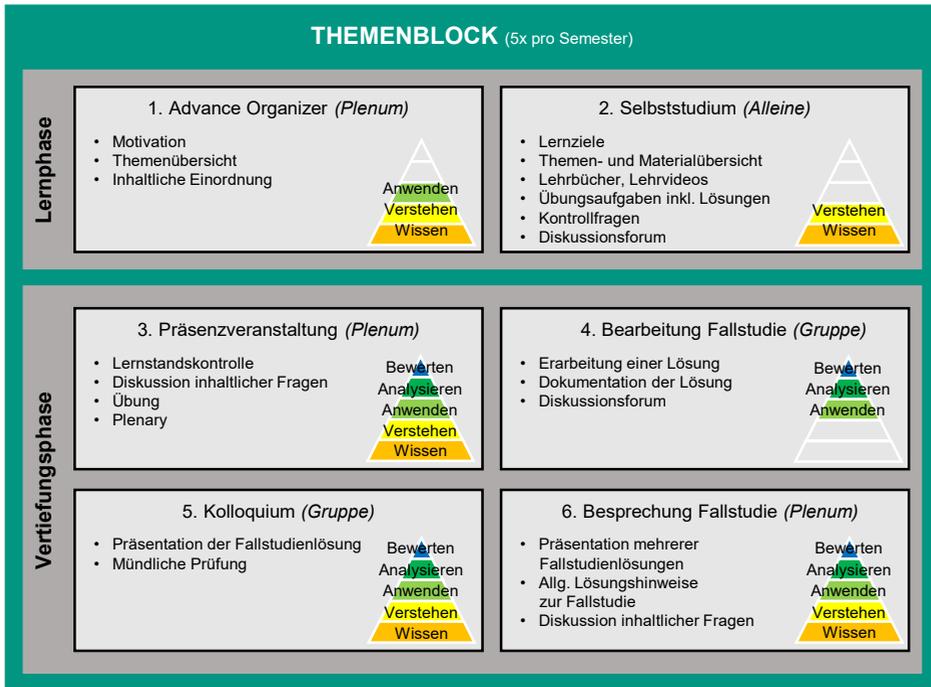


Abb. 2: Didaktische Bausteine der Lehrmethode

### 3.1 Advance Organizer

Der Themenblock beginnt mit einem Advance Organizer (AUSUBEL, 1974), der einen ersten inhaltlichen Überblick über den Themenblock gibt. Ausgehend vom Advance Organizer stellen die Lehrenden im Plenum die wesentlichen Lerninhalte des Themenblocks vor, setzen diese zu den Lerninhalten der anderen Themenblöcke in Bezug und zeigen anhand einiger konkreter Anwendungsbeispiele die Relevanz der theoretischen Lerninhalte in der Praxis auf.

### 3.2 Selbststudium

Während des einwöchigen Selbststudiums erarbeiten sich die Studierenden selbstständig die Lerninhalte des Themenblocks. Hierfür werden themenblockspezifische Lernziele definiert und Lehrbücher, Lehrvideos sowie eine strukturierte Themen- und Materialübersicht zur Verfügung gestellt. Mithilfe von Kontrollfragen können die Studierenden während der Lernphase selbstständig die Erreichung der Lernziele überprüfen. Eine Sammlung von Übungsaufgaben mit Lösungswegen ermöglicht es den Studierenden, das erworbene Wissen erstmalig anzuwenden. Ein Diskussionsforum auf einer digitalen Lernplattform bietet den Studierenden die Möglichkeit, während des Selbststudiums Fragen mit Mitstudierenden und Lehrenden zu diskutieren.

### 3.3 Präsenzveranstaltung

Die Präsenzveranstaltung dient der gemeinsamen und interaktiven Vertiefung und Anwendung des im Selbststudiums erworbenen Wissens. Zunächst wird im Rahmen der *Lernstandskontrolle* der aktuelle Wissensstand anhand ausgewählter Multiple-Choice-Fragen überprüft. Die Lehrenden identifizieren so die inhaltlichen Lücken der Studierenden und können auf diese im Folgenden schwerpunktmäßig eingehen.

Der Fokus der anschließenden Übung liegt auf der Vertiefung und Einübung des „Handwerkszeugs“ des Themenblocks. Die Studierenden bearbeiten selbstständig – in Einzelarbeit oder in Gruppen – Übungsaufgaben unterschiedlichen Schwierigkeitsgrads, die überwiegend die Stufen „Wissen“, „Verstehen“ und „Anwenden“ der Lerntaxonomie nach Bloom adressieren. Die Lehrenden unterstützen die Studieren-

den je nach individuellem Bedarf und besprechen die Lösungen der Übungsaufgaben anschließend gemeinsam im Plenum.

Im anschließenden *Plenary* werden erstmals realitätsnahe offene Fragestellungen des Themenblocks thematisiert. Die Studierenden erwerben die zu deren Lösung notwendigen Problemlösungskompetenzen und wenden diese erstmalig an. Hierzu bearbeiten die Studierenden selbstständig – in Einzelarbeit oder in Gruppen – Aufgaben mit offener Fragestellung der Stufen „Anwenden“, „Analysieren“ und „Bewerten“ der Lerntaxonomie nach Bloom. Die Lehrenden unterstützen die Studierenden je nach individuellem Bedarf. Die Studierenden tragen ihre unterschiedlichen Lösungsansätze und Ergebnisse anschließend im Plenum zusammen, um diese gemeinsam zu diskutieren und einzuordnen.

### **3.4 Bearbeitung der Fallstudie**

Nach der Präsenzveranstaltung sind die Studierenden in der Lage, die Lerninhalte des Themenblocks im Rahmen einer Fallstudie auf realitätsnahe offene Fragestellungen zu übertragen und für diese tragfähige Lösungen zu entwickeln. Die Fallstudie wird in Gruppen von jeweils 4–5 Studierenden bearbeitet. Die Studierenden erarbeiten gruppenindividuelle Fallstudienlösungen, die in Abhängigkeit von der konkreten Aufgabenstellung die getroffenen Annahmen, den gewählten Lösungsansatz, dessen Anwendung auf die betrachtete Fragestellung sowie die Ergebnisse und die daraus resultierenden Implikationen umfassen. Die Studierenden werden hierbei durch ein Diskussionsforum auf der digitalen Lernplattform unterstützt. Dort können Fragen zur Fallstudie mit Mitstudierenden und Lehrenden diskutiert werden. Die erarbeitete Fallstudienlösung muss bis zu einem festgelegten Termin abgegeben werden und ist Teil der Prüfungsleistung (vgl. Abschnitt 4.1).

### **3.5 Kolloquium**

Das Kolloquium findet in den Fallstudiengruppen statt. Zunächst stellen die Studierenden in einer fünfminütigen Kurzpräsentation ihre erarbeitete Fallstudienlösung vor, die Teil der Prüfungsleistung ist (vgl. Abschnitt 4.1). Anschließend werden die Studierenden in einer halbstündigen Gruppenprüfung zu den Lerninhalten des Themenblocks geprüft (vgl. Abschnitt 4.2).

### 3.6 Besprechung der Fallstudie

Der Themenblock wird mit der Besprechung der Fallstudie abgeschlossen. Um die Bandbreite möglicher Lösungsansätze aufzuzeigen und zu diskutieren, präsentieren ausgewählte Gruppen ihre Fallstudienlösung im Plenum. Bei Bedarf ergänzen die Lehrenden allgemeine Lösungshinweise und thematisieren ausgewählte Aspekte der Fallstudie nochmals im Detail. Außerdem werden die verbliebenen offenen Fragen beantwortet.

## 4 Prüfungsleistungen

Die Studierenden sollen in der Lehrveranstaltung Fachwissen und Methodenkompetenzen erwerben, um alleine und im Team tragfähige Lösungsansätze für offene Fragestellungen erarbeiten und diskutieren zu können. Um diese Kompetenzen vollständig zu prüfen, setzt sich die Prüfungsleistung aus folgenden Teilleistungen zusammen:

- Teilleistung Fallstudien (40 % der Prüfungsleistung),
- Teilleistung mündliche Prüfungen (20 % der Prüfungsleistung),
- Teilleistung Abschlussfallstudie (40 % der Prüfungsleistung).

Bei der Konzeption der Teilleistungen wird sichergestellt, dass diese die Gütekriterien einer Prüfung nach METZGER und NÜESCH (2004, S.6) erfüllen. Das Kriterium der *Validität* stellt sicher, dass die Prüfung das beurteilt, was sie zu beurteilen vorgibt. Hierzu zählt, dass die Prüfung die Lernziele repräsentativ abdeckt (Inhaltsvalidität) und dass mithilfe der Prüfungsaufgaben gültige Rückschlüsse auf den Ausprägungsgrad der im Lernziel beschriebenen Kompetenzen möglich sind (Konstruktvalidität). Das Kriterium der *Zuverlässigkeit* formuliert den Anspruch, dass die erbrachte Prüfungsleistung fehlerfrei erfasst wird. Hierzu zählt die sorgfältige Auswahl und Zusammensetzung der Prüfungsaufgaben zu einer Prüfung mittleren Schwierigkeitsgrads (Reliabilität). Außerdem darf das Verhalten der Prüfenden während der Prüfung keinen Einfluss auf das Prüfungsergebnis haben (Durchführungsobjektivität) und die Zuteilung der Punkte zu den Antworten muss unabhängig von den Prüfenden sein (Auswertungsobjektivität). Das Kriterium der

Ökonomie stellt sicher, dass der Nutzen einer Prüfung in vertretbarem Rahmen zu dessen Aufwand steht und das Kriterium der *Chancengleichheit* fordert die Gleichwertigkeit der Prüfungen aller Prüflinge.

## 4.1 Teilleistung Fallstudien

Die Teilleistung Fallstudien setzt sich aus fünf Einzelleistungen Fallstudie – eine je Themenblock – zusammen. Die Bewertung der Teilleistung Fallstudien entspricht dem arithmetischen Mittel der Bewertungen der besten vier der fünf erbrachten Einzelleistungen. Der Fokus der Einzelleistung Fallstudie liegt auf der Überprüfung der erworbenen Fach- und Methodenkompetenzen zur Erarbeitung tragfähiger Lösungsansätze für realitätsnahe offene Fragestellungen des jeweiligen Themenblocks im Team. Sie umfasst die in der Gruppe erarbeitete Fallstudienlösung und deren Präsentation im Kolloquium. Mit dieser Teilleistung wird sichergestellt, dass sich die Studierenden in der Vertiefungsphase intensiv mit der Anwendung des erworbenen Wissens auf offene Fragestellungen des jeweiligen Themenblocks auseinandersetzen.

### 4.1.1 Konzeption

Die Aufgabenstellungen der Fallstudien werden basierend auf den themenblock-spezifischen Lernzielen (Inhaltsvalidität) konzipiert. Hierbei ist darauf zu achten, dass die einzelnen Aufgaben der Fallstudie unterschiedliche Stufen der Lerntaxonomie nach Bloom abdecken (Reliabilität). Dies wird durch mehrere aufeinander aufbauende und sich hinsichtlich des Schwierigkeitsgrads steigernde Teilaufgaben erreicht. Weiterhin sind die Aufgaben so ausgewählt, dass über die verschiedenen didaktischen Bausteine eines Themenblocks hinweg alle themenblock-spezifischen Lernziele gleichmäßig abgedeckt werden (Konstruktvalidität). Bei der Erstellung der Fallstudie werden verschiedene Lösungsansätze für die einzelnen Aufgaben (Inhaltsvalidität) entwickelt und dokumentiert und die aufgabenspezifischen Punkteschemata festgelegt (Auswertungsobjektivität). Zur Bewertung der Einzelleistung Fallstudie werden mehrere prüfungsform-spezifische Bewertungsschemata eingesetzt (vgl. Abb. 3 und 4), um eine objektive Bewertung der erbrachten Leistungen sicherzustellen (Auswertungsobjektivität).

### 4.1.2 Bewertung

Die Bewertung der Einzelleistung Fallstudie setzt sich aus der Bewertung der *Fallstudienlösung* (75 % der Einzelleistung Fallstudie) und der *Präsentation der Fallstudienlösung* (25 % der Einzelleistung Fallstudie) zusammen.

In den Gruppenabgaben der Fallstudienlösungen werden die Korrektheit, Vollständigkeit und Nachvollziehbarkeit des dargestellten Lösungsansatzes bewertet. Da es sich bei den Aufgabenstellungen der Fallstudie um offene Fragestellungen handelt, liegt der Fokus der Bewertung auf dem gewählten Lösungsansatz und der kritischen Auseinandersetzung mit den gefundenen Ergebnissen. Im Einzelnen werden die begründete Auswahl eines geeigneten Modells, dessen konsistente Parametrierung und nachvollziehbare Anwendung auf die betrachtete Problemstellung, die Korrektheit der berechneten Ergebnisse sowie deren kritische Diskussion beurteilt. Zudem muss Fremdwissen nach wissenschaftlichen Standards gekennzeichnet sein. Für jede Gruppenabgabe werden der individuelle Lösungsweg sowie die individuelle Argumentation nachvollzogen, etwaige Folgefehler berücksichtigt und anteilig Punkte vergeben. Die Bewertung erfolgt basierend auf dem Bewertungsschema der Fallstudienlösung (vgl. Abb. 3) und den aufgabenspezifischen Punkteschemata. In der Präsentation der Fallstudienlösung werden Aufbau und Inhalt des Vortrags, Darstellung und Aufbereitung der Inhalte sowie der Vortragsstil bewertet. Außerdem ist darauf zu achten, dass sich die Studierenden die Vortragsinhalte gleichmäßig untereinander aufteilen und die vorgegebene Vortragsdauer einhalten. Das Bewertungsschema der Präsentation der Fallstudienlösung ist in Abb. 4 dargestellt. Jedes Bewertungskriterium wird mit einer Punktzahl zwischen 0 und 6 Punkten bewertet. Alle Bewertungsschemata werden den Studierenden transparent gemacht.

Bewertungskriterium	Beobachtungsanker		
	unzureichend	befriedigend	gut
<b>Auswahl eines geeigneten Modells / Methode</b>	- Ungeeignetes Modell ohne konsistente Begründung gewählt	- Ungeeignetes Modell ausgewählt, jedoch konsistent begründet - Geeignetes Modell gewählt, dies aber nicht begründet	- Es wird ein geeignetes Modell gewählt und dessen Wahl nachvollziehbar und schlüssig begründet
<b>Parametrisierung des Modells / der Methode</b>	- In der Aufgabenstellung gegebene Werte werden den Modellparametern falsch zugeordnet - Die Wahl nicht gegebener Parameter wird nicht oder nicht nachvollziehbar begründet oder ist inkonsistent zum betrachteten Anwendungsfall - In der Aufgabenstellung gegebene Informationen werden nicht genutzt und dafür neue Werte "erfunden" - Bei der Recherche nach fehlenden Daten und Parametern werden unseriöse oder nicht zugängliche Quellen verwendet	- Teilweise korrekte Parametrisierung des Modells - Teilweise nicht nachvollziehbare / nicht konsistente Begründung der Parameterwahl - Mindestens eine nicht seriöse oder nicht zugängliche Quelle genutzt	- Konsistente Parametrisierung des Modells aus den in der Aufgabenstellung gegebenen Werten - Die Wahl nicht explizit angegebener Parameter wird nachvollziehbar begründet und ist konsistent zum betrachteten Anwendungsfall - Es werden seriöse und allgemein zugängliche Quellen verwendet
<b>Anwendung des Modells / der Methode</b>	- Die Vorgehensweise zur Anwendung des Modells kann nicht nachvollzogen werden	- Die Vorgehensweise zur Anwendung des Modells kann nur teilweise oder schwer nachvollzogen werden	- Die Vorgehensweise zur Anwendung des Modells ist nachvollziehbar dargestellt bzw. dargelegt
<b>Ergebnisse</b>	- Wesentliche Kenngrößen sind falsch berechnet - Die Einheiten der berechneten Kenngrößen fehlen oder sind falsch	- Kenngrößen teilweise korrekt berechnet - Mindestens eine Einheit nicht korrekt berechnet / nicht angegeben	- Kenngrößen sind korrekt berechnet - Zu allen Kenngrößen sind konsistente Einheiten angegeben
<b>Diskussion / Interpretation der Ergebnisse</b>	- Die Ergebnisse werden nicht oder nicht nachvollziehbar diskutiert - Die Diskussion ist inkonsistent zu den Ergebnissen - Die aufgeführten Punkte widersprechen wissenschaftlichen Erkenntnissen oder physikalischen Gesetzen - Wesentliche Aspekte werden nicht aufgegriffen	- Die Ergebnisse werden nur einseitig diskutiert, wesentliche Erkenntnisse werden nicht aufgegriffen - Es werden weniger wichtige Aspekte zu ausführlich oder wichtige Aspekte nur oberflächlich diskutiert - Die Diskussion ist teilweise inkonsistent zu den Ergebnissen	- Die Ergebnisse werden nachvollziehbar diskutiert - Die relevanten Aspekte werden prägnant diskutiert - Die Diskussion ist konsistent zu den Ergebnissen - Es werden Aspekte aufgegriffen, die über den Inhalt der Aufgabenstellung hinausreichen
<b>Kennzeichnung fremden Gedankenguts</b>	- Fremdwissen wird ungekennzeichnet verwendet	- Fremdwissen ist nicht nach wissenschaftlichen Standards gekennzeichnet	- Fremdwissen ist nach wissenschaftlichen Standards gekennzeichnet

Abb. 3: Bewertungsschema der Fallstudienlösung

Bewertungskriterium	Beobachtungsanker		
	unzureichend (0 - 2 Punkte)	befriedigend (3 - 4 Punkte)	gut (5 - 6 Punkte)
<b>Aufteilung in der Gruppe</b>	<ul style="list-style-type: none"> <li>- Zeitlich ungleichmäßige Aufteilung unter Gruppenmitgliedern</li> <li>- Vortragende unterbrechen oder korrigieren sich gegenseitig, keine oder unschöne Übergänge zwischen den Vortragenden</li> <li>- Gesamtvortrag wirkt nicht aus einem Guss</li> </ul>	Nur 1-2 der Beobachtungsanker sind erfüllt	<ul style="list-style-type: none"> <li>- Zeitlich gleichmäßige Aufteilung unter den Gruppenmitgliedern</li> <li>- Vortragende treten als Team auf, gute Übergänge zwischen den Vortragenden</li> <li>- Gesamtvortrag wirkt wie aus einem Guss</li> </ul>
<b>Vortragsstil</b>	<ul style="list-style-type: none"> <li>- Vortragende sprechen undeutlich und in einer unangemessenen Lautstärke</li> <li>- Auftreten, Bewegung und Gestik der Vortragenden wirken unruhig oder unkonzentriert</li> <li>- Vortragende suchen keinen Blickkontakt zur Zuhörerschaft</li> <li>- Vortrag wird nicht frei gehalten</li> </ul>	Nur 2-3 der Beobachtungsanker sind erfüllt	<ul style="list-style-type: none"> <li>- Vortragende sprechen deutlich und in angemessener Lautstärke</li> <li>- Auftreten, Bewegungen und Gestik wirken ruhig und konzentriert</li> <li>- Vortragende suchen den Blickkontakt der Zuhörerschaft</li> <li>- Vortrag wird frei gehalten</li> </ul>
<b>Darstellung</b>	<ul style="list-style-type: none"> <li>- Struktur der Präsentationsfolien ist nicht erkennbar, nicht logisch und uneinheitlich</li> <li>- Gesprochenes Wort wird nicht durch sinnvolle Bilder, Grafiken und Tabellen ergänzt, abgebildete Bilder, Grafiken und Tabellen bleiben unerwähnt</li> <li>- Bilder, Grafiken und Tabellen sind schwer lesbar und haben eine geringe Qualität</li> <li>- Textmenge ist unangemessen</li> <li>- Schriftart, -größe und Farben behindern die Lesbarkeit</li> </ul>	Nur 3-4 der Beobachtungsanker sind erfüllt	<ul style="list-style-type: none"> <li>- Präsentationsfolien sind klar, logisch und einheitlich strukturiert</li> <li>- Bilder, Grafiken und Tabellen ergänzen das gesprochene Wort sinnvoll</li> <li>- Bilder, Grafiken und Tabellen sind lesbar und haben eine hohe Qualität</li> <li>- Textmenge ist angemessen</li> <li>- Schriftart, -größe und Farben erhöhen die Lesbarkeit</li> </ul>
<b>Aufbau und Inhalt</b>	<ul style="list-style-type: none"> <li>- Zentrale Punkte der Fallstudie bleiben unerwähnt, nebensächlichen Details wird zu viel Raum gewährt</li> <li>- Vortrag ist unstrukturiert, es ist schwer, dem Vortrag zu folgen</li> </ul>	<ul style="list-style-type: none"> <li>- Zentrale Punkte der Fallstudie werden teilweise angemessen dargestellt</li> <li>- Vortragsstruktur ist teilweise unklar</li> </ul>	<ul style="list-style-type: none"> <li>- Zentrale Punkte der Fallstudie werden nachvollziehbar beschrieben</li> <li>- Vortrag ist klar strukturiert und fördert dadurch das Verständnis</li> </ul>
<b>Zeiteinhaltung</b>	<ul style="list-style-type: none"> <li>- Vortragende unter-/ überschreiten das vorgegebene Zeitlimit deutlich</li> <li>- Vortragende referieren schnell und hektisch, aufgrund des Zeitdrucks bleiben wesentliche Inhalte unklar bzw. werden nicht dargestellt</li> </ul>	<ul style="list-style-type: none"> <li>- Vortragende unter-/ überschreiten das vorgegebene Zeitlimit leicht</li> <li>- Vortragende referieren ruhig, einige Inhalte bleiben aufgrund des Zeitdrucks unklar oder werden zu kurz dargestellt</li> </ul>	<ul style="list-style-type: none"> <li>- Vortragende nutzen die zur Verfügung stehende Zeit exakt</li> <li>- Vortragende referieren ruhig und haben ausreichend Zeit, um die Inhalte umfassend darstellen zu können</li> </ul>

Abb. 4: Bewertungsschema der Präsentation der Fallstudienlösung

## 4.2 Teilleistung mündliche Prüfungen

Die Teilleistung mündliche Prüfungen setzt sich aus fünf Einzelleistungen mündliche Prüfung – eine mündliche Prüfung je Themenblock – zusammen. Die Bewertung der Teilleistung mündliche Prüfungen entspricht dem arithmetischen Mittel der Bewertungen der besten vier der fünf erbrachten Einzelleistungen mündliche Prüfung. Die Einzelleistung mündliche Prüfung konzentriert sich auf die Prüfung des erworbenen Fachwissens. Diese Teilleistung stellt sicher, dass sich die Studierenden während der Lernphase jedes Themenblocks intensiv mit dessen Lerninhalten auseinandersetzen. Dies ist die zentrale Herausforderung des *Inverted Classroom*, da dessen Vertiefungsphase eine sorgfältige Vorbereitung der Lerninhalte in der Lernphase voraussetzt (HANDKE & SPERL, 2012, S. 49).

### 4.2.1 Konzeption

In der mündlichen Prüfung werden jedem Studierenden vier individuelle Fragen unterschiedlichen Schwierigkeitsgrads gestellt: Eine Frage bezieht sich auf das Verständnis der gruppenindividuellen Fallstudienlösung, um die Mitarbeit aller Studierenden bei deren Erarbeitung zu überprüfen. Zwei Fragen thematisieren die inhaltlichen Grundlagen des Themenblocks. Außerdem wird eine Transferfrage zu den Lerninhalten des Themenblocks gestellt.

Die Prüfungsfragen werden aus den themenblock-spezifischen Fragekatalogen ausgewählt, die neben den schriftlich ausformulierten Fragen auch deren Lösung sowie deren Einordnung in die Lerntaxonomie nach Bloom umfassen (Inhaltsvalidität). Die Fragen werden bereits vorab ausgewählt und gruppiert. So ist sichergestellt, dass jedem Studierenden die gleiche Anzahl an Fragen unterschiedlicher Stufen der Lerntaxonomie nach Bloom gestellt (Reliabilität) und bei jedem Studierenden mehrere Themenbereiche und Lernziele des Themenblocks abgedeckt werden (Inhaltsvalidität).

Bewertungskriterium	Beobachtungsanker		
	unzureichend (0 - 2 Punkte)	befriedigend (3 - 4 Punkte)	gut (5 - 6 Punkte)
<b>Verständnis des Lösungswegs</b>	- Lösung unvollständig oder inkorrekt erklärt - Rückfragen zur Lösung nicht beantwortet	- Lösung ohne Fokus auf die wesentlichen Schritte erklärt - Rückfragen teilweise oder ungenau beantwortet	- Nachvollziehbare Lösung mit Fokus auf die wesentlichen Schritte prägnant erklärt - Rückfragen präzise und vollständig beantwortet
<b>Grundlagen</b>	- Modelle, Formeln, Grafiken und Fachbegriffe nicht genannt - Wirkzusammenhänge nicht erläutert und nicht analysiert	- Modelle, Formeln, Grafiken und Fachbegriffe genannt und erklärt - Wirkzusammenhänge teilweise erläutert und analysiert	- Modelle, Formeln, Grafiken und Fachbegriffe präzise und vollständig genannt, erklärt und angewandt - Wirkzusammenhänge präzise und vollständig erläutert und analysiert
<b>Transfer</b>	- Keine Verbindung zwischen Transferfrage und Modell hergestellt	- Verbindung zwischen Transferfrage und Modell hergestellt	- Verbindung zwischen Transferfrage und Modell hergestellt und Auswirkungen beurteilt

Abb. 5: Bewertungsschema der mündlichen Prüfung

Die Studierenden dürfen die Prüfungsfragen nicht untereinander weitergeben. Die Studierenden werden in mehreren Fragerunden geprüft, wobei jedem Studierenden eine Prüfungsfrage je Fragerunde gestellt wird und die Reihenfolge der Studierenden in jeder Fragerunde variiert (Durchführungsobjektivität). Die Antworten der Studierenden werden protokolliert (Durchführungsobjektivität). Anhand des Protokolls und des Bewertungsschemas der mündlichen Prüfung beurteilen die Prüfenden die Leistung der Studierenden nach der mündlichen Prüfung gemeinsam (Auswertungsobjektivität).

#### 4.2.2 Bewertung

Es werden die Korrektheit, Vollständigkeit und Genauigkeit der Antworten der Studierenden auf deren individuelle Prüfungsfragen bewertet. Das Bewertungsschema der mündlichen Prüfung ist in Abb. 5 dargestellt. Jede Prüfungsfrage wird mit einer Punktzahl zwischen 0 und 6 Punkten bewertet.

### **4.3 Teilleistung Abschlussfallstudie**

Der Fokus der Teilleistung Abschlussfallstudie liegt auf der Überprüfung der erworbenen Fach- und Methodenkompetenzen zur selbstständigen Erarbeitung tragfähiger Lösungsansätze für realitätsnahe offene Fragestellungen aller Lerninhalte der Veranstaltung. Hierzu bearbeiten die Studierenden nach Abschluss der fünf Themenblöcke am Ende des Semesters eine Fallstudie – die sogenannte Abschlussfallstudie – in Einzelarbeit und unter Prüfungsbedingungen einer schriftlichen Klausur. Die Bearbeitungszeit der Abschlussfallstudie ist auf vier Stunden begrenzt. Die Studierenden dürfen jegliche nicht-elektronische Materialien, wie Lehrbücher, Übungsaufgaben und persönliche Aufschriebe nutzen.

Die Teilleistung Abschlussfallstudie wird nach denselben Prinzipien konzipiert wie die Fallstudien der Teilleistung Fallstudien (vgl. Abschnitt 4.1.1). Zusätzlich werden die übergeordneten Lernziele der Lehrveranstaltung berücksichtigt, um deren vollständige und gleichmäßige Prüfung sicherzustellen. Zur Bewertung von Korrektheit, Vollständigkeit und Nachvollziehbarkeit der abgegebenen Fallstudienlösung werden die Kriterien der Fallstudienlösungen der Teilleistung Fallstudien herangezogen (vgl. Abb. 3).

## **5 Evaluation des Lehrkonzepts**

Systematische Lehrevaluationen werden am KIT in allen Lehrveranstaltungen mit dem Ziel durchgeführt, die Qualität der Lehre sowohl mittels quantitativer als auch qualitativer Auswertungen zu beurteilen. Um die Erreichung der gesetzten Lernziele zu messen, wird der vom KIT eingesetzte Evaluationsbogen um die von RAUPACH et al. (2011) vorgeschlagene Methode der „vergleichenden Selbsteinschätzung“ der Studierenden erweitert. Hierbei beurteilen die Studierenden ihre erworbenen Kompetenzen hinsichtlich der definierten Lernziele selbst. Auf dem Evaluationsbogen haben die Studierenden weiterhin die Möglichkeit, in Freitextfeldern Feedback zur Durchführung der Lehrveranstaltung zu geben. Im Folgenden werden die Lehrevaluationen aus den Jahren 2018 (29 Teilnehmende) und 2019 (25 Teilnehmende) ausgewertet. Die Lehrevaluationen wurden jeweils am Ende des Semesters, jedoch vor der Abschlussfallstudie durchgeführt.

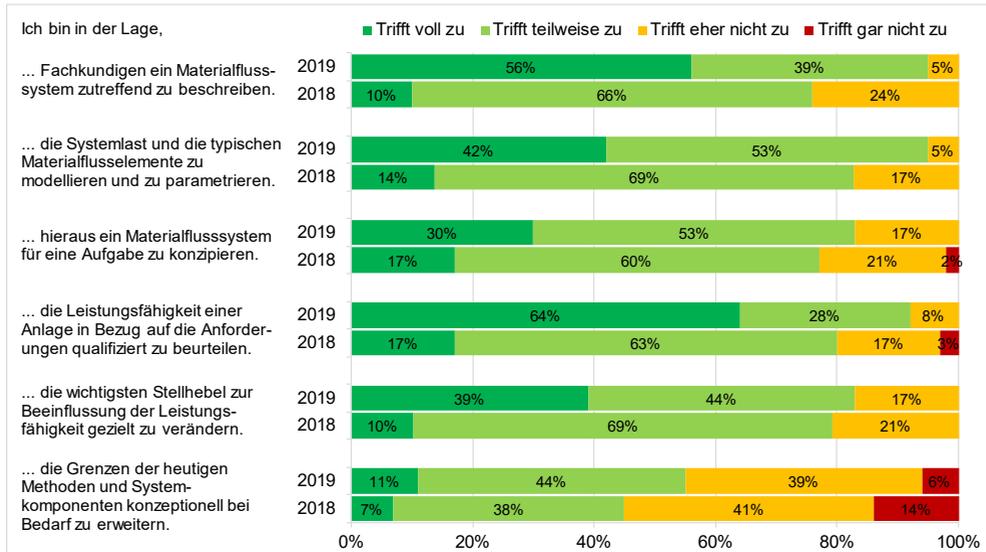


Abb. 6: Auswertung der Lehrevaluationen in der vergleichenden Selbsteinschätzung der Studierenden

Abb. 6 zeigt die Ergebnisse der vergleichenden Selbsteinschätzung in Bezug auf die Lernziele der Veranstaltung. Aus den Ergebnissen geht hervor, dass in beiden Jahren die Mehrheit der Befragten die Lernziele als erreicht ansieht. Der Anteil der Befragten, der der Erreichung der gesetzten Lernziele voll oder teilweise zustimmt, liegt bei fünf der sechs Lernziele deutlich über 70 Prozent. Das sechste Lernziel – die Grenzen der heutigen Methoden konzeptionell zu erweitern – sieht jeweils etwa die Hälfte der Befragten als voll oder teilweise erreicht an. Gleichzeitig ist hier der Anteil der Befragten am größten, die das Lernziel als nicht erreicht ansehen. Dies überrascht nicht, da dieses Lernziel in der Lerntaxonomie nach Bloom am weitesten oben angesiedelt ist. Die Zustimmungswerte sind im Jahr 2019 im Vergleich zu 2018 in allen Lernzielen teilweise deutlich angestiegen. Dies ist hauptsächlich auf einen Stabilisierungseffekt des Lehrkonzepts bei wiederholter Durchführung zurückzuführen.

Die Auswertung der Freitextantworten in den Evaluationsbögen unterstützt die These, dass das Lehrkonzept zur Kompetenzentwicklung geeignet ist. Die Mehrheit der Befragten gibt an, in der Veranstaltung viel gelernt zu haben. Einen großen Lernerfolg erreichen die Befragten durch das eigenständige Erfassen des Stoffes und die selbstständige Bearbeitung offener Fragestellungen. Als wesentliche Faktoren für den Lernerfolg werden die fachliche Diskussion und der Austausch mit Mitstudierenden genannt. Das legt die Schlussfolgerung nahe, dass die Bereitstellung gut vorbereiteter Arbeitsmaterialien sowie die enge fachliche Betreuung der Studierenden (sowohl während als auch außerhalb der Präsenzveranstaltungen) wesentlich für den Erfolg des Lehrkonzepts sind. Konkretes Verbesserungspotenzial sehen die Befragten bei der Gestaltung des Feedbacks zu den Fallstudienlösungen. Hier wird der Wunsch geäußert, im Feedback individueller auf den jeweils gewählten Lösungsansatz einzugehen und detaillierter zu erklären, weshalb die maximale Punktzahl nicht erreicht wurde. Als Kritik am Lehrkonzept wird der im Vergleich zu anderen Lehrveranstaltungen als deutlich überdurchschnittlich wahrgenommene Arbeitsaufwand während der Vorlesungszeit angegeben. Während dies von allen Befragten hervorgehoben wird, gibt etwa die Hälfte der Befragten an, dass auch der erzielte Lernerfolg deutlich über dem Durchschnitt anderer Lehrveranstaltungen liegt, und möchte die Lehrveranstaltung daher an Mitstudierende weiterempfehlen.

## 6 Fazit

Der Beitrag stellt ein Lehrkonzept vor, dessen Ziel es ist, dass Studierende der Ingenieurwissenschaften am KIT Kompetenzen für die effektive Planung und den effizienten Betrieb der zukünftigen Logistiksysteme erwerben. Den Kern des Lehrkonzepts bildet die selbstständige Bearbeitung offener Fragestellungen sowie die Prüfung der hierdurch erworbenen Kompetenzen. Die Lerninhalte werden in fünf aufeinander aufbauende Themenblöcke aufgeteilt, die sich jeweils aus einer Lern- und einer Vertiefungsphase zusammensetzen. In den Lernphasen erarbeiten sich die Studierenden die Lerninhalte selbstständig. Die Inhalte werden in den Präsenzveranstaltungen gemeinsam vertieft und die Studierenden so an die Bearbeitung offener Fragestellungen herangeführt. Jeder Themenblock wird mit der Bearbeitung einer Fallstudie abgeschlossen, in der die Studierenden in Gruppen tragfähige Lösungen für realitätsnahe offene Fragestellungen entwickeln und in einem anschließenden

Kolloquium gegen kritische Nachfragen verteidigen. Sowohl die Bearbeitung der Fallstudien als auch die Kolloquien sind Prüfungsleistungen, in denen die erworbenen Kompetenzen der Studierenden im Sinne der definierten Lernziele geprüft werden. Zum Abschluss der Lehrveranstaltung werden die erworbenen Kompetenzen in einer Abschlussfallstudie geprüft, in der offene Fragestellungen zu allen Lerninhalten der Veranstaltung in Einzelarbeit bearbeitet werden.

Der Lernerfolg der Studierenden ist mit den Lehrevaluationen aus zwei Jahrgängen belegt. Ein Großteil der befragten Studierenden gibt an, in der Veranstaltung viel zu lernen und sieht die definierten Lernziele als erfüllt an. Der Erfolg der Kompetenzorientierung des Lehrkonzepts – im Sinne der *employability* der Studierenden – zeigt sich in den überdurchschnittlichen Leistungen, die Studierende nach ihrer Teilnahme an der Lehrveranstaltung in ihren Abschlussarbeiten oder als wissenschaftliche Hilfskraft erbringen.

Das vorgestellte Lehrkonzept basiert auf vier Erfolgsfaktoren: Die Grundlage ist die Abstimmung der Lehr- und Lernmethoden sowie der kompetenzorientierten Prüfungen mit den angestrebten Lernzielen nach dem *Constructive Alignment*. Die Einarbeitung in die Themenblöcke während der Phase des Selbststudiums erfolgt nachweislich besser in Einzelarbeit und fördert die Selbstorganisation der Lernenden (SPANNAGEL & FREISLEBEN-TEUTSCHER, 2016). Die Gruppenarbeit in der Vertiefungsphase fördert die soziale Kompetenz, die nachweislich nicht nur einen starken Einfluss auf den Lernerfolg beim kooperativen Lernen hat (GREEN & GREEN, 2005, und COHEN, 1988), sondern auch Voraussetzung für erfolgreiches Lernen und Arbeiten in Gruppen ist (JURKOWSKI, 2010). Zuletzt ist der enge und kontinuierliche Austausch zwischen den Studierenden und den Lehrenden wesentlich, unter anderem durch die Möglichkeit zur partizipativen Mitgestaltung der Studierenden in der Präsenzveranstaltung und das regelmäßige Feedback zum Lernerfolg in den über das Semester verteilten kompetenzorientierten Prüfungen.

Vor dem Hintergrund, dass Kompetenzvermittlung in Studienfächern häufig nur gelingen kann, wenn es sich nicht um Einzelinitiativen handelt (SPANNAGEL & FREISLEBEN-TEUTSCHER, 2016, und HANDKE & SCHÄFER, 2012), möchten wir andere Lehrverantwortliche – vor allem in den Ingenieurwissenschaften – ermutigen, die Methoden der konventionellen Lehre zu überdenken und die eigenen Lehrveranstaltungen nach Vorbild des hier vorgestellten Lehrkonzepts umzugestalten.

## Danksagung

Wir bedanken uns bei Katharina Fleischer-Dörr, Benedikt Fuß und Zäzilia Seibold, die an der initialen Entwicklung des Lehrkonzepts maßgeblich beteiligt waren. Der Abteilung für Personalentwicklung und Berufliche Ausbildung des KIT danken wir für die fachkundige Begleitung bei der Weiterentwicklung des Lehrkonzepts.

## 7 Literaturverzeichnis

- Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, R., & Wittrock, M.** (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Arnold, D., & Furmans, K.** (2019). *Materialfluss in Logistiksystemen*. Berlin: Springer.
- Ausubel, D. P.** (1974). *Psychologie des Unterrichts*. Weinheim: Beltz.
- Barr, R. B., & Tagg, J.** (1995). From Teaching to Learning – A New Paradigm for Undergraduate Education. *Change*, 27(6), 12–26.
- Biggs, J., & Tang, C.** (2007). *Teaching for quality learning at university: what the student does*. New York: McGraw-Hill/Society for Research into Higher Education.
- Bloom, B., Engelhart, M., Furst, E., Hill, W., & Krathwohl, D.** (1972). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim: Beltz Studienbuch.
- Cohen, J.** (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York: Taylor & Francis.
- Green, N., & Green, K.** (2005). *Kooperatives Lernen im Klassenraum und im Kollegium. Das Trainingsbuch*. Seelze-Velber: Kallmeyer.
- Handke, J., & Schäfer, A. M.** (2012). *E-Learning, E-Teaching und E-Assessment in der Hochschullehre – Eine Anleitung*. München: R. Oldenbourg Verlag.
- Handke, J., & Sperl, A.** (2012). *Das Inverted Classroom Model: Begleitband zur ersten deutschen ICM-Konferenz*. München: Oldenbourg.
- Jurkowski, S.** (2010). *Soziale Kompetenz und Lernerfolg beim kooperativen Lernen*. Kassel: kassel university press.

**Lage, M. J., Platt, G. J., & Treglia, M.** (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The journal of economic education*, 31(1), 30–43.

**Metzger, C., & Nüesch, C.** (2004). *Fair prüfen: Ein Qualitätsleitfaden für Prüfende an Hochschulen*. St. Gallen: Institut für Wirtschaftspädagogik.

**Raupach, T., Münscher, C., Beissbarth, T., Burckhardt, G., & Pukrop, T.** (2011). Towards outcome-based programme evaluation: Using student comparative selfassessments to determine teaching effectiveness. *Medical Teacher*, 33(8), e446–e453.

**Spannagel, C.** (2013). *Flipped Classroom – Inverted Classroom*. <https://youtu.be/3ddbzXKfTTE>

**Spannagel, C., & Freisleben-Teutscher, C. F.** (2016). Inverted classroom meets Kompetenzorientierung. *Kompetenzorientiert Lehren und Prüfen. Basics – Modelle – Best-Practices. Tagungsband zum 5. Tag der Lehre an der FH St. Pölten am 20.10.2016* (S. 57–67). St. Pölten: Fachhochschule St. Pölten.

**Weinert, F.** (2014). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. *Leistungsmessungen in Schulen* (S. 17–31). Weinheim – Basel: Beltz Verlag.

**Wildt J., & Wildt, B.** (2011). Lernprozessorientiertes Prüfen im „Constructive Alignment“. Ein Beitrag zur Förderung der Qualität von Hochschulbildung durch eine Weiterentwicklung des Prüfungssystems. In *Neues Handbuch Hochschullehre* (S. 1–46). Berlin: Raabe.

## Autor\*in\*en



Uta MOHRING || Karlsruher Institut für Technologie (KIT),  
Institut für Fördertechnik und Logistiksysteme (IFL) ||  
Gotthard-Franz-Str. 8, D-76131 Karlsruhe

[mohring@kit.edu](mailto:mohring@kit.edu)



Christoph JACOBI || Karlsruher Institut für Technologie (KIT),  
Institut für Fördertechnik und Logistiksysteme (IFL) ||  
Gotthard-Franz-Str. 8, D-76131 Karlsruhe

[jacobi@kit.edu](mailto:jacobi@kit.edu)



Prof. Dr.-Ing. Kai FURMANS || Karlsruher Institut für Technolo-  
gie (KIT), Institut für Fördertechnik und Logistiksysteme (IFL) ||  
Gotthard-Franz-Str. 8, D-76131 Karlsruhe

[furmans@kit.edu](mailto:furmans@kit.edu)

Mirjam BRASSLER<sup>1</sup> (Hamburg)

# Die Prüfung von interdisziplinären Kompetenzen im Rahmen des Prinzips des Constructive Alignments

## Zusammenfassung

Interdisziplinäres Lernen hat das Potenzial, Studierenden ein ganzheitliches Verständnis von Wissenschaft zu vermitteln, ihre Kompetenzen für den Arbeitsmarkt zu steigern und sie zu befähigen, komplexe Probleme zu lösen. Aufgrund der unterschiedlichen Lehr- und Prüfungskulturen fällt es Lehrenden schwer, interdisziplinäre Lehre im Team-Teaching umzusetzen. Der Artikel geht der Frage nach, wie man mithilfe des Prinzips des Constructive Alignments erfolgreich interdisziplinäre Prüfungsformate gestalten kann. Dazu werden auf Basis des Prinzips interdisziplinäre Lernziele (Verständnis disziplinärer Perspektiven, Fähigkeiten der fachübergreifenden Integration der Inhalte und die Reflexion) und dazu im Einklang stehende Prüfungsmethoden (Prüfung der Tiefe der Einzeldisziplinen, dem Mehrwert durch Integration und der kritischen Reflexion) und Lehr-Lern-Methoden (Methoden zum Kennenlernen und Verstehen der Einzeldisziplinen, zum Zusammenarbeiten und Reflektieren) abgeleitet. Auf Basis der theoretischen Herleitung werden Gestaltungsempfehlungen für die Formulierung interdisziplinärer Lernziele und für die Auswahl geeigneter Prüfungsmethoden zur Erreichung der Lernziele abgeleitet, die Lehrende dabei unterstützen, gemeinsam erfolgreich interdisziplinäre Prüfungen zu planen und umzusetzen.

## Schlüsselwörter

Interdisziplinäres Prüfen, Interdisziplinäres Lernen, Interdisziplinäre Kompetenz, Constructive Alignment, Interdisziplinäres Team-Teaching

---

<sup>1</sup> E- Mail: [mirjam.brassler@uni-hamburg.de](mailto:mirjam.brassler@uni-hamburg.de)



## **Assessing interdisciplinary competencies by designing with constructive alignment**

### **Abstract**

Interdisciplinary competence is important for academia, employability, and sustainable development. Due to differences in teaching and assessment cultures and traditions, educators experience difficulties in interdisciplinary team-teaching. This paper investigates how educators can successfully plan and execute interdisciplinary assessments following the principle of constructive alignment. Based on this principle, we define interdisciplinary learning goals as the further development of interdisciplinary competence (understanding disciplinary perspectives, interdisciplinary skills of integration and reflection), aligned assessments (depth of insight, additional value of integration and reflection) and teaching methods (addressing understanding, teamwork, and critical reflection). Based on these theoretical concepts, we explain the implications and make recommendations for designing and implementing interdisciplinary teaching and assessment in interdisciplinary team-teaching.

### **Keywords**

interdisciplinary assessment, interdisciplinary learning, interdisciplinary competence, constructive alignment, interdisciplinary team-teaching

## **1 Einleitung**

Der Bildungsauftrag im Studium adressiert sowohl die Vorbereitung auf eine berufliche Tätigkeit als auch die Persönlichkeitsentwicklung in Auseinandersetzung mit Wissenschaft und die Befähigung zur Bearbeitung von gesellschaftlichen Kernproblemen. Interdisziplinäres Lernen adressiert alle drei Dimensionen des Bildungsauftrages: Interdisziplinäres Lernen bereitet Studierende auf eine interdisziplinäre Tätigkeit vor, die zunehmend auf dem Arbeitsmarkt nachgefragt wird (NANCAR-

ROW et al., 2013), erlaubt eine Persönlichkeitsentwicklung vor dem Hintergrund einer holistischen Auseinandersetzung mit Wissenschaft, die zunehmend interdisziplinär ist (VAN NOORDEN, 2015), und befähigt zur interdisziplinären Bearbeitung von gesellschaftlichen Schlüsselproblemen wie dem Klimawandel, der Korruption oder der Geschlechterungerechtigkeit (UN, 2015), deren Komplexität eine interdisziplinäre Herangehensweise erfordert (LEDFORD, 2015).

Trotz dieser vielen Chancen und Vorteile für die Studierenden stehen die Lehrenden in der interdisziplinären Lehre vor vielen Herausforderungen und Hindernissen. Unterschiedliche Lehr- und Prüfungskulturen führen bei Lehrenden im Team-Teaching zu Konflikten und behindern die erfolgreiche Gestaltung und Durchführung einer gemeinsamen Lehrveranstaltung und Prüfung (JENKINS & STONE, 2019; LINDVIG & ULRIKSEN, 2019; SCHARLAU & HUBER, 2019). Der Artikel geht der Frage nach, wie Lehrende im interdisziplinären Team-Teaching mithilfe des Prinzips des *Constructive Alignments* erfolgreich interdisziplinäre Prüfungsformate gestalten können, um die studentische Entwicklung der interdisziplinären Kompetenz zu prüfen. Dazu wird zunächst in einer theoretischen Herleitung das Prinzip des *Constructive Alignments* aus der pädagogischen Psychologie mit erziehungswissenschaftlichen Befunden zum interdisziplinären Lehren und Prüfen verknüpft und ein Dreiklang aus interdisziplinären Lernzielen, Lehr-Lern-Methoden und Prüfungsmethoden zur Weiterentwicklung der interdisziplinären Kompetenz entwickelt. Auf Basis der theoretischen Herleitung werden im Anschluss Gestaltungsempfehlungen zum interdisziplinären Prüfen im Team-Teaching abgeleitet und eine Prüfungsmethode wird als Praxisbeispiel vorgestellt.

Der Mehrwert des vorliegenden Beitrags liegt in der Implementierung des Prinzips des *Constructive Alignments* in der interdisziplinären Lehre, indem dieser Ansatz der Psychologie mit der pragmatisch-konstruktionistische Theorie zum interdisziplinären Lernen der Erziehungswissenschaft verknüpft wird. Diese Analyse ermöglicht eine erstmalige theoretisch fundierte Sammlung an Leitfragen zur Formulierung von interdisziplinären Lernzielen und im Einklang stehenden Prüfungskriterien über die Disziplingrenzen Lehrender hinweg.

## 2 Herausforderungen im interdisziplinären Lehren und Prüfen

Interdisziplinäres Lernen beschreibt einen Prozess, in dem „Lernende Informationen, Daten, Methoden, Werkzeuge, Perspektiven, Konzepte und/oder Theorien von zwei oder mehr Disziplinen integrieren, um Produkte zu erstellen, Phänomene zu erklären, oder Probleme zu lösen, in einer Art, die mit einer einzelnen Disziplin nicht möglich wäre“ (BOIX MANSILLA, 2010, S. 289; eigene Übersetzung). Die Aufgabe im interdisziplinären Lehren ist es folglich, einen Lehr-Lern-Raum zu schaffen, in dem Studierende ihr unterschiedliches disziplinäres Wissen integrieren können, um gemeinsame Ergebnisse zu erzielen. Dies kann z. B. zwischen Studierenden der Physik und Politikwissenschaften geschehen, die zum Thema „Nukleare Sicherheit“ ein Szenario-Papier für die UN erstellen. Ein weiteres Beispiel wären Studierende aus der Psychologie und der Betriebswirtschaftslehre, die gemeinsam eine Handlungsempfehlung für die organisationale Führung eines Betriebs zum Thema „Gesundheitsförderliche Arbeitsmaßnahmen“ entwerfen.

Um die Wissenschaftlichkeit der Einzeldisziplinen in der Lehre abzusichern, werden interdisziplinäre Lehr-Lern-Formate oft im Team-Teaching durchgeführt. Team-Teaching bedeutet, dass Lehrende gemeinsam Lehre planen, vorbereiten und durchführen. Dabei sind beide anwesend, haben wechselnde Rollen in Bezug auf Leitung und Unterstützung und tragen gemeinsam die Verantwortung, wobei die Aufgabenverteilung flexibel gestaltet werden kann (HALFHIDE, 2009). Kommen die Lehrenden aus unterschiedlichen Einzeldisziplinen, können sie sich zu einem interdisziplinären Team für die Lehre zusammenschließen. Die interdisziplinäre Begegnung kann Lehrenden ermöglichen, eine neue Perspektive auf den eigenen Forschungsgegenstand zu erlangen, neue Lehr-Lern-Methoden kennenzulernen und zu erproben (BRAßLER, 2020).

In der Durchführung eines interdisziplinären Team-Teachings ergeben sich durch die unterschiedlichen Fachkulturen viele Herausforderungen für die Lehrenden. Durch die unterschiedlichen Lehrtraditionen können Lehrende verschiedene Vorstellungen davon haben, was „gute“ Lehre ausmacht (JENKINS & STONE, 2019) und in welcher Form Wissenschaft in der Lehre stattfinden sollte (SCHARLAU & HUBER, 2019). Eine unterschiedliche Lehrkultur zeigt sich zusätzlich in der Beziehung zwischen Lehrenden und Studierenden und unterschiedlichen Fachsprachen, die

zu Missverständnissen führen können. Im interdisziplinären Team-Teaching können Lehrende Schwierigkeiten haben, geeignete interdisziplinäre Lehrgegenstände zu identifizieren, die thematisch in beiden Disziplinen verankert sind (BRAßLER, 2020). Zusätzlich können ihnen Kenntnisse und Erfahrungen in der Implementierung von geeigneten interdisziplinären Lehrmethoden fehlen (GOLDING, 2009). Die Aushandlung in der gemeinsamen Planung kann äußerst zeitaufwändig und konfliktreich sein, was von Lehrenden oft unterschätzt wird (WADE & STONE, 2010). Interdisziplinäre Lehre ändert die Rolle der Lehrenden, was in diesen ein Gefühl des Kontrollverlusts verursachen kann (LINDVIG & ULRIKSEN, 2019).

Eine besondere Herausforderung ist die gemeinsame Gestaltung und Durchführung von Prüfungen in interdisziplinären Lehrveranstaltungen. Die unterschiedlichen Fachkulturen schließen auch unterschiedliche Prüfungskulturen mit ein (LINDVIG & ULRIKSEN, 2019; SCHARLAU & HUBER, 2019). Dazu können unterschiedliche Erwartung und Voraussetzung von Basiswissen, andere Zielsetzung von Prüfungen, unterschiedliche Beziehungen von Theorie und Praxis sowie unterschiedliche (hierarchische) Beziehungen von Prüfenden und Prüfling zählen (SCHARLAU & HUBER, 2019). Hinzu kommen Unterschiede in den Benotungstraditionen, unterschiedliche Möglichkeiten der Prüfungsgestaltung auf Basis der jeweiligen Prüfungsordnungen sowie unterschiedliche tradierte Prüfungsmethoden (BRAßLER, 2020). All diese Unterschiede versuchen Lehrende im interdisziplinären Team-Teaching gemeinsam zu überwinden, um zu einem gemeinsamen Prüfungsformat zu gelangen. LINDVIG und ULRIKSEN (2019) kommen zu dem Schluss, dass Lehrende im interdisziplinären Lehren andere Prüfungsformate brauchen als im monodisziplinären Lehren.

## 3 Theoretische Herleitung des Constructive Alignments im interdisziplinären Lehren und Prüfen

### 3.1 Das Gestaltungsprinzip „Constructive Alignment“

Ein theoretischer Ansatz aus der pädagogischen Psychologie zur Planung und Durchführung von Lehrformaten ist das Prinzip des *Constructive Alignments* (BIGGS & TANG, 2011). Lehre erfüllt dieses Prinzip, wenn Lernziele kompetenzorientiert formuliert und die Lehr-Lern-Methoden und Prüfungsmethoden in Kohärenz mit diesen Lernzielen ausgewählt werden. Die Idee hinter diesem Prinzip ist, dass Studierende und Lehrende Lehre „von hinten“ denken. Sowohl die Studierenden als auch die Lehrenden fokussieren sich dabei auf die Erreichung der Lernziele.

Die Lehrende, der Lehrende oder das Lehrendenteam überlegt sich dazu in einem ersten Schritt, was die Studierenden im Anschluss an die Lehrveranstaltung können sollen. Dazu formuliert die Lehrende oder der Lehrende kompetenzorientierte Lernziele. Im Anschluss wird überlegt, welche Prüfungsmethoden geeignet sind, um das Erreichen der vorab definierten Lernziele festzustellen. Die Lehrperson wählt entsprechende Prüfungsmethoden aus und kommuniziert diese zusammen mit den beabsichtigten Lernzielen an die Studierenden. Die Idee dahinter ist, dass nun alle Beteiligten – die Lehrperson und die Studierenden – ihr Verhalten in den Lehr- und Lernaktivitäten auf das Ziel ausrichten. Die Lehrperson wählt Lehr-Lern-Methoden aus, die die Studierenden dabei unterstützen, die Lernziele zu erreichen. Die Studierenden möchten adäquat auf die Prüfung vorbereitet sein und nutzen die Lehr-Lern-Methoden bewusst zur eigenen Vorbereitung.

Das Prinzip des *Constructive Alignments* lässt sich auf die Gestaltung von interdisziplinärem Lehren und Prüfen im Team-Teaching übertragen. Im Sinne des Prinzips ist interdisziplinäres Lehren und Prüfen dann erfolgreich, wenn das Lehrendenteam (a) vorab gemeinsam kompetenzorientierte Lernziele in Bezug auf das interdisziplinäre Lernen formuliert, (b) diese in der gewählten interdisziplinären Prüfung abfragt, und (c) passende interdisziplinäre Lehr-Lern-Methoden auswählt, die die Entwicklung dieser Kompetenzen ermöglichen.

### **3.2 Die Weiterentwicklung der interdisziplinären Kompetenz als Lernziel**

Folgt man dem Kompetenzmodell von LATTUCA, KNIGHT und BERGOM (2013), setzt sich die interdisziplinäre Kompetenz aus drei Subfacetten zusammen. Zum einen beinhaltet sie das Verständnis von unterschiedlichem disziplin-basiertem Wissen und unterschiedlichen Methoden, Erwartungen und Grenzen. Zum anderen beschreibt sie interdisziplinäre Fähigkeiten wie das Anerkennen unterschiedlicher disziplinärer Perspektiven sowie die Verbindung, Integration und Synthese dieser Perspektiven, um Ansätze für eine Problemlösung zu entwickeln. Die dritte Facette betrifft das reflektierende Verhalten, in dem das eigene Denken und die eigenen Lösungsstrategien hinterfragt und gegebenenfalls angepasst werden.

### **3.3 Interdisziplinäre Prüfungsformate**

Die Wahl des Prüfungsformats wird im Sinne des Prinzips des *Constructive Alignments* im Einklang mit den vorab definierten Lernzielen getroffen. Dementsprechend sollte bei einer Prüfung im Rahmen des interdisziplinären Lernens die Weiterentwicklung der interdisziplinären Kompetenz innerhalb der drei Subfacetten (Kompetenzmodell nach LATTUCA et al., 2013) festgestellt werden können. Den interdisziplinären Lernzielen folgend, sollten demnach in der interdisziplinären Prüfung a) das Verständnis von unterschiedlichen disziplinären Wissensinhalten, Methoden, Erwartungen und Grenzen, b) die interdisziplinäre Fähigkeit, unterschiedliche disziplinäre Perspektiven zu bedenken und zu einer Problemlösung zu integrieren sowie c) die Fähigkeit, das eigene Verhalten reflektieren zu können, adressiert werden.

In der Diskussion über die Prüfung interdisziplinären Lernens haben sich in der Erziehungswissenschaft folgende drei Kriterien für die Feststellung des interdisziplinären Lernerfolgs durchgesetzt: (1) Tiefe in den Einzeldisziplinen, (2) Mehrwert durch Integration und (3) kritische Reflexion (BOIX MANSILLA & DURAISINGH, 2007). Diese drei Kriterien entsprechen den einzelnen Subfacetten der interdisziplinären Kompetenz. Die Tiefe in den Einzeldisziplinen prüft das Verständnis von unterschiedlichen disziplinären Wissensinhalten, Methoden, Erwartungen und Grenzen. Der Mehrwert durch Integration prüft die interdisziplinäre Fähigkeit, unterschiedliche disziplinäre Perspektiven zu bedenken und zu einer Problemlösung zu integrieren. Mithilfe der Betrachtung der kritischen Reflexion kann die Fähig-

keit, das eigene Verhalten zu reflektieren, in der Prüfung adressiert werden. Sind die interdisziplinären Lernziele sowie das interdisziplinäre Prüfungsformat gewählt, so können im Sinne des Prinzips des *Constructive Alignments* Lehr-Lern-Methoden identifiziert werden, die im Einklang mit den Lernzielen stehen.

### 3.4 Interdisziplinäre Lehr-Lern-Methoden

In der Frage, welche interdisziplinären Lehr-Lern-Methoden sich eignen, damit Studierende interdisziplinäre Kompetenzen entwickeln, kann die pragmatisch-konstruktionistische Theorie zum interdisziplinären Lernen einige Antworten liefern (BOIX MANSILLA, 2016). Die Theorie wurde in der Erziehungswissenschaft an der Universität Harvard entwickelt und zielt darauf ab, dass Lernende ein interdisziplinäres Verständnis bzw. eine interdisziplinäre Denkart gewinnen. Dazu sind entsprechend der Theorie vier Elemente notwendig: ein interdisziplinäres Ziel, eine disziplinäre Grundlage, Integration und ein kritischer Stand.

Gemäß dem *Constructive Alignment* sollten die gewählten Lehr-Lern-Methoden im Einklang mit den intendierten Lernzielen stehen. Die vier Elemente der pragmatisch-konstruktionistischen Theorie sind dabei kohärent mit den drei Subfacetten der interdisziplinären Kompetenz und können daher einander zugeordnet werden. So kann die theoretische Untermauerung des Lernprozesses (pragmatisch-konstruktionistische Theorie) mit der theoretischen Fundierung der interdisziplinären Kompetenz (Kompetenzmodell nach LATTUCA et al., 2013) verknüpft werden. Die einzelnen Elemente sprechen die einzelnen Facetten der interdisziplinären Kompetenz an: Die disziplinäre Grundlage adressiert das Verständnis in den Einzeldisziplinen, das interdisziplinäre Ziel und die Verhandlung der Integrationselemente umfasst das Verbinden, Integrieren und Synthetisieren der Disziplinen und der kritische Stand adressiert die Reflexion des eigenen Denkens und der eigenen Lösungsstrategien.

Folglich sollten Lehrende Lehr-Lern-Methoden einsetzen, die das fachübergreifende Kennenlernen und das gegenseitige Verstehen fördern (Disziplinäre Grundlage). Außerdem sollten Lehrende Lehr-Lern-Methoden wählen, die Studierenden ermöglicht fachübergreifend zielorientiert zusammenzuarbeiten (Interdisziplinäres Ziel und Integration). Zudem ist entscheidend, dass Lehrende Lehr-Lern-Methoden einsetzen, die Studierenden ermöglichen, den interdisziplinären Lern- und Arbeitsprozess zu reflektieren (Kritischer Stand).

### 3.5 Constructive Alignment im interdisziplinären Lehren und Prüfen

Aus der theoretischen Herleitung mithilfe des Prinzips des *Constructive Alignments* im interdisziplinären Lehren und Prüfen ergibt sich ein Dreiklang aus den intendierten interdisziplinären Lernzielen, den interdisziplinären Prüfungsmethoden und den interdisziplinären Lehr-Lern-Methoden (siehe Abb. 1).

Das intendierte Lernziel des Verständnisses disziplinärer Perspektiven steht im Einklang mit Lehr-Lern-Methoden, die eine disziplinäre Grundlage in Form des Kennenlernens und Verstehens der Einzelwissenschaften adressieren, und mit Prüfungsmethoden, die die Tiefe in den Einzelwissenschaften abfragen (dunkelgrau). Das intendierte Lernziel der Weiterentwicklung interdisziplinärer Fähigkeiten erfordert den Einsatz von interdisziplinären Lehr-Lern-Methoden, die das Zusammenarbeiten in Form einer gemeinsamen Zieldefinition und Integration fördern, sowie den Einsatz von interdisziplinären Prüfungsmethoden, die den Mehrwert durch Integration prüfen (grau).

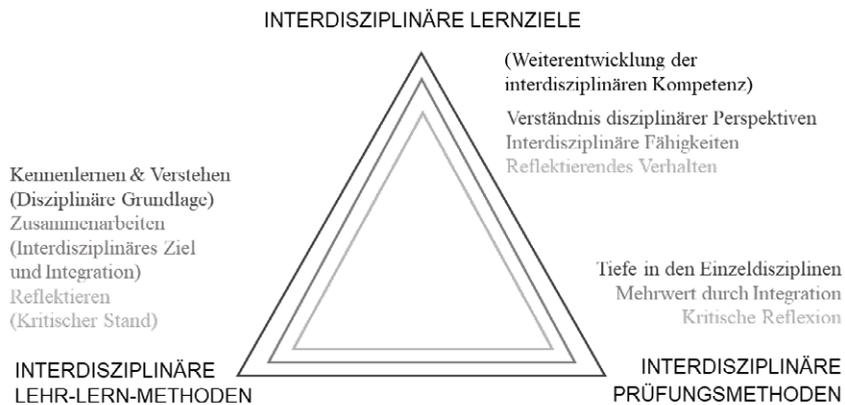


Abb. 1: Constructive Alignment im interdisziplinären Lehren und Prüfen (nach BRAßLER, 2020)

Das intendierte Lernziel der Weiterentwicklung reflektierenden Verhaltens steht im Einklang mit interdisziplinären Lehr-Lern-Methoden, die einen kritischen Stand in der individuellen Reflexion ermöglichen, und mit interdisziplinären Prüfungs-methoden, die eine kritische Reflexion abfragen (hellgrau).

## 4 Ableitung von Gestaltungsempfehlungen im interdisziplinären Team-Teaching

### 4.1 Gemeinsame Formulierung von interdisziplinären Lernzie-len

Folgt man dem *Constructive Alignment* im interdisziplinären Lehren und Prüfen können Lehrende in einem ersten Schritt interdisziplinäre Lernziele formulieren, die die Weiterentwicklung der interdisziplinären Kompetenz der Studierenden be-treffen. Dies ermöglicht Lehrenden im interdisziplinären Team-Teaching gemein-sam Lernziele zum Verständnis von unterschiedlichem disziplin-basiertem Wissen und unterschiedlichen Methoden, Erwartungen und Grenzen zu formulieren. Zum anderen können sie gemeinsam Lernziele zu interdisziplinären Fähigkeiten formu-lieren, die das Anerkennen unterschiedlicher disziplinärer Perspektiven sowie die Verbindung, Integration und Synthese dieser Perspektiven, um Ansätze für eine Problemlösung zu entwickeln, betreffen. Außerdem können sie gemeinsam Lern-ziele zum reflektierenden Verhalten formulieren, das Studierende im eigenen Den-ken und den eigenen Lösungsstrategien hinterfragen und gegebenenfalls anpassen. Dabei kann im Team-Teaching das Vorwissen der Studierenden in allen beteiligten Disziplinen sowie die Vorerfahrung im interdisziplinären Lernen und Arbeiten be-rücksichtigt werden.

Um im interdisziplinären Team-Teaching gemeinsame Lernziele zu formulieren, können Leitfragen, die man gemeinsam durchgeht, helfen. Zur Formulierung in-terdisziplinärer Lernziele in Bezug auf das *Verständnis disziplinärer Perspektiven* können im Lehrendenteam folgende Leitfragen nützlich sein:

- Bis zu welchem Grad sollen die Studierenden unterschiedliche Konzepte, Theorien, Befunde, Bilder, Daten, Methoden, Techniken, Instrumente, Prüfmetho-

den, Applikationen, Analogien, Diskurse und Spezialisierungen der jeweiligen Eigen- und Fremddisziplin kennen und verstehen können?

- Inwieweit sollen Studierende Inhalte ihrer eigenen Disziplin der jeweils anderen verständlich darstellen können?
- Inwieweit sollen Studierende Gemeinsamkeiten und Unterschiede zwischen den Disziplinen identifizieren und einordnen können?
- Inwieweit sollen Studierende fachfremde Fachliteratur lesen und in Grundzügen wiedergeben können?

Möchten Lehrende im interdisziplinären Team-Teaching gemeinsame Lernziele in Bezug auf die *interdisziplinären Fähigkeiten* formulieren, kann die gemeinsame Beantwortung folgender Leitfragen nützlich sein:

- Inwieweit sollen Studierende disziplinäre Inhalte, Perspektiven oder Erkenntnisse oder Methoden in Beziehung zur eigenen Disziplin setzen können?
- Inwieweit sollen Studierende im Team interdisziplinäre Lern- und Arbeitsziele und/oder Hypothesen formulieren können?
- Inwieweit sollen Studierende Gemeinsamkeiten und Unterschiede sowie Widersprüche zwischen den Disziplinen identifizieren, einordnen und Schlüsse daraus ziehen können?
- Inwieweit sollen Studierende zwischen den Disziplinen vermitteln und den Dialog koordinieren können?
- Inwieweit sollen Studierende die gemeinsame Arbeit über Disziplinen hinweg eigenständig koordinieren können?
- Inwieweit sollen Studierende Inhalte verschiedener Disziplinen mit einem Mehrwert integrieren können?
- Inwieweit sollen Studierende praktische interdisziplinäre Lösungsansätze entwickeln und umsetzen können?

Möchten Lehrende im interdisziplinären Team-Teaching gemeinsame Lernziele in Bezug auf das *reflektierende Verhalten* formulieren, kann die gemeinsame Beantwortung folgender Leitfragen nützlich sein:

- Inwieweit sollen Studierende die durch das interdisziplinäre Lernen erlangten Erkenntnisse über die eigene Disziplin (bspw. bezüglich ihrer Grenzen) benennen können?
- Inwieweit sollen Studierende die eigene persönliche Entwicklung und individuellen Lernfortschritte durch das interdisziplinäre Lernen und Arbeiten darstellen können?
- Inwieweit sollen Studierende die Inhalte, Methoden und Perspektiven der jeweils anderen Disziplin wertschätzen können?

## 4.2 Gemeinsame Prüfung der interdisziplinären Lernziele

Folgt man dem Constructive Alignment im interdisziplinären Lehren und Prüfen werden in einem zweiten Schritt geeignete interdisziplinäre Prüfungsmethoden gewählt, die im Einklang zu den Lernzielen stehen. Entsprechend den Erläuterungen weiter oben überlegen sich die Lehrenden im interdisziplinären Team-Teaching, wie sie die Tiefe des erreichten Wissens in den beteiligten Einzeldisziplinen, den Mehrwert der Integration der unterschiedlichen disziplinären Inhalte und die kritische Reflexion des interdisziplinären Lern- und Arbeitsprozesses überprüfen möchten. Um die Anreize für den interdisziplinären Lernprozess zu setzen, können die Lehrenden gemeinsame Kriterien zur Notenfindung wählen, die die Weiterentwicklung der interdisziplinären Kompetenz in ihren drei Subfacetten abbilden.

Ausgehend von den gemeinsam vorab formulierten interdisziplinären Lernzielen können sich die Lehrenden im Team-Teaching überlegen, welche Kriterien für ihre Prüfung stimmig sind.

Die *Tiefe des Wissens in den Einzeldisziplinen* bei den Studierenden kann anhand folgender Kriterien geprüft werden:

- Verständnis der Fremddisziplin (Inhalte, Perspektiven oder Erkenntnisse oder Methoden)
- Qualität und Quantität der genannten Literatur der jeweiligen Einzelwissenschaften
- Grad der Elaboration der Gemeinsamkeiten und Unterschiede der jeweiligen Einzelwissenschaften

Der *Mehrwert durch Integration* der unterschiedlichen disziplinären Inhalte kann anhand folgender Kriterien geprüft werden:

- Aufbereitung und Begründung interdisziplinärer Fragestellungen und Hypothesen
- Grad der Elaboration der Integration der Inhalte, Perspektiven, Ergebnisse und Gegenstände der Einzelwissenschaften
- Neuheit und Nützlichkeit der kreativen Lösungen und Ideen durch die fachübergreifende Integration

Die *kritische Reflexion* kann anhand folgender Kriterien geprüft werden:

- Grad der kritischen Auseinandersetzung mit der eigenen Arbeit bzw. dem interdisziplinären Ergebnis
- Grad der kritischen Auseinandersetzung mit den Grenzen in der eigenen Disziplin
- Verweise auf andere Disziplinen
- Reflexion des eigenen interdisziplinären Lernprozesses
- Reflexion der eigenen Lern- und Arbeitsstrategien

Wichtig ist, dass die Lehrenden ein Prüfungsformat wählen, das die Studierenden der unterschiedlichen Fachdisziplinen gleichermaßen anspricht und niemanden bevorteilt. In der Wahl einer Prüfungsmethode ist es entscheidend, die jeweiligen Prüfungskulturen, in denen die Studierenden sozialisiert wurden, zu berücksichtigen und zu integrieren. Im Anschluss können die Lehrenden interdisziplinäre Lehr-Lern-Methoden auswählen, die die Lernziele erfüllen und optimal auf die Prüfung vorbereiten.

### 4.3 Ein Praxisbeispiel: Interdisziplinäre mündliche Gruppenprüfung

In der Prüfungsmethode „Interdisziplinäre mündliche Gruppenprüfung“ (BRAßLER, 2022) lösen Studierende einen interdisziplinären, komplexen Problemfall im Team. Den Problemfall konzeptionieren die Lehrenden gemeinsam in der Schnittstelle zwischen den beteiligten Disziplinen. Optimalerweise haben die Lehrenden im Semester die Lehr-Lern-Methode „Interdisziplinäres Problembasiertes Lernen“ eingesetzt, das die Studierenden auf diese Prüfung vorbereitet.

Mithilfe der Prüfungsmethode können Lehrende die Erreichung folgender interdisziplinärer Lernziele überprüfen:

- *Verständnis disziplinärer Perspektiven:* Inhalte der Eigen- und Fremddisziplin lesen, erfassen und wiedergeben können, Gemeinsamkeiten und Unterschiede identifizieren können
- *Interdisziplinäre Fähigkeiten:* interdisziplinäre Lösungsansätze im Dialog generieren können
- *Reflektierendes Verhalten:* Grenzen der eigenen Disziplin aufdecken können

Die Prüfung gliedert sich in sechs Schritte, die auf zwei Prüfungsphasen aufgeteilt sind: eine individuelle Vorbereitungszeit und eine Diskussion im interdisziplinären Studierendenteam. Zunächst erhalten alle Studierenden des Teams die Problemstellung. (1) In der individuellen Vorbereitungszeit stehen den Studierenden 15 Minuten zur Verfügung, um sich die interdisziplinäre Problemstellung durchzulesen und bei Bedarf Notizen auf bereitgestelltem Papier zu machen. Danach beginnt die eigentliche mündliche Prüfung. (2) Zunächst definieren die Studierenden gemeinsam ihr interdisziplinäres Problem zwischen den beteiligten Disziplinen, um im Anschluss in einem (3) multidisziplinären Brainstorming verschiedene Lösungsansätze aus den verschiedenen Disziplinen zu diskutieren. Dabei bringen sie Literatur aus der Eigen- und Fremddisziplin an, die sie in der Lehrveranstaltung gelesen haben. Sie haben die Möglichkeit, ihre Ideen auf einer Pinnwand schriftlich festzuhalten. (4) In einem nächsten Schritt ordnen, verbinden und reflektieren sie ihre Ideen, indem sie Gemeinsamkeiten und Unterschiede zwischen den Disziplinen identifizieren. (5) Sie diskutieren die verschiedenen Ansätze und verknüpfen ihre Ideen. (6) Zum Abschluss formulieren sie ein integratives Team-Statement, das ihre interdisziplinäre Lösung für

die Problemstellung auf den Punkt bringt und gleichzeitig mögliche Grenzen aufzeigt.

- *Tiefe in den Einzeldisziplinen*: Qualität und Quantität der genannten Literatur; Pluspunkte für die Vorstellung der jeweils fachfremden Literatur
- *Mehrwert durch Integration*: Grad der Elaboration der Integration der Inhalte, Anzahl der Ideen
- *Kritische Reflexion*: Grad der kritischen Auseinandersetzung mit der Eigen- und Fremddisziplin, des eigenen interdisziplinären Ergebnisses und Verweise auf andere Disziplinen.

Es ist wichtig, dass die Lehrenden der unterschiedlichen Disziplinen während der Prüfung als Prüfende anwesend sind, da nur sie die fachliche Richtigkeit in ihrer Disziplin feststellen können.

Je nach Prüfungsordnung sind Gruppen- oder Einzelnoten denkbar. Zur Bestimmung der Einzelnote kann eine Bewertung der individuellen Leistung (anhand der Quantität der genannten Fachartikel und der Qualität der Elaboration der fachfremden Fachartikel) mit der Bewertung der Gruppenleistung (Integration und kritische Reflexion) kombiniert werden.

## 5 Diskussion

Interdisziplinäre Kompetenzen sind relevant für die Arbeit in Forschung und Praxis. Damit Studierende diese Kompetenzen entwickeln können, braucht es Lehrende, die interdisziplinäre Lehrveranstaltungen planen und umsetzen. Durch ihre unterschiedlichen Disziplinzugehörigkeiten und damit verbundenen unterschiedlichen Lehr- und Prüfungskulturen erleben Lehrende im interdisziplinären Team-Teaching vermehrt Konflikte (JENKINS & STONE, 2019; LINDVIG & ULRIKSEN, 2019; SCHARLAU & HUBER, 2019). Der vorliegende Artikel untersucht, wie Lehrende mithilfe des Prinzips des *Constructive Alignments* erfolgreich interdisziplinäre Prüfungsformate gestalten können, um die studentische Entwicklung der interdisziplinären Kompetenz zu prüfen. Definieren die Lehrenden vorab, inwieweit Studierende Wissen in ihrer eigenen Fachdisziplin und der Fremddisziplin aneignen sollen, wie die Integration gelingen kann und wie der Prozess und die gemeinsame Arbeit

reflektiert werden können, können Lehrende und Studierende ihr Lehr- und Lernverhalten darauf ausrichten.

Mit der gemeinsamen Definition von Lernzielen in Bezug auf die drei Facetten der interdisziplinären Kompetenz haben Lehrende im Team-Teaching eine gemeinsame Grundlage zu Beginn der gemeinsamen Arbeit. In der Diskussion um das intendierte Verständnis, Integration und Reflexion lernen die Lehrenden sich und ihre jeweiligen Disziplinen kennen. Dies kann interdisziplinäre Kommunikation stützen und die kulturübergreifende Begegnung stärken. Außerdem reflektieren die Lehrenden ihre eigene interdisziplinäre Kompetenz, indem sie feststellen, inwieweit sie ihre unterschiedlichen Perspektiven kennen, verstehen und integrieren und Grenzen der eigenen Disziplin erkennen können.

Die Festlegung auf gemeinsame Kriterien in der Prüfung, die die einzelnen Facetten der interdisziplinären Kompetenz adressieren, integrieren die Lehrenden ihre unterschiedlichen Perspektiven und schaffen so eigenen gemeinsamen Erwartungsraum (SCHARLAU & HUBER, 2019). Auf Basis der Kriterien haben die Lehrenden eine gemeinsame Zielsetzung der Prüfung. Die Adressierung der einzelnen Facetten schafft eine neue Ausrichtung der Prüfung für beide Lehrende und erlaubt somit auch eine neue gemeinsame Prüfungskultur und/oder Methode, die es zur Realisierung von Interdisziplinarität im Prüfungskontext braucht (LINDVIG & ULRIKSEN, 2019). Indem die Lehrenden in der Prüfung – wie im Praxisbeispiel – Möglichkeiten der Multidisziplinarität, also dem Nebeneinander der Inhalte der Disziplinen, und der Interdisziplinarität, als der Verknüpfung dieser Inhalte, sowie der Reflexion des Prozesses geben, ermöglichen sie eine Ausrichtung der Prüfung auf die Feststellung der Weiterentwicklung der interdisziplinären Kompetenz. Um die inhaltliche Richtigkeit und Genauigkeit in den Einzeldisziplinen sowie deren Grenzen festzustellen, braucht es die Anwesenheit und aktive Bewertung aller Lehrenden in der Durchführung der interdisziplinären Prüfung.

Dieses kleinschrittige und zielorientierte Vorgehen im *Constructive Alignment* kann interdisziplinären Konflikten präventiv entgegenwirken und Lehrenden ein Gefühl von Kontrolle geben (LINDVIG & ULRIKSEN, 2019). Nichtsdestotrotz bleiben große Unterschiede zwischen den Disziplinen, die in der gemeinsamen Lehre und Prüfung bedacht werden müssen. Das Prinzip des *Constructive Alignments* kann aber eine Erleichterung in der Aushandlung dieser Unterschiede schaffen.

## 6 Literaturverzeichnis

- Biggs, J., & Tang, C.** (2011). *Teaching for quality learning at university* (4. Aufl.). Buckingham: The Society for Research into Higher Education & Open University Press.
- Boix Mansilla, V.** (2010). Learning to synthesize: The development of interdisciplinary understanding. In R. Frodeman, J. T. Klein, C. Mitcham & J. B. Holbrook (Hrsg.), *Oxford handbook of interdisciplinarity* (S. 288–306). Oxford: Oxford University Press.
- Boix Mansilla, V.** (2016). Interdisciplinary learning. A cognitive-epistemological foundation. In R. Frodeman & J. Klein (Hrsg.), *Oxford handbook of interdisciplinarity* (2. Aufl., S. 261–275). Oxford, UK: Oxford University Press.
- Boix Mansilla, V., & Duraisingh, E. D.** (2007). Targeted assessment of students' interdisciplinary work: An empirically grounded framework proposed. *The Journal of Higher Education*, 78(2), 215–237.
- Braßler, M.** (2020). *Interdisziplinäres Lehren und Lernen – 50 Methoden für die Hochschullehre*. Weinheim: Beltz Juventa.
- Braßler, M.** (2022). Mündliche Gruppenprüfung im interdisziplinären Problem-basierten Lernen (iPBL). In J. Gerick, A. Sommer & G. Zimmermann (Hrsg.) *60 Prüfungsformate für die Hochschullehre. Kompetent Prüfungen gestalten* (S. 226–230). Münster: WAXMANN-Verlag.
- Golding, C.** (2009). *Integrating the disciplines: Successful interdisciplinary subjects*. Centre for the Study of Higher Education. Abgerufen von: [https://udayton.edu/el/aboutoel/\\_images/integrating-the-disciplines.pdf](https://udayton.edu/el/aboutoel/_images/integrating-the-disciplines.pdf)
- Halfhide, T.** (2009). Teamteaching. *Bündner Schulblatt*, 3, 4–8.
- Jenkins, N., & Stone, T. E.** (2019). Interdisciplinary responses to climate change in the university classroom. *Sustainability*, 12(2), 100–103. <https://www.liebertpub.com/doi/10.1089/sus.2018.0033>
- Lattuca, L. R., Knight, D. B., & Bergom, I. M.** (2013). Developing a measure of interdisciplinary competence. *International Journal of Engineering Education*, 29(3), 726–739.
- Ledford, H.** (2015). How to solve the world's biggest problems. *Nature*, 525(7569), 308–311.

**Lindvig, K., & Ulriksen, L.** (2019). Different, difficult and local: A review of interdisciplinary teaching activities. *The Review of Higher Education*, 43(2), 697–725. <https://doi.org/10.1353/rhe.2019.0115>

**Nancarrow, S. A., Booth, A., Ariss, S., Smith, T., Enderby, P., & Roots, A.** (2013). Ten principles of good interdisciplinary teamwork. *Human resources for health*, 11, 1–11.

**Scharlau, I., & Huber, L.** (2019). Welche Rolle spielen Fachkulturen heute? Bericht von einer Erkundungsstudie. *die hochschullehre*, 2019(5), 315–354.

**United Nations (UN)** (2015). *Transforming our world. The 2030 agenda for sustainable development*. Beschlüsse und Entscheidungen der United Nations Sustainable Development Summit 2015, 25.–27. September 2015, New York, US.

**van Noorden, R.** (2015). Interdisciplinary research by the numbers. *Nature*, 525(7569), 306–307.

**Wade, B. H., & Stone, J. H.** (2010). Overcoming disciplinary and institutional barriers: An interdisciplinary course in economic and sociological perspectives on health issues. *The Journal of Economic Education*, 41(1), 71–84. <https://doi.org/10.1080/00220480903382198>

## Autorin



Dr. Mirjam BRASSLER || Universität Hamburg, Institut für Psychologie || Von-Melle-Park 5, D-20146 Hamburg

<https://www.psy.uni-hamburg.de/arbeitsbereiche/arbeits-und-organisationspsychologie/team/brassler-mirjam.html>

[mirjam.brassler@uni-hamburg.de](mailto:mirjam.brassler@uni-hamburg.de)

Silke BRASELMANN<sup>1</sup> (Jena), Jolene MATHIESON (Leipzig) & Oliver MOISICH (Jena)

# Multimodal take-home exams in online teaching and beyond: constructive and professional alignment in teacher education

## Abstract

The authors propose an alternative to in-class exams (ICEs) based on the higher-order levels of Bloom's taxonomy as well as both constructive and professional alignment. In light of the COVID-19 pandemic, instructors were faced with restructuring both their teaching and assessment. The paper argues that take-home exams (THEs) tailored to the necessities of individual courses are advantageous to an online learning environment in comparison to an ICE and the pitfalls that accompany it, such as online proctoring, lower-level assessment, and the relative lack of constructive/professional alignment. In addition, THEs provide instructors with the possibility to utilize a variety of multimodal material that authentically reflect learner-centered assessment. The authors will continue to implement THEs post-pandemically as they allow for incorporating activities and strategies that are useful to students in their teaching careers.

## Keywords

multimodal take home exam, online teaching, constructive alignment, professional alignment, TEFL teacher education

---

1 email: [silke.braselmann@uni-jena.de](mailto:silke.braselmann@uni-jena.de)



# 1 Introduction

The COVID-19 pandemic has presented teachers in higher education with a number of challenges, including finding suitable assessment procedures in times of online teaching and learning. At the beginning of summer term 2020, and despite recent efforts to incorporate new classroom techniques and integrate multiple media as well as new forms of technology into university teaching practice, final exams were still a highly traditional, pen-and-paper format in some disciplines. As a result, many teachers in higher education began to experiment with online-proctoring strategies in order to avoid (or at least minimize) cheating in the new online environment by filming hands, tracking eye movements, using multiple hardware devices, and comparing handwritings. In addition to the considerable amount of time and effort required of both teachers and students within the online-proctoring regime, these procedures quickly led to heated debates about data protection and privacy in Germany (cf. HIMMELRATH, 2021; SCHWARTMANN, 2020). Despite these challenges, however, and out of sheer necessity, the accelerated digitalization of the educational system as result of the pandemic may have sparked the long-awaited “mini-revolution” (WILLIAMS & WONG, 2009, p. 227) in the university sector towards more innovative and competence-based forms of teaching and assessing. A growing body of current research aims to explore the potential of applying some of these changes and adjustments to post-pandemic teaching and learning (cf. KORDTS et al., 2021). This essay would like to contribute to this ongoing discussion by proposing the implementation of multimodal take-home exams (THEs) – in which students are provided with multimedia material and asked to complete a number of complex tasks over a prolonged time span – as an alternative assessment format that can be used in both online and classroom teaching, especially for teacher education.

This article draws upon our experiences restructuring and teaching the introductory seminar, “Teaching English as a Foreign Language (TEFL),” at the University of Jena for three consecutive semesters during the COVID-19 pandemic. In this specific course – as in many other traditionally assessed seminars – students would have normally written a proctored in-class exam (ICE) at the end of the semester, which, in large part, assesses student performance on the lower levels of Bloom’s taxonomy. Bloom’s well-known taxonomy provides a hierarchically organized description of learning that functions within two dimensions. The ‘knowledge dimension’ comprises of factual (e. g. knowledge of terminology), conceptual, procedural, and me-

ta-cognitive learning. The ‘cognitive process dimension’ focusses on the processes involved in learning. This second dimension is subdivided into the categories of remembering, understanding, applying, analyzing, evaluating, and creating. On the lowest level – remembering – students are asked to retrieve, recall, and recognize relevant knowledge (revised by ANDERSON et al., 2001). Being critical of online proctoring and wary of its invasiveness, and interested in assessing student performance on the higher levels of Bloom’s taxonomy, the authors of this paper, who each taught the same introductory course simultaneously in different groups, chose to redesign the final exam altogether. Based on the multiliteracies approach to teaching (KALANTZIS et al., 2016), we designed a multimodal take-home exam (THE) that focuses on both the knowledge and cognitive process dimensions.

In line with the prevailing theory of constructive alignment (CA) (cf. BIGGS, 2014) in higher education and in an effort to embody what we call *professional alignment* (PA) in teacher education, we adjusted the structure of the entire seminar and focused on fostering personal and professional competences and skills in the appropriate application of acquired knowledge. In the following, we discuss the benefits and challenges of THEs in comparison to traditional, proctored ICEs and argue for the underlying concepts of constructive and professional alignment in exam design. As a good-practice example with didactic insights, we will outline our design and assessment guidelines and further explore the potential that multimodal take-home exams bear for other disciplines and course formats.

## **2 Advantages and challenges of (multimodal) take-home exams**

In respect to exam design, *authenticity* lies at the center of the current paradigm of competence-based assessment (cf. HALBHERR et al., 2016, p. 252), which focuses on realistic, authentic examples and problems that students evaluate and solve in an exam. It therefore seems anachronistic that while computers and digital media are an integral part of both students’ and teachers’ personal and professional lives, modern technology still plays little to no role in examinations in many disciplines. In a study on the efficacy of final examinations, WILLIAMS and WONG (2009) ask whether

exams at university are, in fact, off-limits for innovation – and their investigation is still relevant today:

“[G]iven it is still the most commonly administered summative assessment instrument in universities today, is there some other special intrinsic value attached to a closed-book, invigilated exam that justifies its continued use?” (p. 228)

As we do not ascribe “special intrinsic value” to closed-book, pen-and-paper ICEs and further wish to incorporate the principle of authenticity into all aspects of our exam design, we designed a THE that includes digital and online material. THEs have been commonplace in the humanities at U.S.-American universities since the mid-1980s, and their advantages in comparison to ICEs have been widely discussed in anglophone research: ICEs are mainly characterized by a strict time limit (2–6 hours) and the stress that this necessarily and purposefully imposes on students. In traditional ICEs, learners are required to activate (factual) knowledge under an artificial amount of pressure, without help or opportunities for collaboration. While scholars in favor of ICEs and closed-book exams (CBEs) mostly stress the connection between expert performance and “rich, well-organized content knowledge of a subject” (DURNING et al., 2016, p. 1) as well as the ability of students to retrieve and activate knowledge under pressure, ICEs have been frequently criticized for several reasons. One important point of criticism is the artificially constructed, high amount of pressure inherent in the ICE setting that has been shown to have an adverse impact on students’ performance (cf. BENGTSSON, 2019). Further, higher level exam tasks which foster an active, creative, and lasting engagement with the seminar content and acquired knowledge are frequently omitted in ICEs due to time restraints and other pragmatic factors. As such, ICEs are “not suitable for assessing students’ performance on the higher levels of Bloom’s taxonomy scale,” nor can they promote anything but superficial learning, and thus are “not consonant with the prevailing theory of ‘constructive alignment’ in higher education” (IBID, p. 1).

Take-home exams are designed in opposition to ICEs: Due to a prolonged time span and unlimited access to textbooks, the Internet as well as other resources, they invite students to more thoroughly reflect on the tasks and materials at hand as well as make intelligent editorial and interpretive decisions regarding the applicability of specific sources and sets of knowledge gleaned from the wealth of material available to them. In their discussion of open-book and open-Internet exams (whether in the classroom or at home), WILLIAMS and WONG argue that when “learners

are presented with unstructured problems that require the application of relevant skills and knowledge, rather than selection from predetermined options as is the case with multiple-choice tests,” authentic and competence-based assessment can more accurately capture an “understanding of learning processes in terms of real-life performance as opposed to a display of inert knowledge” (2009, p. 229). Such exams, they conclude, “emphasize the importance of learner-directed discovery of knowledge” (IBID). In addition to these salient points, we further argue that the inclusive, complex, realistic environments provided by THEs allow for a substantially higher degree of not only constructive alignment but also professional alignment in teacher education, meaning that the exam itself is heuristically valuable. It functions as a “highly educational task in which students experience a deep learning process” (LÓPEZ et al., 2011, para. 3) as well as a highly professional assessment tool in which students and instructors alike experience a reflection process on exam design.

Studies have further shown that open-book and THE settings minimize student anxiety and that working with textbooks and other material can increase student confidence in their ability to successfully complete exam tasks (cf. JOHANNES et al., 2017). This is not surprising since such inclusive environments would more readily appeal to authentic and thus more tangible and less arbitrary learning and work strategies than ICEs. Students can work at a time of day and in a physical space that best suits their time management needs. What is more, they can analyze and apply both familiar as well as newly researched, supplementary material; and in cases where feasible, helpful, and reflective of professional praxis, they can cooperate with peers. Given these environmental advantages, THE exam questions and tasks can more robustly address all course material and “force students to higher level thinking, to apply knowledge to novel situations and synthesize material” (BENGTSSON, 2019, p. 9). In regard to assessing learning objectives, THE task structures can result in a more comprehensive and accurate testing of the targeted, multi-faceted learning objectives customarily engineered into a syllabus.

While there are many advantages to implementing THEs as an assessment format, they have also been contested. One main caveat is the issue of cheating: Recent studies support the general concern among university teachers that dishonesty, cheating, and unwanted cooperation increase outside the university classroom (STEFENSEN & SCHUSTER, 2020, p. 604–610). ICEs are thus mostly used to avoid unethical student behavior (i. e. cheating and unwanted cooperation). Yet, during

the first COVID-19 semester, the need to create online-exams that offered the same degree of proctoring led to time-consuming and often ineffective or legally problematic attempts at online-proctoring. However, even before the digital semesters of the pandemic, ICEs by no means guaranteed a cheating-free exam setting – on the contrary, as many so-called ‘cheating-scandals’ have shown (cf. PÉREZ-PENÁ, 2013). While it has been pointed out that a “non-proctored exam conducted in a closed dorm room with an Internet access is the perfect setup for frame-ups and imposture” (BENGTSSON, 2019, p. 11), it is important to note that THEs often do not specifically prohibit or sanction cooperation between students. Instead, we take the position that because THE task structures assess student performance on the higher levels of the learning taxonomy – especially understanding, applying, and creating – cooperation does not necessarily compromise their final performance but may instead enhance their output and ability to activate and apply factual knowledge in dialogical, cooperative learning (cf. also JOHNSON et al., 2015). As THEs test higher levels of Bloom’s taxonomy and consist of open-ended questions that require essay-style answers designed to foster higher-order cognitive skills (HOCS), the possibilities for unethical student behavior are minimized. While we condemn unethical phenomena such as pens-for-hire and plagiarism, we want to challenge the notion of ‘unethical student behavior’ – especially in teacher education. We would rather promote a culture of trust, cooperation, and reflection as these values align constructively and professionally in our field.

This, however, leads to the next point of criticism – the format’s practicability. Considering the time needed for designing THEs such as choosing multimodal material and formulating comprehensive reflection questions as well as for grading answers in prose, they may be better suited for smaller group sizes and for certain disciplines. However, as this article argues, due to their commitment to competence-based, authentic, and aligned assessment, THEs appear to be more rewarding for both teachers and students. As we found during three semesters of assessing our courses with THEs, this is also due to the guiding principles of constructive alignment for higher education and professional alignment for teacher education. The prevalent theory of constructive alignment (CA) is an outcome-based approach in which teaching is designed to help students reach the pre-defined outcome and assessment is designed to “enable clear judgements as to how well those outcomes have been attained” (BIGGS, 2014, p. 5) – not only are teaching and assessment aligned, but both are also aligned to the intended learning outcomes (IBID, p. 9). For the design and imple-

mentation of a THE, this implies that the learning outcome of the course is defined and openly communicated at the beginning of the semester.

Constructive alignment also goes hand in hand with what we came to call ‘professional alignment’ in teacher education. For teacher education, this means an alignment of the assessment with professional skills such as ethical and personal responsibility, self-reflection, critical thinking, information and time management, written communication, and appropriate use of different media. Guiding principles for professional alignment of outcome, teaching, and assessment are proposed as follows:

1. Establishing a culture of trust.
2. Taking one’s own teaching into account within the assessment regime, i. e. testing what has been taught and preparing students for the exam throughout the entire course.
3. Allowing for differentiation and designing an exam that is suitable for most learning strategies and learner types.
4. Applying a pedagogy of multiliteracies (KALANTZIS et al., 2016) in order to foster a deep engagement with multiple types of media and focusing on creative application, critical analysis, experiencing of old and new media and content, and conceptualization by theorizing and naming.
5. Creating tasks that authentically prepare the students for their profession (e. g. lesson planning, developing multiliteracies, researching own material).

## 3 The design and assessment of our THE

### 3.1 Exam design

As our THE is meant to function as a learner-centered assessment that constructively and professionally aligns with our course content, the test itself also shapes the seminar syllabus (cf. also RICH et al., 2014). Thus, designing the THE also entails designing the course. Our specific course was an introductory seminar that is a first-level prerequisite for all TEFL students of education, entitled “Teaching English as a Foreign Language (TEFL).” The course content and final examination are based entirely on the textbook *Teaching English* (GRIMM et al., 2015) and the assignments were specifically designed around an incremental series of tasks to hone the skills needed for the final examination.

The assignments function as formative summative assessments (FSA), which measure student progress throughout the semester in order to modify instruction and to enhance student performance by reviewing student work in class, producing “both quantitative and qualitative feedback from the students about their comprehension” (WININGER, 2005, p. 164). In our seminars, the FSAs take the form of tailor-made worksheets on a chapter-by-chapter and week-by-week basis that students must complete before commencing the week’s seminar session. The worksheets consist of closed, semi-open, and open task types, ranging from multiple-choice questions to gap filling to project work such as creating lesson plans based on differing media types including literature, film, music, GIFS, cartoons, etc. The worksheet tasks aim to target 1) knowledge of terminology and concepts and their creative application, 2) critical analysis of methodologies, 3) researching and analyzing both traditional and new media and content, and 4) inferencing by theorizing lesson objectives and tasks, and specifically relating practical examples to their underlying theories on (second) language acquisition and didactics. The seminar sessions consist of 13 weeks of online meetings where students, either in group work or in teacher-student interaction, present solutions to the worksheet tasks and optimize their responses and lesson plans based on peer and instructor feedback. A master worksheet is then collocated for that week’s chapter and uploaded in a shared digital forum. By the end of the semester, all worksheets including the lesson plans produced in project work are avail-

able to the students during the THE. The final exam takes place in the final week of the semester and students are given a time span of 48 hours to complete the THE.

The final examination entails two main sections – ‘Application’ and ‘Reflection’ – whereby ‘Application’ is further divided into two parts which comprise of open tasks that operate on the higher levels of Bloom’s learning taxonomy by requiring comprehension, application, and creative production in order to be completed: 1) designing a lesson plan based on specific media prompts provided by the instructors and according to the PWP (pre-, while-, and post-activities) method and lesson-plan template taught in class, and 2) elucidating in an open-essay format their theorizations and arguments on the efficacy of their chosen design and how it contributes to their defined learning objectives. Students are asked to use the specific terminology, theoretical concepts, and factual knowledge they acquired in their course work reading through, working with, and internalizing the contents of GRIMM et al. (2015). This section is especially attractive in terms of design because the material chosen by the instructors can provide a higher degree of multimodality than traditional ICEs would allow. In our exam, we provide students with six specific teaching materials, mostly in the form of URLs, from multiple media types, of which they can choose one to create their lesson plan, but they are also sanctioned to include any relevant supplementary media/materials based on their own research. Designing *multimodal* THEs, as we suggest here, that includes different kinds of material ranging from YouTube videos, GIFs, and TikToks to cartoons and poems, enables students to reflect on their material preferences (e. g. audio-visual, textual, etc.), their own learning styles, and their strategies for approaching their material. The THE thus also aims to produce the confident and transparent application of online sources as learning material. In the collection and choice of material provided for and by the students, THEs adhere to a multiliteracies pedagogy that enables students to actively and creatively engage with multiple text types (KALANTZIS et al., 2016). Further, the provided material can easily be updated and rotated on a semester basis with little time and effort.

The ‘Reflection’ section comprises of a semi-open, professionally-aligned task where students are asked to write a short essay on one of two questions that tests 1) factual knowledge of a specific concept, theory, or media type, 2) their ability to reflect on and create an informed opinion based on knowledge, and 3) their ability to illustrate

how didactic theory informs practical solutions and vice versa. The questions are each accompanied by an illustrative cartoon that students can use and build upon.

In designing our THE, we include the more professionally-aligned tasks that are too time-consuming in an ICE. The flexibility of this design further allows for a greater variety of topical, interesting, and multimodal material that speaks to the students and their realities, therefore engaging and motivating students to a higher degree. By doing so, THEs create a rich and multimodal learning experience that requires and encourages reflection on students' processes of learning and adheres to the principles of the post-method pedagogy.

### **3.2 Student and instructor assessment of exam design**

Course participants provided feedback to the instructors directly after they had met the deadline for submission. This feedback proves invaluable for the adjustment of future THE designs and will be the focus of the following section. In addition, we present some self-reflective remarks on the experiences surrounding THE from an instructor's perspective.

While students were working on the exam's assignments, they were able to pose questions to the instructor. Depending on the instructor, students could write an email with questions, or post questions on a server that was specifically installed for this course. The latter in particular mimicked situations during ICEs when students would ask questions to the instructor, as they were able to either ask a question in a direct message (visible only to the instructor) or ask in chat (visible to everyone). Instructors were thus able to react to questions whose answers may be beneficial for all participants in case they clear up uncertainties with the exam that were shared by more than one student. Likewise, instructors confirmed the timely submission of students' papers via email or text chat, which in turn reassured students of their submission status. We observed that the lack of a submission on paper was a potential cause for anxiety on the students' side, so confirmation via email or text seems beneficial.

Students evaluated that they found the assignments' goals self-evident and feasible within the provided time window. However, many also claimed to be short on time; almost all handed in their digital papers shortly before the deadline and individual

students remarked that they needed almost all of the allocated time to work on the tasks. We observed the same issue from some papers' lengths – while most submissions were in the range of five to six pages, some students came close to handing in a paper that is the equivalent to a term paper in terms of length (ten pages). As such, the pressure in ICEs as observed by BENGTSSON (2019, p. 1) appears to still be present in THEs. It remains inconclusive whether this pressure is due to the inherent design of THEs or reliant on student expectations inherent to a course's final exam, regardless of its design. As regards the preparation of classes for a THE, instructors should therefore not only take care of preparing their courses' content in accordance with the requirements of a THE, but also prepare their students on what to expect and how to approach the allocated time for such an exam. Students may appreciate not only a hard deadline that tells them when they ought to submit their paper, but also an estimation of how long they should actually work on the assignments within that time frame.

Feedback for the material offered in the exam was universally positive. When asked, students assessed that they found the material in the exam to be of great variety in terms of both topic and medium. In particular, one assignment that revolved around designing a TEFL lesson plan with material provided in the exam let students choose between various media, such as: GIFs with grammar puns to design a grammar lesson; a video speech at the DNC by Michelle Obama; a short clip from the US television show *The Simpsons* that commented on the January 6, 2021 insurrection at the United States Capitol. The last example especially proved to be popular since the exam took place mere weeks after the incident. Implementation of these media, among many others, shows the flexibility of a THE in comparison to traditional ICEs, where this high-level application of multiliteracies is virtually impossible.

Opponents of THEs may argue that the format makes examination easier for students. After all, an open-book exam with more time to work on assignments is bound to give students advantages that they would otherwise lack in a proctored, in-class examination. We found the opposite to be the case; in fact, the average grade in our THEs is close to averages of former ICEs for the same module. Figure 1 shows these averages for summer and winter term 2019 (in-class) and summer term 2020 to summer term 2021 (take-home) on a 1-to-5 scale, where 1.0 is the best grade achievable and 5.0 is a failed exam. The averages show that there is no significant difference between ICEs and THEs in terms of student performance. We hypothesize that the

slightly lower average performance in summer 2020 may be due to the relative novelty of a THE to students or due to mostly individual circumstances surrounding the emerging pandemic that may have affected students. The results after summer 2020 remain stable around the 2.2 to 2.3 range.

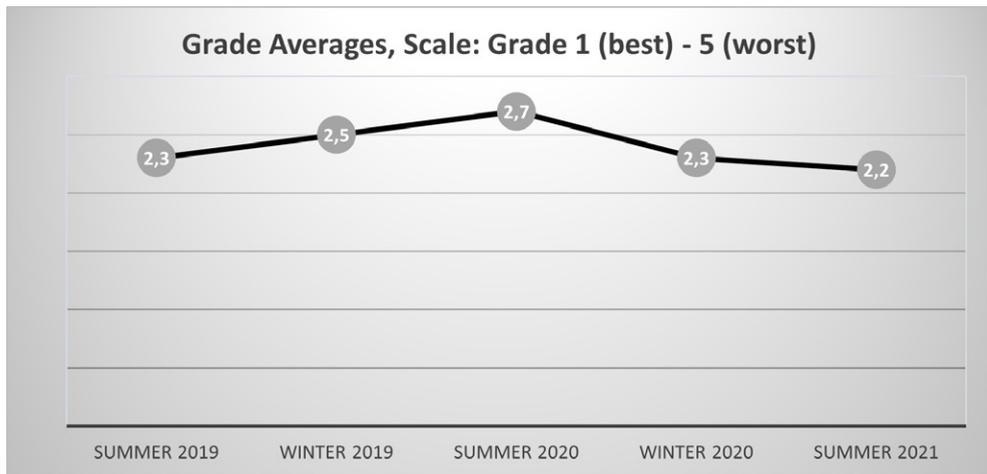


Fig. 1: Grade averages for course exams (illustration by the authors)

## 4 Conclusion

We mentioned above that THEs activate higher levels of Bloom’s taxonomy and, in addition, foster professional alignment for teacher education, by which we imply that the exam presents teacher degree students with a best-practice example of how to assess a course’s performance. This meta-didactic strategy proved successful in our TEFL course. However, such a conclusion does not necessarily mean that THEs are a viable solution for all kinds of courses in all kinds of disciplines. Modules with output-oriented curricula and open-designed tasks will benefit more from THEs than modules with closed-designed tasks. As an example, exams that ask students to prove a mathematical theory will not benefit from a THE since, traditionally, mathematical proofs by necessity allow only for a very limited number of ways to arrive at the correct conclusion. With that in mind, we propose that our exam design is quite flexible across didactic seminars of all disciplines – designing lesson plans and discussing didactic topics with an informed background (based on seminar input) are open to some degree in terms of students’ creativity. As such, the well-known teacher idiom, “teaching to the test,” gains a new meaning in light of THEs: We still teach to the test; however, due to professional alignment, the test now incorporates useful didactic heuristics and models that will be useful to teacher degree students in their careers.

Our experience with THEs also reveals a constructive alternative to online proctoring. While traditional exam formats require a high degree of effort from both teachers and students as well as constant online monitoring during the exam, which raise questions of efficiency, data protection, and ethics, a THE renders such issues a moot point by design. We have seen that students arrive at the same conclusion. While the concept of a THE arose out of necessity in a time of digital solutions where in-class courses were simply impossible, we have found the practicability of a take-home exam to be far more valuable than being an emergency alternative to in-class exams. As we prepared the first in-class seminars for the same module in almost two years, we decided to retain the exam’s THE format. While a take-home exam at first seemed uncalled for in an in-class structure, we determined that the advantages of a THE exam outweigh the traditional ICE exam even with an in-class course preceding examination. In conclusion, we are determined to refine the THE design further by implementing it into in-class courses.

## 5 References

- Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., & Wittrock, M.** (2001). *A Taxonomy for Teaching, Learning, and Assessing: A Revision of Blooms Taxonomy of Educational Objectives*. New York: Longman.
- Bengtsson, L.** (2019). Take-home exams in higher education: A systematic review. *Education Sciences*, 9(4), 1–16. [https://www.researchgate.net/publication/337093684\\_Take-Home\\_Exams\\_in\\_Higher\\_Education\\_A\\_Systematic\\_Review](https://www.researchgate.net/publication/337093684_Take-Home_Exams_in_Higher_Education_A_Systematic_Review)
- Biggs, J.** (2014). Constructive alignment in university teaching. *HERDSA: Review of Higher Education*, 1, 1–22. <https://www.herdsa.org.au/herdsa-review-higher-education-vol-1/5-22>
- Durning, S., Dong T., Ratcliffe, T., Schuwirth, L., Artino, A., Boulet, J. R., & Eva, K.** (2016). Comparing open-book and closed-book examinations: A systematic review. *Academic Medicine*, 91(4), 583–599. [https://journals.lww.com/academicmedicine/Fulltext/2016/04000/Comparing\\_Open\\_Book\\_and\\_Closed\\_Book\\_Examinations\\_.37.aspx](https://journals.lww.com/academicmedicine/Fulltext/2016/04000/Comparing_Open_Book_and_Closed_Book_Examinations_.37.aspx)
- Grimm, N., Meyer, M., & Volkmann, L.** (2015). *Teaching English*. Tübingen: Narr Verlag.
- Halbherr, T., Dittmann-Domenichini, N., Piendl T., & Schlienger, C.** (2016). Authentische, kompetenzorientierte Online-Prüfungen an der ETH Zürich. *Zeitschrift für Hochschulentwicklung*, 11(2), 247–269. <https://zfhe.at/index.php/zfhe/article/view/890>
- Himmelrath, R.** (2021). Überwachte Online-Klausuren: Ist volle Kontrolle möglich?. *Frankfurter Allgemeine Zeitung*. <https://www.faz.net/aktuell/karriere-hochschule/hoersaal/duerfen-studenten-bei-online-pruefungen-per-video-ueberwacht-werden-17234467.html>
- Johannes, B., Dinkens, A., & Moore, J.** (2017) A systematic review comparing open-book and closed-book examinations: Evaluating effects on development of critical thinking skills. *Nurse Education in Practice*, 27, 89–94. <https://www.sciencedirect.com/science/article/abs/pii/S1471595317305486?via%3Dihub>
- Johnson, C. M., Green, K. A., Galbraith, B. J., & Anelli, C. M.** (2015). Assessing and Refining Group Take-Home Exams as Authentic, Effective Learning Experiences. *Journal of College Science Teaching*, 44(5), 61–71.

**Kalantzis, M., Cope, B., Chan, E., & Dalley-Trim, L.** (2016). *Literacies*. 2<sup>nd</sup> ed. Cambridge: Cambridge University Press.

**Kordts, R., Wagner, R., Sidler, C., Tinsner-Fuchs, K., Dilger, B., & Brahm, T.** (2021). Cultivating a culture of experimentation in higher-education teaching and learning: Evaluation of recent experiences and transfer to the new-normal. *Zeitschrift für Hochschulentwicklung*, 16(3). <https://zfhe.at/index.php/zfhe/issue/view/71>

**López, D., Sánchez-Carracedo, F., Cruz, J., & Fernández, A.** (2011). A take-home exam to assess professional skills. *Proceedings – Frontiers in Education Conference 2011*.

**Pérez-Peña, R.** (2013). Students Disciplined in Harvard Scandal. *The New York Times*. <https://www.nytimes.com/2013/02/02/education/harvard-forced-dozens-to-leave-in-cheating-scandal.html>

**Rich, J. D., Colon, A. N., Mines, D., & Jivers, K. L.** (2014). Creating learner-centered assessment strategies for promoting greater student retention and class participation. *Frontiers in Psychology*, 5, Article 595.

**Schwartmann, R.** (2020). Zwischen Unmut und Demut: Ein Interview über die Rechts(un)sicherheit von Online-Prüfungen und deren Möglichkeiten. *Hochschulforum Digitalisierung*. <https://hochschulforumdigitalisierung.de/de/blog/interview-rechtsunsicherheit-online-pruefungen>

**Steffensen, B., & Schuster, K.** (2020). Studentische (Un-)Redlichkeit beim Erwerb von Studienleistungen. *Die Hochschullehre. Jahresheft 2020*, 604–613.

**Williams, J., & Wong, A.** (2009). The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology*, 40(2), 227–236.

**Wininger, S. R.** (2005). Using your test to teach: formative summative assessment. *Teaching Psychology*, 32, 164–166.

## Authors



Dr. Silke BRASELMANN || Friedrich Schiller University of Jena,  
Department of English and American Studies ||  
Ernst-Abbe-Platz 8, D-07743 Jena

[https://www.iaa.uni-jena.de/institut/mitarbeiter\\_innen/braselmann\\_silke](https://www.iaa.uni-jena.de/institut/mitarbeiter_innen/braselmann_silke)

[silke.braselmann@uni-jena.de](mailto:silke.braselmann@uni-jena.de)



Jolene MATHIESON || Leipzig University, Department of British  
Studies || Beethovenstraße 15 , D-04107 Leipzig

[https://www.philol.uni-leipzig.de/institut-fuer-anglistik/institut/team/](https://www.philol.uni-leipzig.de/institut-fuer-anglistik/institut/team/jolene.mathieson@uni-leipzig.de)

[jolene.mathieson@uni-leipzig.de](mailto:jolene.mathieson@uni-leipzig.de)



Oliver MOISICH || Friedrich Schiller University of Jena, Depart-  
ment of English and American Studies || Ernst-Abbe-Platz 8,  
D-07743 Jena

[https://www.iaa.uni-jena.de/institut/mitarbeiter\\_innen/moisich\\_oliver](https://www.iaa.uni-jena.de/institut/mitarbeiter_innen/moisich_oliver)

[oliver.moisch@uni-jena.de](mailto:oliver.moisch@uni-jena.de)

Bernhard SPANGL<sup>1</sup>, Dóra KÉRTESZ & Christian F. FREISLEBEN-TEUTSCHER (Wien/St. Pölten)

# Online Two-stage Exams als Prüfungsmethode in Statistik-Lehrveranstaltungen

## Zusammenfassung

Kollaborative Prüfungsszenarien sorgen während der Prüfung durch zusätzliches Feedback für eine weitere Lernerfahrung. Two-stage Exams im Besonderen kombinieren die individuelle Einzelleistung mit Gruppendiskussionen und erlauben es den Studierenden bei Unsicherheit in mehreren Versuchen die richtige Lösung zu erarbeiten. Anhand zweier Statistik-Lehrveranstaltungen wird die Überführung von papierbasierten Einzelprüfungen in Online Two-stage Exams während der Covid-19 Pandemie beschrieben und die Ergebnisse analysiert. Es konnte gezeigt werden, dass die Studierenden in mehrfacher Hinsicht von diesem Format profitieren und noch während der Prüfung ein Lernzuwachs erfolgt.

## Schlüsselwörter

Online Two-stage Exam, kollaborative Prüfungsmethode, Moodle, Videokonferenz, Mint-Fächer

---

<sup>1</sup> E-Mail: [bernhard.spangl@boku.ac.at](mailto:bernhard.spangl@boku.ac.at)



## Online two-stage exams as an examination method in statistics courses

### Abstract

Collaborative examination scenarios provide additional learning experience during an exam by providing additional feedback. Two-stage exams, in particular, combine the individual performance of students with group discussions and allow students to work out a correct solution using several attempts, if they are uncertain. This paper presents the transition from individual, paper-based exams to online two-stage exams in two statistics courses during the Covid-19 pandemic and investigates the examination results. The results show that students benefit from this method in multiple ways and that learning takes place during the examination.

### Keywords

online two-stage exam, collaborative examination method, Moodle, video conferencing, STEM disciplines

## 1 Einleitung

Constructive Alignment beschreibt das Zusammenspiel zwischen den Lernergebnissen, Lehrformaten und Prüfungsmodalitäten einer Lehrveranstaltung in wechselseitiger Abhängigkeit voneinander (BIGGS & TANG, 2011). Innerhalb dieser Triade des Modells bewirkt jegliche Änderung eines der drei Elemente in der Regel auch eine Überprüfung und Anpassung der beiden anderen. Aufgrund ihres hohen Stellenwerts spielen Prüfungen als formale Leistungsüberprüfungen eine Schlüsselrolle im studentischen Lernverhalten (BIGGS & TANG, 2011), weshalb sie geplant als Mittel zur Förderung des studentischen Lernens eingesetzt werden können (EFU, 2019).

Für anwendungsorientierte Lehrveranstaltungen der sogenannten MINT-Fächer aus Mathematik, Ingenieurwesen, Naturwissenschaften und Technik erweisen sich aktive und darunter kollaborative Lehr-Lernformate als förderlich für die Problemlösungskompetenzen der Studierenden (HELLER, KEITH & ANDERSON, 1992;

LINDSLEY et al., 2016). Eine Anpassung der Prüfungsmethoden im Sinne des Constructive Alignment bewirkt durch die Integration von kollaborativen Elementen in das Prüfungsdesign eine konsequente Anbindung an die Lehr-Lernmethoden der Lehrveranstaltung (RIEGER & HEINER, 2014).

Die während der Covid-19-Pandemie erfolgte Umstellung der Hochschullehre auf Online-Lehre ab dem Sommersemester 2020 bewirkte auch einen Umbruch im Prüfungsbetrieb durch den Wegfall der Möglichkeit Prüfungen in physischer Anwesenheit und mit Prüfungsaufsicht vor Ort abzuhalten. Für eine sinnvolle Übertragung der Prüfungen bestehender Statistik-Lehrveranstaltungen auf dislozierte Onlineprüfungen mussten diese nach der Neugestaltung zur Vermeidung einer Schlechterstellung der Studierenden eine Reihe von Anforderungen erfüllen, nämlich (a) die konsequente Fortführung des aktiven und kollaborativen Ansatzes der jeweiligen Lehrveranstaltung, (b) die Beschränkung der Überwachung der Studierenden auf ein notwendiges und noch akzeptables Maß zur Wahrung einer ordnungsgemäßen Durchführung, (c) die Förderung von Lernprozessen auch während der Prüfung und (d) die automatisierte Auswertbarkeit der Prüfungen in und mithilfe der Lernplattform Moodle.

Der Einsatz von Moodle Online-Tests im Rahmen eines kollaborativen Settings sollte ein Erreichen der gesetzten Lernergebnisse sicherstellen und gleichzeitig die Aktivierung der Studierenden auch während der Prüfungen ermöglichen, sodass diese nicht nur, sondern auch durch die und während der Prüfungen lernen. Es war naheliegend, die bereits vorhandenen Prüfungsfragen auch online unter Gewährung erlaubter Hilfsmittel einzusetzen.

## **2 Online Two-stage Exams als kollaboratives Prüfungsformat**

Two-stage Exams sind eine Variante kollaborativer Prüfungsformate, bei denen die Studierenden zunächst ihre Prüfung als Einzelleistung ablegen, um unmittelbar danach in Kleingruppen die gleichen oder überwiegend die gleichen Prüfungsfragen gemeinsam zu beantworten. Die Prüfungsnote wird entweder vorab als fixer, gewichteter Durchschnitt der Einzel- und Gruppenleistung festgelegt (CHEN & KINNIBURGH, 2019; MILLER & JAMES, 2019; NEWTON et al., 2019) oder

in Abhängigkeit der Teilleistungen variabel gestaltet (ZIPP, 2007; LEIGHT et al., 2012). Bei fixer Gewichtung macht die Einzelleistung dabei zwischen 75 bis 90 Prozent der Prüfungsnote aus, die restlichen Prozent des Notenanteils entfallen auf die Gruppenleistung (RIEGER & HEINER, 2014; CHEN & KINNIBURGH, 2019; MILLER & JAMES, 2019; NEWTON et al., 2019).

Während der Gruppenphase können die Studierenden zum einen ihre zuvor erbrachten Einzelleistungen nachträglich auf ihre Richtigkeit überprüfen, zum anderen müssen sie ihre Antworten begründen und vor ihren Peers argumentieren, um innerhalb einer Gruppe Konsens über die zu wählende Lösung zu erzielen, da nur eine Antwort pro Gruppe gewertet wird. Diese Auseinandersetzung mit den unterschiedlichen Standpunkten innerhalb der Gruppe, um zu einer gemeinsamen, einstimmigen Lösung zu kommen, und insbesondere das Feedback der anderen Gruppenmitglieder auf die vorgebrachten Sichtweisen und Argumente bewirkt eine erneute Auseinandersetzung mit dem Lernstoff, die mit dessen tiefergehenden Verarbeitung einhergeht, sodass ein Lernen noch während der Prüfung erfolgt (GILLEY & CLARKSTON, 2014). Die von der Gruppe gefundene Lösung kann unmittelbar auf ihre Richtigkeit geprüft und bei Bedarf unter Punkteabzug so oft abgeändert werden, bis die Frage korrekt beantwortet ist oder keine Punkte mehr für die Frage erzielt werden können. Der Mehrwert dieses unmittelbaren, automatisierten Feedbacks liegt in der Aufklärung von Missverständnissen und etwaigen Fehlkonzepten noch unmittelbar während der Prüfung (EFU, 2019).

Die Zuteilung der Studierenden zu den einzelnen Gruppen erfolgt vorzugsweise auf Basis der erzielten Leistungen (Hausübungen, Wissensüberprüfungen etc.) während des Semesters, sodass die Mitglieder einer Gruppe in etwa das selbe Leistungsniveau aufweisen. Damit soll vermieden werden, dass sehr gute Studierende eher schwache Studierende mittragen und so das tatsächliche Leistungsniveau verzerren (ZIPP, 2007).

Durch unterschiedliche Testfragen in den einzelnen Gruppen, bei denen eventuell nur die Angabe und/oder die Antwortmöglichkeiten leicht variieren, kann sichergestellt werden, dass sich die Studierenden nur innerhalb der Gruppe, nicht aber zwischen den Gruppen austauschen. Auch ist es nicht zwingend notwendig, sowohl für die Einzelphase als auch für die Gruppenphase dieselben Testfragen zu stellen, z. B. könnten im Individualteil zu einem Themenbereich mittels offener Fragen Zusammenhänge abgefragt und in der Gruppenphase zum selben Themenbereich mittels

Konzeptfragen mit vorgegebenen Antwortmöglichkeiten das Verständnis überprüft werden. Jedoch sollten in der Gruppenphase Fragen verwendet werden, die den Studierenden einen Rückschluss auf die Richtigkeit ihrer abgegebenen Antworten während der Einzelphase ermöglichen. Andernfalls entfällt das zeitnahe Feedback auf die Fragen dieser Phase und es kommt weder zu einer Korrektur von Missverständnissen noch zu einem Erfolgsgefühl.

### **3 Material und Methoden**

Der konkrete Einsatz der vorgestellten Lehr-Lern- und Prüfungsformate erfolgte im Studienjahr 2020/21 im Rahmen der Bachelor-Einführungslehveranstaltung „Angewandte Statistik“ und der Master-Vertiefungslehveranstaltung „Statistische Versuchsplanung“. Beide Lehrveranstaltungen wurden nach dem Modell „Inverted Classroom“ (LAGE, PLATT & TREGLIA, 2000) und zur Gänze online abgehalten. Die Studierenden wurden zu Beginn beider Lehrveranstaltungen über das Format „Inverted Classroom“ und den Prüfungsmodus „Two-stage Exam“ informiert.

Die wöchentliche Erarbeitung von Themen durchlief typischerweise vier Phasen: Vorbereitung, Just-in-Time-Teaching (SIMKINS & MAIER, 2010), Vertiefung und Feedback. In der Vorbereitungsphase erarbeiteten sich die Studierenden die Theorie anhand vorgegebener Literatur individuell im Selbststudium vor der Lehrveranstaltungseinheit. Ihre Kommentare und Fragen zu den Lerninhalten wurden schriftlich an den Vortragenden übermittelt, sodass dieser bereits vor der Online-Präsenz eine erste Einschätzung zum Kenntnisstand der Studierenden treffen konnte. In der zweiten Phase, die als Just-in-Time-Teaching konzipiert war, erfolgte mittels Konzeptfragen unter Verwendung von Audio-Response-Systemen in einem „Peer Instruction“-Setting (MAZUR, 1997) die Vertiefung, Korrektur sowie die Festigung des erarbeiteten Wissens. Die Studierenden beantworteten während der Online-Präsenz wiederholt Konzeptfragen zunächst individuell mithilfe der Abstimmungsfunktion in Zoom. Bei zu wenigen richtigen Antworten galt es in der darauf folgenden Diskussion in Kleingruppen von vier bis sechs Personen die Gruppenmitglieder durch fachliche Argumente von der Richtigkeit der getroffenen Wahl zu überzeugen. Die an die Gruppendiskussion anschließende erneute Abstimmung über die ursprüngliche Frage ging schließlich mit einer Auflösung der Frage einher. „Peer

Instruction“ war somit eine direkte Vorbereitung auf die Gruppenphase des „Two-stage Exams“. Die restliche Zeit der Online-Einheit wurde zur Vertiefung der Inhalte genutzt (Phase 3). Am Ende jeder Einheit schließlich konnten die Studierenden anonym ein kurzes 3-Punkte-Feedback auf die Lehrveranstaltungseinheit geben. Die Rücklaufquoten lagen im Durchschnitt bei ca. 20 Prozent.

Die Prüfungen der beiden Lehrveranstaltungen<sup>2</sup> waren als „Online Two-stage Exams“ organisiert, die mithilfe der Lernplattform Moodle durchgeführt wurden. Diese beiden Teilprüfungen, die jeweils unterschiedliche Themengebiete behandelten, konnten entweder zu verschiedenen Terminen (in der Mitte des Semesters und am Semesterende) oder gemeinsam zu Semesterende<sup>3</sup> absolviert werden. Die Prüfungsaufsicht und die Hilfestellung bei Problemen erfolgte mithilfe der Video-Konferenz-Software Zoom.

Die Prüfungseinstellungen während der individuellen ersten Phase erlaubten weder Feedback noch eine Rückmeldung des Punktestandes an die Studierenden. Die Studierenden durchliefen die zweite Phase der Prüfung unmittelbar nach der ersten. Hierfür wurden sie in der laufenden Zoom-Sitzung in Breakout-Sessions einer Gruppe von vier bis sieben Personen zugeteilt. In jeder Gruppe wurde einem Gruppenmitglied als Gruppensprecher\*in die Möglichkeit eingeräumt, den eigenen Bildschirm für die Dauer der zweiten Phase mit der restlichen Gruppe zu teilen. Anschließend durchliefen die Gruppen parallel einen der ersten Phase ähnlichen Online-Test, wobei nur der\*die Gruppensprecher\*in die Möglichkeit hatte, die Antwort der Gruppe in Moodle einzugeben. Im Unterschied zur ersten Phase hatten die Studierenden die Möglichkeit eines (mehrmaligen) Prüfens der eingegebenen Antwort(en) mit adaptivem Punkteabzug. Als Feedback erhielten die Studierenden „richtig“ oder „falsch“. Die Prüfungsdauer betrug 20 Minuten für die Einzelphase und 40 Minuten für die Gruppenphase. Im Anschluss an alle abgehaltenen Two-stage Exams wurde von den teilnehmenden Studierenden ein Feedback zu diesem Prüfungsmodus eingeholt.

---

2 Die Gesamtnote der Lehrveranstaltung „Angewandte Statistik“ setzte sich aus den beiden Teilprüfungen und mehreren, über das gesamte Semester verteilten praktischen Projektarbeiten, auf die hier nicht näher eingegangen wird, zusammen.

3 Der Wiederholungstermin der ersten Teilprüfung zu Semesterende fand zeitlich vor dem zweiten Teilprüfungstermin statt.

### 3.1 Die Lehrveranstaltung „Angewandte Statistik“

Die Lehrveranstaltung „Angewandte Statistik“ im Ausmaß von 2 ECTS ist für Studierende des Bachelorstudiums Landschaftsplanung und Landschaftsarchitektur an der Universität für Bodenkultur Wien verpflichtend. Sie wird als integrierte Lehrveranstaltung (VU) mit immanentem Prüfungscharakter abgehalten und findet regulär laut Studienplan im 5. Semester statt. Der Lehrveranstaltungstyp „VU“ sieht mehrere Leistungsfeststellungen während des Semesters vor. In der Lehrveranstaltung „Angewandte Statistik“ erfolgen diese in Form von kurzen Tests vor jeder Einheit und mehreren, über das Semester verteilten praktischen Projektarbeiten, die es erlauben, die Studierenden aufgrund dieser Leistungen für die erste Teilprüfung in homogene Gruppen einzuteilen. Im Wintersemester 2020/21 waren 124 Studierende zu dieser Veranstaltung angemeldet. Die statistischen Kennzahlen der Prüfungsergebnisse sind in Tabelle 1 aufgelistet.

Tab. 1: Statistische Kennzahlen der Prüfungsergebnisse (Lehrveranstaltung „Angewandte Statistik“).

	1. Teilprüfung						2. Teilprüfung			
	Testgruppe A		Testgruppe B		Wiederholung		Testgruppe A		Testgruppe B	
	E	G	E	G	E	G	E	G	E	G
Anzahl <sup>4</sup>	38	7	53	9	7	2	42	8	51	8
Mögliche Punkteanzahl	13.00		13.00		12.00 <sup>5</sup>		12.00		12.00	
Min. Punkteanzahl	6.00	12.17	6.00	11.67	8.00	11.00	6.00	10.50	2.00	9.00
Mittelwert	10.67	12.65	10.49	12.49	10.43	11.14	9.74	11.43	7.57	11.02
Standardabweichung	1.73	0.27	1.53	0.53	1.51	0.24	1.74	0.48	2.13	0.96
Max. Punkteanzahl	13.00	13.00	13.00	13.00	12.00	11.50	12.00	12.00	11.00	12.00

4 In der Spalte „E“ (Einzelphase) ist die Anzahl der Studierenden, in der Spalte „G“ (Gruppenphase) die Anzahl der Breakout-Sessions angegeben.

5 Eine der 13 Testfragen der ersten Teilprüfung wurde zum Wiederholungstermin gestrichen und nicht durch eine Alternative ersetzt.

### 3.2 Die Lehrveranstaltung „Statistische Versuchsplanung“

Die Lehrveranstaltung „Statistische Versuchsplanung“ im Ausmaß von 3 ECTS ist für Studierende des Masterstudiums Nutzpflanzenwissenschaften an der Universität für Bodenkultur Wien verpflichtend. Sie wird als Vorlesung (VO) mit abschließender Prüfung abgehalten. Da dieser Lehrveranstaltungstyp keine weiteren Leistungsfeststellungen vorsieht, erfolgte die Gruppenzuteilung der Studierenden für die erste Teilprüfung im Unterschied zur Lehrveranstaltung „Angewandte Statistik“ zufällig. Für die zweite Teilprüfung konnte eine Zuordnung zu einer Gruppe aufgrund der während der ersten Teilprüfung erzielten Leistungen erfolgen. Im Sommersemester 2021 waren 78 Studierende angemeldet. Die statistischen Kennzahlen der Prüfungsergebnisse sind in Tabelle 2 aufgelistet.

Tab. 2: Statistische Kennzahlen der Prüfungsergebnisse (Lehrveranstaltung „Statistische Versuchsplanung“).

	1. Teilprüfung				2. Teilprüfung			
	Testgruppe A		Wiederholung		Testgruppe A		Testgruppe B	
	E	G	E	G	E	G	E	G
Anzahl <sup>6</sup>	26	5	8	2	18	4	13	3
Mögliche Punkteanzahl	12.00		12.00		12.00		12.00	
Min. Punkteanzahl	4.25	11.00	4.00	11.8	4.00	8.34	3.33	8.67
Mittelwert	9.22	11.78	10.44	11.9	7.13	9.84	6.89	10.60
Standardabweichung	2.39	0.42	2.70	0.11	2.05	1.16	2.04	1.34
Max. Punkteanzahl	12.00	12.00	12.00	12.00	10.75	11.42	10.00	11.50

### 3.3 Statistische Analyse

Die statistische Analyse wurde mit der Statistik-Software R (R CORE TEAM, 2021) durchgeführt. Zur Überführung der Gleichheit bzw. Unterschiedlichkeit der Testgruppen zu den einzelnen Teilprüfungsterminen wurden Äquivalenztests (TOST;

<sup>6</sup> In der Spalte „E“ (Einzelfase) ist die Anzahl der Studierenden, in der Spalte „G“ (Gruppenphase) die Anzahl der Breakout-Sessions angegeben.

R Paket TOSTER; LAKENS, 2017) bzw. ein t-Test verwendet. Für die Äquivalenztests wurde als Schranke  $d = 1.5$  gewählt.<sup>7</sup> Für den statistischen Nachweis einer Leistungssteigerung zwischen Einzel- und Gruppenphase wurden gepaarte t-Tests verwendet. P-Werte mit  $p < 0.05$  wurden als statistisch signifikant erachtet. Die Wiederholungstermine wurden aufgrund der geringen Stichprobengrößen nicht statistisch ausgewertet.

## 4 Ergebnisse & Evaluierung

Der folgende Abschnitt analysiert die Prüfungsergebnisse der Studierenden beider Lehrveranstaltungen.

### 4.1 Die Lehrveranstaltung „Angewandte Statistik“

Die durchschnittlichen erreichten Punkte in den beiden Testgruppen der 1. Teilprüfung bezogen jeweils auf die Einzel- (TOST:  $p < 0.001$ ) und Gruppenphase (TOST:  $p < 0.001$ ) sind gleich. Hingegen unterscheiden sich die durchschnittlichen Einzelleistungen in den beiden Testgruppen der 2. Teilprüfung signifikant (t-Test:  $p < 0.001$ ). Die durchschnittlichen Gruppenleistungen in den beiden Testgruppen der 2. Teilprüfung sind wieder vergleichbar (TOST:  $p < 0.001$ ).

Abbildung 1 zeigt die Ergebnisse für das Two-stage Exam im Detail. Auf der vertikalen Achse sind die erzielten Gesamtpunkte aufgetragen. Die Linien entsprechen den einzelnen Studierenden und verbinden die erzielten Gesamtpunkte in der Einzelphase (jeweils links in den Grafiken der Abbildung 1) mit den erreichten Punkten in der Gruppenphase (jeweils rechts in den Grafiken der Abbildung 1). Die Farben entsprechen den Gruppen der Gruppenphase. Die Gruppen sind entsprechend der durchschnittlichen Übungsleistung der Studierenden absteigend angeordnet.

---

<sup>7</sup> Das entspricht in etwa einem Notengrad bei einer Maximalpunktzahl von 12 Punkten und dem Notenschlüssel: „Nicht Genügend“ (< 50%), „Genügend“ (< 62.5%), „Befriedigend“ (< 75%), „Gut“ (< 87.5%), „Sehr Gut“ (sonst).

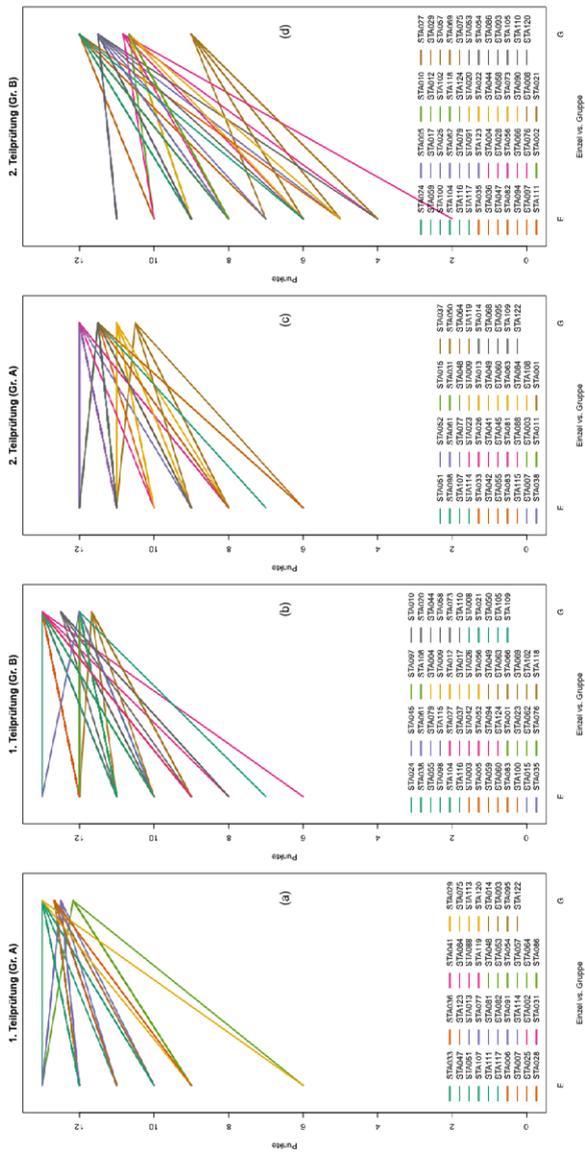


Abb. 1: Punkteverteilung des Two-stage Exams im Detail (Lehrveranstaltung „Angewandte Statistik“). Die Linien entsprechen den einzelnen Studierenden und verbinden die erzielten Gesamtpunkte in der Einzelphase E mit den erreichten Punkten in der Gruppenphase G. Die Farben entsprechen den Gruppen der Gruppenphase.

Bei allen Studierenden kam es in der Gruppenphase im Vergleich zur jeweiligen Einzelleistung zu einer statistisch signifikanten Steigerung der erzielten Punkte (gepaarte t-Tests; 1. Teilprüfung, Testgruppe A: mittlere Steigerung um 1.89 Punkte,  $p < 0.001$ , Testgruppe B: mittlere Steigerung um 2.00 Punkte,  $p < 0.001$ ; 2. Teilprüfung, Testgruppe A: mittlere Steigerung um 1.69 Punkte,  $p < 0.001$ , Testgruppe B: mittlere Steigerung um 3.45 Punkte,  $p < 0.001$ ). Die Verbesserung der Leistungen in der Gruppenphase legt nahe, dass in dieser Phase ein Lernzuwachs stattfand. Diese Annahme deckt sich mit den in der Literatur publizierten Ergebnissen (vgl. etwa GILLEY & CLARKSTON, 2014; LEIGHT et al., 2012; STEARNS, 1996) und dem Feedback der Studierenden. Die hier präsentierten Analyseergebnisse bestätigen auch Auswertungen der Prüfungsergebnisse aus einem früheren Durchgang im Wintersemester 2019/20 (vgl. SPANGL & KERTESZ, 2020).

## **4.2 Die Lehrveranstaltung „Statistische Versuchsplanung“**

Die Gleichheit der durchschnittlich erreichten Punkte in den beiden Testgruppen der 2. Teilprüfung bezogen jeweils auf die Einzel- (TOST:  $p = 0.052$ ) und Gruppenphase (TOST:  $p = 0.061$ ) kann statistisch knapp nicht verifiziert werden.

Wiederum kam es, wie in Abbildung 2 dargestellt, bei fast allen Studierenden in der Gruppenphase im Vergleich zur jeweiligen Einzelleistung zu einer statistisch signifikanten Steigerung der erzielten Punkte (gepaarte t-Tests; 1. Teilprüfung: mittlere Steigerung um 2.56 Punkte,  $p < 0.001$ ; 2. Teilprüfung, Testgruppe A: mittlere Steigerung um 2.71 Punkte,  $p < 0.001$ , Testgruppe B: mittlere Steigerung um 3.71 Punkte,  $p < 0.001$ ).

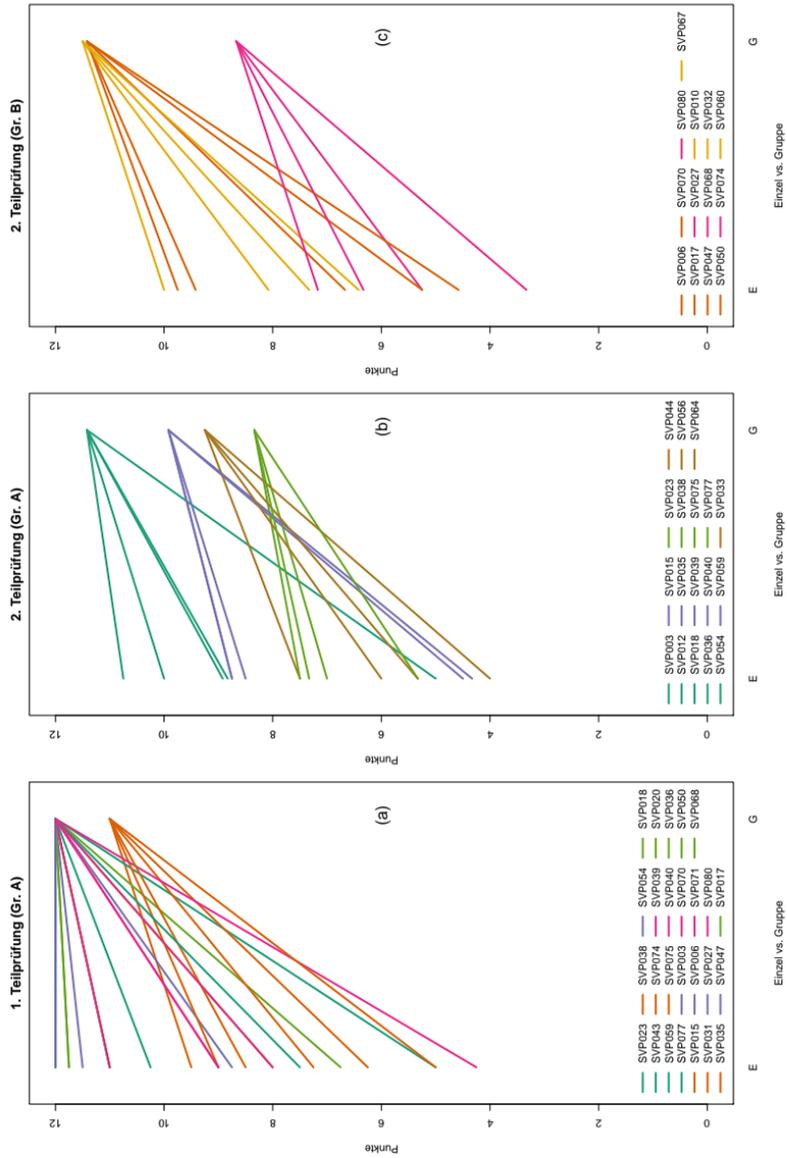


Abb. 2: Punkteverteilung des Two-stage Exams im Detail (Lehrveranstaltung „Statistische Ver- suchsplanung“). Linien und Farbcodierung sind analog zu Abbildung 1.

Da eine Zuordnung der Studierenden zu homogenen Gruppen aufgrund einer fehlenden Leistungsfeststellung nicht möglich war, wurden die Studierenden zufällig in Gruppen aufgeteilt. Diese zufällige Zuordnung der Studierenden wird in Abbildung 2a besonders deutlich. Auch heterogene Gruppen, bezogen auf die individuellen Einzelleistungen innerhalb der Gruppe, erreichten in der Gruppenphase die maximale Punkteanzahl. Leistungsstarke Studierende zogen vermutlich die schwächeren mit und verzerren so das tatsächliche Leistungsniveau. Die leistungsabhängige Zuteilung für die 2. Teilprüfung ist in Abbildungen 2b und 2c gut ersichtlich; die Einzelleistungen in den Gruppen sind deutlich homogener und die Ergebnisse der Gruppenphase entsprechen dem in der Einzelphase erbrachten Leistungsniveau der Studierenden.

## 5 Diskussion

Der Mehrwert des Two-stage Exams entstand für die Studierenden vor allem in der Gruppenphase. Die inhaltliche Auseinandersetzung innerhalb der Gruppe und die Rückmeldungen der anderen Gruppenmitglieder auf die vorgebrachten Argumente bewirken eine erneute Auseinandersetzung mit dem Lernstoff und führen zu einem Wissenserwerb noch während der Prüfung. In ihrem Feedback zum Prüfungsformat bestätigten die Studierenden in Übereinstimmung mit der Literatur eine Festigung des Gelernten und die Korrektur der im Individualteil gemachten Fehler. Die statistische Auswertung in Abschnitt 4 belegt eine Verbesserung der Ergebnisse während der Gruppenphase.

Eine Überwachung der Studierenden während der Online-Prüfung ist nur eingeschränkt möglich. Eine Kollaboration über Kanäle außerhalb der verwendeten Video-Konferenz-Software und Prüfungsplattform kann nicht unterbunden und auch nicht nachgewiesen werden.

Peer Instruction und Two-stage Exams fördern kommunikative und kooperative Kompetenzen der Studierenden (RIEGER & HEINER, 2014; KINNEAR, 2021). Damit Online Two-stage Exams erfolgreich durchgeführt werden können, müssen die Studierenden in der Lage sein, ihre Lösung in Kleingruppen auszuhandeln. Die Studierenden müssen daher zur Vorbereitung auf das Prüfungsformat ausreichend die Möglichkeit erhalten, die hierfür benötigten Kompetenzen zu erwerben bezie-

ungsweise zu vertiefen. Diese Kompetenzen werden jedoch in den Lehrveranstaltungszielen nicht gesondert ausgewiesen. Auch der Beitrag für nachfolgende Lehrveranstaltungen ist unklar.

Die Lernziele der beschriebenen Lehrveranstaltungen adressieren unterschiedliche Kompetenzen, die durch das gewählte Prüfungsformat nur unvollständig abgedeckt werden können. Hier bedarf es einer genaueren Passung der Lernziele, der Lehr-Lernmethoden und der Prüfungsmodalitäten. Nach Absolvierung der vorliegenden Lehrveranstaltungen sollten die Studierenden statistische Aufgabenstellungen rechnerisch lösen und die Ergebnisse interpretieren sowie auf Plausibilität prüfen können. Für die beiden letzteren Ziele wurden die Lehr-Lernmethoden und die Leistungsüberprüfung mithilfe von Two-stage Exams aufeinander abgestimmt. Das praktische Lösen statistischer Fragestellungen erfolgt innerhalb der Lehrveranstaltung „Angewandte Statistik“ im Rahmen von Übungsbeispielen; für die Vorlesung „Statistische Versuchsplanung“ wurde hingegen dieser Kompetenzerwerb in eine eigene Lehrveranstaltung ausgelagert. Die Wechselwirkung zwischen den Übungsbeispielen, die als Rechenaufgaben konzipiert sind, den Peer-Instruction-Einheiten beziehungsweise den Two-stage Exams wurde bisher jedoch weder adressiert noch untersucht. Es ist daher unklar, in welchem Ausmaß die Rechenaufgaben zum Konzeptverständnis beitragen und der Prüfungsvorbereitung dienen. Die Auswirkungen der Prüfungen auf die Problemlösungskompetenzen der Studierenden vor dem Hintergrund einer praktischen Anwendung des Gelernten wurde bisher ebenso wenig berücksichtigt.

In Online Two-stage Exams sind während der Gruppenphase nur geschlossene Prüfungsfragen, die eine sofortige und automatische Benotung ermöglichen, einsetzbar. Andernfalls entfällt das zeitnahe Feedback, das einen Rückschluss auf die Richtigkeit der gegebenen Antworten zulässt, und es kommt weder zu einer Korrektur von Missverständnissen noch zu einem Erfolgsgefühl der Studierenden. Einen entsprechenden Fragenpool vorausgesetzt, erfordert das Format einen Mehraufwand bei der Gruppenzuteilung der Studierenden. Zusätzlich verdreifacht sich die Prüfungsdauer aufgrund der Diskussionen in der Gruppenphase.

Ein langfristiger Lerntransfer aufgrund des Two-stage Exam konnte in der Literatur nicht eindeutig belegt werden (CHEN & KINNIBURGH, 2019). Weiterer Forschungsbedarf besteht in Hinblick auf einen nachhaltigen Lerntransfer und in welchem Ausmaß die Studierenden das erworbene Wissen für ihre Bachelor- (Landschaftsplanung) bzw. Masterarbeit (Nutzpflanzenwissenschaften) anwenden können.

## 6 Fazit

Die Überführung von papierbasierten Einzelprüfungen in Online Two-stage Exams während der Covid-19-Pandemie an der Universität für Bodenkultur Wien unterstrich das große Potenzial für das Erlangen einer gesteigerten Problemlösungskompetenz in Statistik-Lehrveranstaltungen. Das Feedback der Studierenden bestätigte den zusätzlichen Kompetenzerwerb während der Gruppenphase.

Two-stage Exams erfordern das Argumentieren der gewählten Lösung durch die Studierenden. Diese gaben in ihren Rückmeldungen zu den Lehrveranstaltungen an, dass ihnen die Diskussionen während der Gruppenphase besonders gut gefielen. Sie stimmten mehrheitlich auch für die Beibehaltung dieses Prüfungsformats.

Um den Einfluss des Prüfungsmodus Two-stage Exam auf den Lerntransfer näher zu untersuchen, sind im kommenden Abhaltungszyklus der Lehrveranstaltungen Cross-over-Studien, wie in LEIGHT et al. (2012) oder GILLEY & CLARKSTON (2014) beschrieben, geplant.

### Danksagung

Wir danken den anonymen Gutachter\*innen für die wertvollen Anmerkungen, die wesentlich zur Verbesserung des Manuskripts beigetragen haben.

## 7 Literaturverzeichnis

**Biggs, J. B., & Tang, C. S.** (2011). *Teaching for quality learning at university: What the student does* (4. ed). Maidenhead: McGraw-Hill, Society for Research into Higher Education & Open University Press.

**Chen, S., & Kinniburgh, S.** (2019). A Controlled Experiment on Two-Stage Exams in an Introductory Statistics Course. *The International Journal of Assessment and Evaluation*, 26(2), 1–12. <https://doi.org/10.18848/2327-7920/CGP/v26i02/1-12>

**Efu, S. I.** (2019). Exams as Learning Tools: A Comparison of Traditional and Collaborative Assessment in Higher Education. *College Teaching*, 67(1), 73–83. <https://doi.org/10.1080/87567555.2018.1531282>

**Gilley, B., & Clarkston, B.** (2014). Research and Teaching: Collaborative Testing: Evidence of Learning in a Controlled In-Class Study of Undergraduate Students. *Journal of College Science Teaching*, 043(03). [https://doi.org/10.2505/4/jcst14\\_043\\_03\\_83](https://doi.org/10.2505/4/jcst14_043_03_83)

**Heller, P., Keith, R., & Anderson, S.** (1992). Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving. *American Journal of Physics*, 60(7), 627–636. <https://doi.org/10.1119/1.17117>

**Kinnear, G.** (2021). Two-Stage Collaborative Exams have Little Impact on Subsequent Exam Performance in Undergraduate Mathematics. *International Journal of Research in Undergraduate Mathematics Education*, 7, 33–60. <https://doi.org/10.1007/s40753-020-00121-w>

**Lage, M. J., Platt, G. J., & Treglia, M.** (2000). Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment. *The Journal of Economic Education*, 31(1), 30–43. <https://doi.org/10.2307/1183338>

**Lakens, D.** (2017). Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177>

**Leight, H., Saunders, C., Calkins, R., & Withers, M.** (2012). Collaborative Testing Improves Performance but Not Content Retention in a Large-Enrollment Introductory Biology Class. *CBE Life Sciences Education*, 11(4), 392–401. <https://doi.org/10.1187/cbe.12-04-0048>

**Lindsley, J. E., Morton, D. A., Pippitt, K., Lamb, S., & Colbert-Getz, J. M.** (2016). The Two-Stage Examination: A Method to Assess Individual Competence and Collaborative Problem Solving in Medical Students. *Academic Medicine*, 91(10), 1384–1387. <https://doi.org/10.1097/ACM.0000000000001185>

**Mazur, E.** (1997). *Peer instruction: A user's manual*. Upper Saddle River: Prentice Hall.

**Miller, S. T., & James, C. R.** (2019). *What Is The Impact Of Collaborative Exams On Learning And Attitudes In Introductory Astronomy Classes?* 6(1), 13–17. <https://clutejournals.com/index.php/JAESE/article/view/10289>

**Newton, G., Rajakaruna, R., Kulak, V., Albabish, W., Gilley, B. H., & Ritchie, K.** (2019). Two-Stage (Collaborative) Testing in Science Teaching: Does It Improve Grades on Short-Answer Questions and Retention of Material? *Journal of College Science Teaching*, 48(4), 64–73. <https://www.jstor.org/stable/26901301>

**R Core Team** (2021). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>

**Rieger, G., & Heiner, C.** (2014). Examinations That Support Collaborative Learning: The Students' Perspective. *Journal of College Science Teaching*, 043(04). [https://doi.org/10.2505/4/jcst14\\_043\\_04\\_41](https://doi.org/10.2505/4/jcst14_043_04_41)

**Simkins, S., & Maier, M. H. (Hrsg.)**. (2010). *Just-in-time teaching: Across the disciplines, across the academy* (1st ed). Sterling: Stylus Pub.

**Spangl, B., & Kertesz, D.** (2020). Wenn Studierende bei der Statistikprüfung „Juhu“ rufen ...: Two-Stage Exams mit Moodle und Zoom in Zeiten von Corona. In J. Weißenböck, W. Gruber & Ch. Freisleben-Teutscher (Hrsg.), *Digital Learning in Zeiten von Corona – nachhaltiger Entwicklungsschub für die Hochschulen?* (S. 65–76). St. Pölten: Fachhochschule St. Pölten GmbH.

**Stearns, S. A.** (1996). Collaborative Exams as Learning Tools. *College Teaching*, 44(3), 111–112. <https://doi.org/10.1080/87567555.1996.9925564>

**Zipp, J. F.** (2007). Learning by Exams: The Impact of Two-Stage Cooperative Tests. *Teaching Sociology*, 35(1), 62–76. <https://doi.org/10.1177/0092055X0703500105>

## Autor\*in\*en



DI Dr. Bernhard Spangl || Universität für Bodenkultur Wien,  
Department für Raum, Landschaft und Infrastruktur, Institut für  
Statistik || Gregor-Mendel-Str. 33, A-1180 Wien

<https://boku.ac.at/rali/stat>

[bernhard.spangl@boku.ac.at](mailto:bernhard.spangl@boku.ac.at)



DI Dóra Kértész || FH Technikum Wien, Teaching & Learning  
Center || Höchstädtplatz 6, A-1200 Wien

<https://www.technikum-wien.at/das-teaching-and-learning-center-der-fhtw/>

[dora.kertesz@technikum-wien.at](mailto:dora.kertesz@technikum-wien.at)



Mag. Dr. Christian Freisleben-Teutscher || FH St. Pölten, SKILL  
(Service- und Kompetenzzentrum für Innovatives Lehren und  
Lernen) || Campus-Platz 1, A-3100 St. Pölten

<https://skill.fhstp.ac.at/>

[christian.freisleben-teutscher@fhstp.ac.at](mailto:christian.freisleben-teutscher@fhstp.ac.at)

Christian SPODEN<sup>1</sup> (Emden), Aron FINK (Frankfurt), Andreas FREY (Frankfurt), Hanna KÖHLER (Jena) & Patrick NAUMANN (Frankfurt)

# Kompetenzorientierung und Fairness bei individualisierten E-Klausuren

## Zusammenfassung

Die stärkere Individualisierung des Studiums erfordert adaptiv und individuell zugeschnittene Prüfungsformen, die Anforderungen der Kompetenzorientierung sowie der Vergleichbarkeit und Fairness von Prüfungsleistungen erfüllen. Kompetenzorientierte adaptive E-Klausuren stellen hier ein erfolgsversprechendes Konzept dar, welches im vorliegenden Beitrag erläutert wird. Es wird ferner eine Anwendung dargestellt, aus der sich Empfehlungen zur Umsetzung im regulären Klausurbetrieb ableiten lassen. Abschließend werden Erweiterungen und Herausforderungen der Konzeption diskutiert.

## Schlüsselwörter

E-Klausuren, Kompetenzdiagnostik, Computerisiertes Adaptives Testen, Test Equating

---

<sup>1</sup> E-Mail: [christian.spoden@hs-emden-leer.de](mailto:christian.spoden@hs-emden-leer.de)



## **Competence orientation and fairness in individualised e-exams**

### **Abstract**

Individualised pathways in higher education studies require adaptive and customised exams, which meet the requirements of a genuine competence orientation and the comparability and fairness of exams. For this purpose, competence-oriented adaptive e-exams provide a promising concept, which is explained in this paper. In addition, the paper presents an trial implementation, from which recommendations for implementation in regular exam operations can be derived. Possible extensions and challenges of the concept are summarised in the discussion.

### **Keywords**

e-exams; assessment of competencies; computerised adaptive testing; test equating

## **1 Die Individualisierung des Studiums als Herausforderung für das Prüfungswesen**

Das Hochschulwesen in Deutschland ist seit einiger Zeit durch eine stärkere Ausdifferenzierung von Fachgebieten sowie der Aufspaltung der Studiengänge und Vergrößerung des Studienangebotes gekennzeichnet (HACHMEISTER & GREVERS, 2019). In vielen Fächern beschränkten sich Wahlmöglichkeiten lange auf den Studienort mit dem dort ansässigen Lehrpersonal aus einer Gruppe gleichartig konzipierter Diplom- oder Magister-Studiengänge an Universitäten oder Fachhochschulen. Mit dem Wechsel auf Bachelor-/Master-Studiengänge infolge der Bologna-Reform hat sich dies gewandelt. Heute existieren in vielen Fächern nicht nur Hochschul- und Fachhochschulstudiengänge, sondern Studiengänge mit unterschiedlicher Schwerpunktsetzung sowie Studiengänge in neuen Querschnittsbereichen. Auch hat sich die Durchlässigkeit zwischen Studiengängen infolge der Etablierung von Bachelor-/Master-Studiengängen erhöht. Als Resultat eines gestiegenen Bewusstseins um die Diversität der Studierendenschaft im Hinblick auf Lernvoraussetzungen

und -gelegenheiten im Studium entwickelte sich außerdem ein breiteres Angebot unterschiedlicher Studienformen (Präsenz- oder Fernstudium in Voll- oder Teilzeit, berufsbegleitendes Studium), in dem die Studierenden oft auf unterschiedliche Art studieren (zur Mediennutzung ZAWACKI-RICHTER, 2015). Schließlich hat sich etwa in Deutschland der Anteil ausländischer Studierender und damit die Diversität der Lernvoraussetzungen erhöht.

Für das Prüfungswesen ist diese Tendenz zu einer stärkeren Individualisierung des Studiums eine Herausforderung. Einerseits fehlen überzeugende Konzepte, um diesen vielfältigeren Studienwegen bei Prüfungen gerecht zu werden. Andererseits geht mit Individualisierung auch die Frage nach der Vergleichbarkeit und somit der Fairness von Prüfungen, der Aufrechterhaltung wissenschaftlicher Standards und der Umsetzung der geforderten kompetenzorientierten Prüfungen unter diesen Bedingungen einher. Studierende haben nicht nur ein Anrecht auf eine angemessene Bilanzierung der im Studium angeeigneten Kompetenzen, sondern erwarten, dass Prüfungen überzeugende Fairness-Konzepte enthalten (SAMBELL, McDOWELL & BROWN, 1997).

Die Herausforderung liegt nun darin, bei Hochschulprüfungen Individualisierung, Vergleichbarkeit und Fairness der Bewertung sowie Kompetenzorientierung gemeinsam zu realisieren. Ein möglicher Lösungsansatz für diese Problemstellung liegt in der konsequenten Anwendung etablierter und auf der *Item Response Theory* (IRT; van der LINDEN, 2016) basierender Methoden aus den Bereichen Educational Measurement und Psychometrie. Insbesondere bietet sich das computerisierte adaptive Testen (CAT; FREY, 2020) als individualisierter Ansatz für computerbasierte Test- und Prüfungsverfahren an. Mit dem vorliegenden Beitrag werden daher zwei Ziele verfolgt: Erstens wird ein neues Konzept individualisierter und gleichzeitig fairer, kompetenzorientierter Hochschulklausuren beschrieben. Zweitens wird die Umsetzung dieses Konzeptes für Studiengangverantwortliche und Prüfende an einem illustrativen Beispiel in drei Entwicklungsphasen dargestellt. Abschließend werden in der Diskussion Voraussetzungen und Einschränkungen der Nutzbarkeit aufgeführt, aber auch mögliche Erweiterungen der Konzeption vorgeschlagen.

## 2 Konzeption adaptiver kompetenzorientierter E-Klausuren

Kompetenzorientierte adaptive E-Klausuren verknüpfen eine ernsthafte Kompetenzorientierung und eine über Studierendengruppen hinweg konstante und damit faire Berichtsmetrik für Klausurergebnisse in einem Lehrgebiet mit der Individualisierung der Prüfungen durch die computerisiert-adaptive Auswahl der Prüfungsaufgaben (im Folgenden: Items). Sie basieren auf vier Prinzipien: *Erstens* sind adaptive kompetenzorientierte E-Klausuren als Instrumente zum Messen des Umfangs des Kompetenzerwerbs konzipiert. Um den Messgegenstand zu spezifizieren und festzulegen, was in der Klausur geprüft werden soll, müssen die Lernziele einer Lehrveranstaltung anhand von Inhaltsbereichen und kognitiven Anforderungen explizit dargelegt werden. Dies ist notwendig, um das Erreichen kompetenzorientierter Lernziele zu beurteilen und kein reines Wissen abzufragen. Für die Abgrenzung unterschiedlicher kognitiver Anforderungsniveaus sind Lehrzieltaxonomien (z. B. BLOOM, 1956) als Orientierung geeignet. Für jede Kombination von Inhalt und kognitivem Anforderungsniveau sind dann Items zu konstruieren. Im Sinne des *Constructive Alignment* (BIGGS, 1996) ist diese Strukturierung dahingehend bereits nützlich, Lernaktivitäten zu planen und Lernziele, -methoden und Prüfung schon bei der Planung einer Lehrveranstaltung aufeinander zu beziehen. In aller Regel werden zudem kontextualisierte Klausuritems benötigt, die auf typische wissenschaftliche und praktische Tätigkeiten in der jeweiligen Disziplin Bezug nehmen.

Um Aussagen über das Erreichen der in den Modulkatalogen festgelegten Lernziele und Kompetenzanforderungen bei Klausuren ableiten zu können, ist *zweitens* ihre Konzeption als kriteriumsorientierte Tests (HERZBERG & FREY, 2011) und die Auswertung der Klausuren mithilfe der IRT zielführend. Die Modelle der IRT erlauben Aussagen dazu, mit welcher Wahrscheinlichkeit eine mit Lernzielen verknüpfte Aufgabe eines bestimmten Schwierigkeitsgrads ( $\delta_i$ ) von Studierenden mit einer aus den kodierten Itemlösungen (richtige Lösung: 1; falsche Lösung: 0) geschätzten Kompetenz  $\theta_\nu$  (für Student  $\nu$ ) bewältigt werden können. Sie sind damit als testtheoretisches Rahmenwerk für kompetenzorientierte Tests besonders geeignet. Im einparametrischen IRT-Modell wird diese Wahrscheinlichkeit bestimmt als:

$$P(x = 1 | \theta_v, \delta_i) = \frac{e^{(\theta_v - \delta_i)}}{1 + e^{(\theta_v - \delta_i)}}. \quad (1)$$

Die resultierenden Kompetenzschätzungen können in einem zweiten Schritt über vereinfachte Standard-Setting-Verfahren wie der Bookmark-Methode (MITZEL, LEWIS, PATZ & GREEN, 2001) in Notenstufen überführt werden. Die Grenzen zwischen Notenstufen werden mit Schlüsselanforderungen verknüpft, die in bestimmten Items abgebildet werden, sodass die Grenze zwischen Notenstufen an der Schwierigkeit dieser Items festgemacht wird. Mithilfe dieses Vorgehens werden die auf Kompetenzschätzungen basierenden Klausurergebnisse an die individuelle Bewältigung der Kompetenzanforderungen geknüpft. Zweifelhafte Bewertungspraktiken, nach denen das Bestehen der Klausur etwa an einem willkürlich festgesetzten Anteil richtiger Itemlösungen festgemacht wird, werden ausgeschlossen.

*Drittens* ist die Etablierung einer zeitlich konstanten Kompetenzskala und die kohortenübergreifende Verortung von Kompetenzen auf dieser Skala zu nennen. Mithilfe von *Equating*-Methoden (KOLEN & BRENNAN, 2014) ist es wiederum im Rahmenwerk der IRT möglich, den Bewertungsmaßstab einer Klausur auf die Klausurbewertungen in nachfolgenden Kohorten zu übertragen. Hierzu werden bei der Bestimmung der Studierendenkompetenz und, darauf aufbauend, der Noten die Unterschiede in der mittleren Schwierigkeit zwischen unterschiedlichen eingesetzten Itemstichproben berücksichtigt. Notwendig ist dafür, dass ein kleiner Teil der Items (Linkitems) aus vorangegangenen Klausuren erneut eingesetzt und im Rahmen einer *Fixed Parameter Calibration* wieder auf die gleichen Schwierigkeiten geschätzt werden kann. Auf Basis dieser konstanten Schätzer wird die statistische „Verankerung“ der Klausuren über die Zeit vorgenommen. Die Schwierigkeiten werden dabei im ersten Schritt frei geschätzt. Dann werden die Schwierigkeiten über eine lineare Transformation um den geschätzten mittleren Kompetenzunterschied zwischen den Studierendenkohorten verschoben, wodurch für die Linkitems wieder die Schwierigkeiten aus vorherigen Klausuren erreicht werden sollten. Die in der Realität resultierenden Abweichungen werden statistisch auf *Item-Drift* (z. B. DeMARS, 2004) geprüft. In einem weiteren Schritt wird die Fixed Parameter Calibration mit fixierten Schwierigkeiten jener Linkitems, bei denen keine signifikanten Abweichungen im Drift-Test identifiziert wurden, vorgenommen. Abbildung 1 zeigt anhand sogenannter *Wright-Maps* beispielhaft, wie sich die Itemschwierigkeiten (rechts) und die

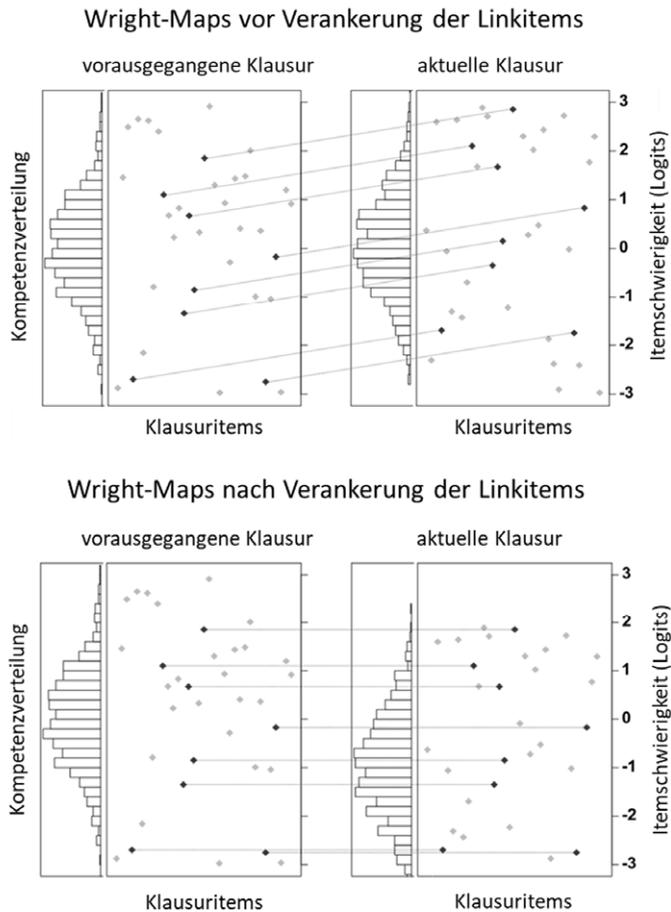


Abb. 1: Wright-Maps vor und nach Verankerung der Linkitems

---

Kompetenzverteilung (links) zwischen freier Skalierung und Fixed Item Parameter Calibration anhand dieser Verankerung verschieben. Wichtig für das Gelingen dieses Verfahrens ist die Vermeidung systematischer Verzerrungen der Schätzung der Itemschwierigkeiten, die durch gezieltes Vorbereiten der Itemlösung durch Studierende bei Bekanntwerden der Items resultieren könnte. Hierzu ist es notwendig, die Klausuritems geheim zu halten und nur Beispielitems freizugeben.

Das Konstanthalten des Bewertungsmaßstabs und die glaubhafte Vermittlung der Tatsache, dass bei einem Zweittermin oder einer Nachklausur keine bessere Note aufgrund leichter Items zu erreichen ist (wie es bei Festlegung der Bestehensgrenze anhand eines festen Prozentsatzes richtiger Lösungen geschehen könnte), dürfte bei Studierenden zur Verringerung der Prokrastination beim Lernen beitragen. Bei gleichbleibenden Lernzielen und Anforderungen in Modulkatalogen leistet der konstante Bewertungsmaßstab unabhängig vom Kompetenzniveau der jeweils getesteten Kohorte insbesondere aber einen Beitrag zur Verbesserung der Fairness (FREY, SPODEN & BORN, 2020). Unter den Studierenden kursierende Itemsammlungen und „Studierendenfunk“ sorgen dafür, dass Studierende die Klausuranforderungen über Jahrgänge vergleichen. Das vorgeschlagene Vorgehen liefert Transparenz in der Vergleichbarkeit von Prüfungsergebnissen über Jahrgänge hinweg. Davon abgesehen sind Vergleiche der durchschnittlichen Kompetenzen über Studierendenkohorten geeignet, eine Form von Monitoring in der Lehre einzusetzen und gegebenenfalls Anpassungen vorzunehmen (siehe Beispiel unten).

*Viertens* ermöglichen adaptive E-Klausuren die individualisierte Itemauswahl durch den Computer. Die computerbasierte Durchführung ist bereits aufgrund der Reduktion des Prüfungsaufwands durch Automatisierung großer Teile der Klausuradministration, -auswertung und Ergebnisrückmeldung (Zeitgewinn ausführlich an einem Beispiel dargestellt in SPODEN, 2021) sowie der nachhaltigeren, weil papierlosen Klausurdurchführung vorteilhaft. Die Möglichkeit für eine adaptive Itemauswahl ergibt sich aus der Eigenschaft der IRT, dass Kompetenzen auf Basis unterschiedlicher Itemstichproben geschätzt und auf derselben Skala abgebildet werden können. Bei üblichen Klausuren wird allen Studierenden, unabhängig von Lernvoraussetzungen, individuellen Schwerpunktsetzungen im Studium oder Vorleistungen, der gleiche Itemsatz vorgelegt. Dies spiegelt nicht die Individualisierung des Studiums wider. Beim CAT werden die Klausuritems hingegen aus einem vorab zusammengestellten Itempool anhand bekannter Informationen wie der Item-

schwierigkeit vom Computer individuell ausgewählt. Meist zielt die Itemauswahl darauf ab, informative Items im Hinblick auf die Verortung der Studierenden auf der Kompetenzdimension vorzulegen und so im gesamten Kompetenzspektrum den Messfehler zu reduzieren. Im einparametrischen Modell bestimmt sich die Iteminformation als

$$I_i(\theta) = P_i(\theta)[1 - P_i(\theta)]. \quad (2)$$

Die maximale Information (Maximum Fisher Item Information Criterion) im Hinblick auf die Verortung der Kompetenz eines Prüflings besitzt ein Item, wenn die Lösungswahrscheinlichkeit gleich der Gegenwahrscheinlichkeit (bzw. die Itemschwierigkeit gleich der Kompetenzausprägung) ist. Der obere Teil von Abbildung 2 verdeutlicht diesen Itemauswahlprozess beim CAT in Form eines Flussdiagramms (nach FREY, 2020); unten ist in Form eines Entscheidungsbaums dargestellt, wie sich die Itemauswahl anhand der Schwierigkeit an der nach jedem bearbeiteten Item jeweils neu geschätzten Kompetenz orientiert. Die Reduktion des Messfehlers im gesamten Kompetenzspektrum durch CAT ist vorteilhaft, da dieser bei Tests ohne CAT-Einsatz im Vergleich zum mittleren Kompetenzbereich üblicherweise an den Rändern der Kompetenzverteilung höher ausfällt (FREY & EHMKE, 2007), so dass kritische diagnostische Entscheidungen (insb. Klausur bestanden / nicht bestanden im unteren Kompetenzbereich) mit einer hohen Fehlerwahrscheinlichkeit getroffen werden. Neben dem Kriterium der maximalen Information kann mit Verfahren des Content Management sichergestellt werden, dass die vorgegebenen Items weiteren Anforderungen genügen. Beispielsweise ist es möglich, Inhaltsbereiche entsprechend zu kodieren, bei der Itemauswahl diese Inhaltsbereiche systematisch abzudecken und so die Repräsentativität der Items in Bezug auf die Lehrinhalte bei adaptiver Itemvorgabe aufrechtzuerhalten. Denkbar ist auch, unterschiedliche Itemkontexte in Codes zu überführen, die vom Computer für eine noch stärker individualisierte Itemauswahl genutzt werden können. In Lehrveranstaltungen finden sich oft Studierende unterschiedlicher Studiengänge. Es wäre möglich, parallele Items zu entwickeln, bei denen der Kontext auf die unterschiedlichen Anforderungen in diesen Studiengängen ausgerichtet wird und der Computer studiengangsspezifisch passende Items auswählt. Als Beispiel zeigt Abbildung 3 ein Übungsitem zu der unten beschriebenen Klausur im Fach Forschungsmethoden der Erziehungswissenschaft mit einem wirtschaftspädagogischen Kontext (Kantine) sowie ein inhalt-

lich gleiches Parallelitern mit schulpädagogischer Kontextualisierung. Analog zum Itemkontext ist denkbar, dass Studierende bei der Klausurbearbeitung die in ihrem Arbeitsbereich jeweils üblichen Hilfsmittel wie fachspezifische Softwarepakete freigeschaltet bekommen.

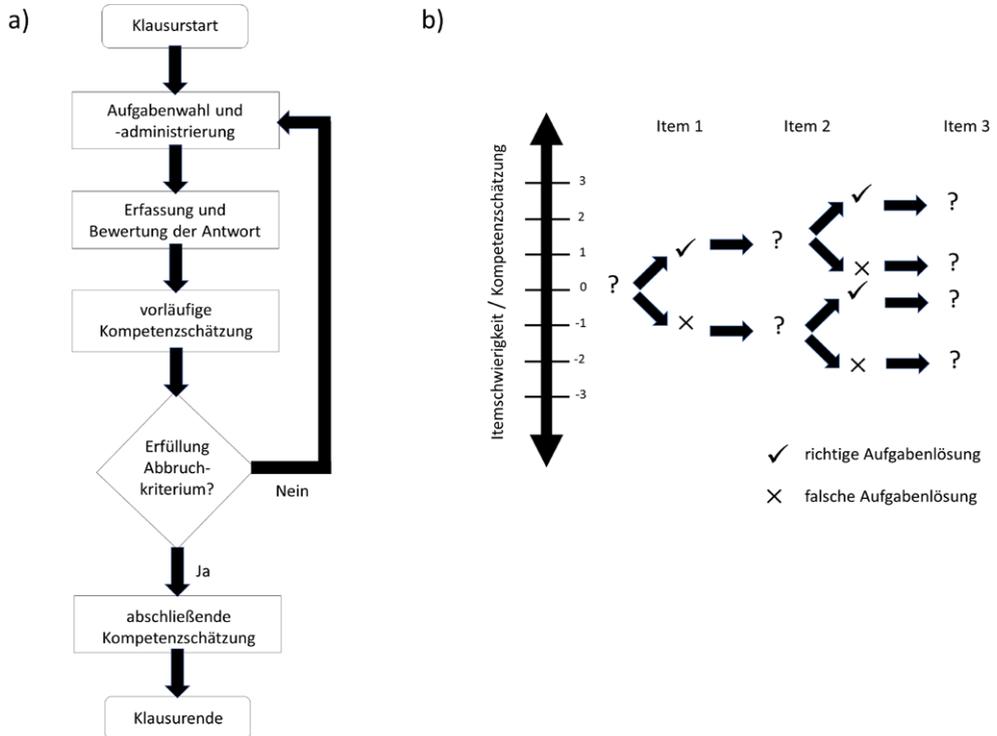


Abb. 2: a) Ablaufschema beim CAT. Angepasst aus „Computerisiertes adaptives Testen“, Frey, 2020, S. 505.  
 b) Entscheidungsbaum zum CAT

Stimuli	<b>Kantine</b> <p>In einem Unternehmen soll entschieden werden, ob eine neue Kantine gebaut werden sollte. Hierzu wird eine Stichprobe der Arbeitnehmer befragt, um abschätzen zu können, ob diese eine neue Kantine wünschen. Die Nullhypothese lautet, dass die Arbeitnehmer an einer neuen Kantine nicht interessiert sind; die Alternativhypothese lautet, dass die Arbeitnehmer interessiert sind. Die Daten zeigen, dass mehr als 40% der Studienteilnehmer an einer neuen Kantine interessiert sind.</p> <p><b>Aufgabe</b></p> <p>Welche der folgenden Aussagen ist eine Konsequenz eines Fehlers 2. Art in diesem Kontext?</p> <p>Bitte markieren Sie die richtige Antwort!</p>	<b>Schulbuchverlag</b> <p>In der Lehrerkonferenz soll entschieden werden, ob die Schule auf die Schulbücher eines neuen Schulbuchverlags wechseln soll. Hierzu erhielt eine Stichprobe der Lehrkräfte kostenfreie Exemplare der neuen Bücher und wurde im Anschluss befragt, um abschätzen zu können, ob die Nutzung der neuen Schulbücher gewünscht ist. Die Nullhypothese lautet, dass die Lehrkräfte an den neuen Schulbüchern nicht interessiert sind; die Alternativhypothese lautet, dass die Lehrkräfte Schulbücher zukünftig vom neuen Verlag beziehen wollen. Die Daten zeigen, dass mehr als 40% der befragten Lehrkräfte an einem Wechsel des Schulbuchverlages interessiert sind.</p> <p><b>Aufgabe</b></p> <p>Welche der folgenden Aussagen ist eine Konsequenz eines Fehlers 2. Art in diesem Kontext?</p> <p>Bitte markieren Sie die richtige Antwort!</p>
	<b>Antwortoptionen</b>	<p>[ a ] Das Unternehmen zieht es nicht in Betracht eine Kantine zu bauen, obwohl es dies sollte.</p> <p>[ b ] Das Unternehmen zieht es nicht in Betracht eine Mensa zu bauen, weil es dies nicht sollte.</p> <p>[ c ] Das Unternehmen zieht es in Betracht eine Kantine zu bauen, obwohl es dies nicht sollte.</p> <p>[ d ] Das Unternehmen zieht es in Betracht eine Kantine zu bauen, weil es dies sollte.</p>

Abb. 3: Item aus einer Klausur „Forschungsmethoden der Erziehungswissenschaft“ in zwei kontextualisierten Fassungen

Einschränkend ist zu Möglichkeiten der Individualisierung zu bedenken, dass E-Klausuren im Hinblick auf die Authentizität, also die Überprüfbarkeit der Zuordnung einer Prüfungsleistung zu einem Prüfling, und die Integrität, die sich auf die Sicherstellung der Nicht-Veränderung der Daten nach der Prüfung bezieht, anderen Anforderungen als papierbasierte Klausuren genügen müssen. Einige Lösungen hierfür werden im Anwendungsbeispiel unten beschrieben. Grundsätzlich sind bei E-Klausuren außerdem Vorkehrungen zur sicheren und störungsfreien Durchführung zu treffen, etwa eine Absicherung gegen Stromausfall. Zumeist dürfte diese Absicherung in universitären Computerpools gut umsetzbar sein.

### **3 Anwendungsbeispiel: Entwicklung einer individualisierten Klausur zu empirischen Forschungsmethoden**

Aufbauend auf dieser Kurzdarstellung des Konzeptes (Zielsetzung 1) wird nun die Umsetzung kompetenzorientierter adaptiver E-Klausuren an einem Anwendungsbeispiel illustriert (Zielsetzung 2). Die Implementation erfolgte in drei Phasen.

#### **3.1 Phase 1: Kompetenzorientierte Klausuren als kriteriumsorientierter Test**

Das Assessment Framework für sechs angeschlossene Klausurzyklen zu Forschungsmethoden der Erziehungswissenschaft bezog sich auf Designs, Datenerhebungen und Inhalte der einführenden Statistik für die Anwendung im Bereich Bildung und Erziehung. Für die Klausuren wurde ein Itempool generiert, der zwischen 2012 und 2017 von zunächst 49 auf 124 Items anwuchs. Die Items repräsentierten überwiegend die Stufen Wissen, Verständnis und Anwendung und wurden kontextualisiert mit Bezug zu Anwendungsbereichen der Erziehungswissenschaft konstruiert. Die Klausuren wurden wie folgt administriert (FREY, SPODEN & BORN, 2020): In jeder Klausur mit einer Prüfungszeit von 90 Minuten wurden 35 und 40 Items vorgegeben, von denen mindestens 15 Items als Linkitems aus vorausgegangenen Klausuren übernommen wurden. Die Linkitems wurden genutzt, um die Klausuren, wie

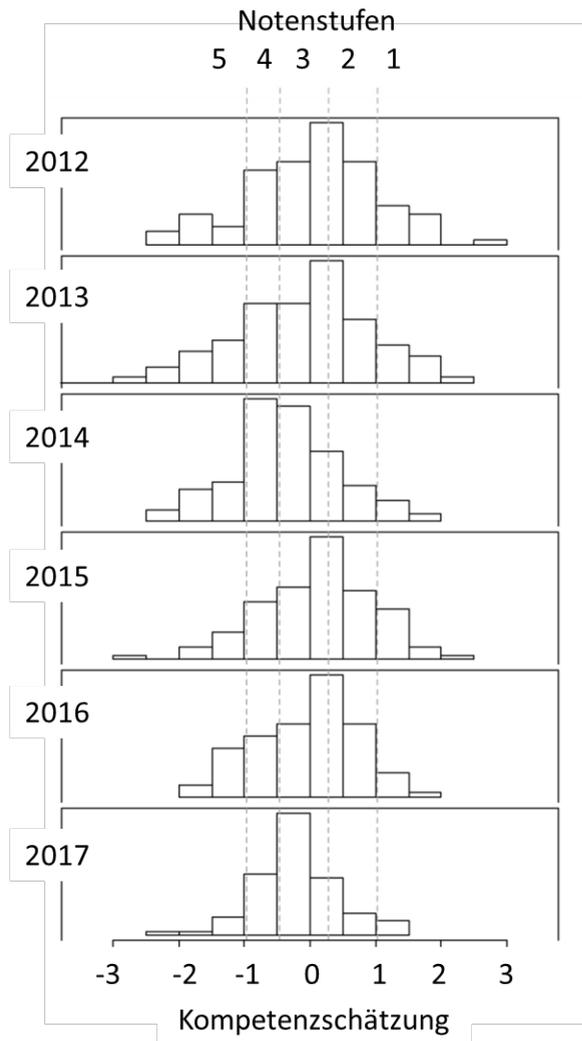


Abb. 4: Notenstufen und Kompetenzschätzungen der Klausur „Forschungsmethoden der Erziehungswissenschaft“ (2012–2017). Angepasst aus FREY, SPODEN & BORN, 2020, S. 482.

zuvor erläutert, über ein Equating-Verfahren statistisch zu verbinden. Die für diese Items geschätzten Schwierigkeiten wurden auf Item-Drift geprüft, wobei jeweils für mindestens 10 Items eine Verankerung auf die früheren Schwierigkeiten möglich war. Auf Basis eines *Common Item Nonequivalent Group-Designs* wurde so eine über sechs Jahrgänge reliable Skala etabliert. Für die Notenvergabe wurden auf der Kompetenzskala Grenzen zwischen Notenstufen anhand der Bookmark-Methoden eingezogen. Die Ergebnisse sind in Abbildung 4 ersichtlich. Durch den Vergleich der durchschnittlichen Kompetenzen der Studierendenkohorten auf einer gemeinsamen Kompetenzdimension wurden Unterschiede zwischen Studierendenkohorten offenbart, die mit Veränderungen in der Lehre in Verbindung standen. Der Kompetenzeinbruch 2014 kann auf eine Reduzierung der Anzahl begleitender Tutorien zurückgeführt werden. Die niedrigen Kompetenzausprägungen 2017 lassen sich durch eine Umstellung der Lehre in einem häufig gewählten Nebenfach auf neue, im Vergleich zur vorigen Veranstaltung nicht mehr mit Forschungsmethoden in Verbindung stehende Inhalte erklären. In der Zusammenschau zeigen die Ergebnisse, dass die Etablierung einer längsschnittlich angelegten, reliablen Kompetenzskala im Klausurbetrieb möglich ist und sich interessante Einblicke zur Erklärung beobachteter Klausurergebnisse und Schlussfolgerungen zur Anpassung der Lehre ableiten lassen.

### **3.2 Phase 2: Kompetenzorientierte E-Klausuren als kriteriumsorientierter computerbasierter Test**

Um die individualisierte Itemauswahl beim CAT zu realisieren, müssen Charakteristika der Items inklusive der Schwierigkeit bekannt sein. Letztere muss aus empirischen Daten bei einer computerbasierten Testadministration geschätzt werden (Testkalibrierung). Unter der Bedingung kleiner Stichproben, wie sie im Hochschulwesen häufig anzutreffen sind, existieren zwei Möglichkeiten: Zum einen kann die kontinuierliche Kalibrierungsstrategie (FINK et al., 2018) genutzt werden. Zum anderen ist eine gemeinsame Kalibrierung über verschiedene Studienorte möglich, bei der oft mehr Items kalibriert werden können. Die Itemschwierigkeiten wurden hier in Vorbereitung eines adaptiven Tests zunächst auf Basis einer solchen kooperativen und computerbasierten Datenerhebung geschätzt. Der Test wurde an drei Standorten unter leicht unterschiedlichen Durchführungsbedingungen administriert: In Jena als Präsenzklausur mit 69 Studierenden der Erziehungswissenschaft; in Bo-

chum und Ulm als Online-Übungsklausur mit 27 beziehungsweise 43 Psychologiestudierenden. An jedem Standort wurden 30 Items eingesetzt, 15 Items (50%) wurden gemeinsam an allen Standorten administriert, um die Skalierung über drei Standorte vorzunehmen. Somit existierten vier Cluster von Items (gemeinsame Linkitems sowie spezifische Items an den drei Standorten). Die Klausuren wurden mithilfe der KAT-HS-App (FINK et al., 2021) vorgegeben. Bei der Klausur in Jena wurde die Kommunikation nach außen durch Nutzung des Safe Exam Browsers (HALBHERR et al., 2016) mit eingeschränkten Nutzungsrechten unterbunden. Die Schätzung der Itemschwierigkeiten erfolgte mithilfe des Mehrfacetten-Rasch-Modells mit den Facetten Itemschwierigkeit und Universitätsort. Gemeinsame Schwierigkeiten konnten letztlich nur für sieben Linkitems bestimmt werden, bei denen keine Unterschiede in der Lösungswahrscheinlichkeit zwischen den Standorten unter Kontrolle der durchschnittlichen Leistungen der Studierendengruppen gefunden wurden, also kein *Differential Item Functioning* (DIF; z. B. OSTERLIND & EVERSON, 2009) vorlag. Mithilfe dieser 7 Linkitems war es möglich, Itemschwierigkeiten für den kompletten Itempool zwischen -1.508 und 4.058 zu schätzen. Die mittlere Kompetenz der Studierenden wurde in Jena auf 0.097, in Bochum auf 0.393 und in Ulm auf -0.489 geschätzt (negative Werte entsprechen hier besseren Kompetenzschätzungen). Diese deutlichen Unterschiede sowie der hohe Anteil von Items, die bei der standortübergreifenden Skalierung DIF aufwiesen, legen nahe, dass die kontinuierliche Item-Kalibrierung für die Vorbereitung eines adaptiven Tests zu bevorzugen ist.

### **3.3 Phase 3: Kompetenzorientierte E-Klausuren als kriteriumsorientierter, computerbasierter adaptiver Test**

In Phase 3 wurde eine computerbasierte und teilweise adaptive E-Klausur mit 10 Testversionen à 25 Items aus fünf Item-Clustern realisiert. Der CAT-Algorithmus realisierte die adaptive Itemauswahl im Safe Exam Browser mit eingeschränkten Nutzungsrechten und der KAT-HS-App entsprechend dem Maximum Fisher Item Information Criterion. Zur Identitätsfeststellung wurden Studierendenausweis und Matrikelnummer genutzt. Die Integrität der Prüfungsleistung und die Funktionsfähigkeit des Prüfungssystems wurden anhand von Log-Daten sichergestellt. Die Prüfungs- und Log-Datei der Studierenden wurde auf einem sicheren Universitäts-

server mit einem Hashwert, der Informationen zur Prüfung in Ganzzahlen abbildet, hinterlegt.

Die Klausur wurde von  $N = 84$  Erstsemesterstudierenden (87% weiblich, 13% männlich) mit Hauptfach Erziehungswissenschaft (Bachelor) in Jena bearbeitet, denen vorab die Möglichkeit eingeräumt wurde, das elektronische Klausursystem zu erproben. Die mittlere geschätzte Kompetenzausprägung lag bei  $M = 0.03$  ( $SD = 0.79$ ), die Reliabilität der Klausur lag bei  $rel_{WLE} = .722$ . Diese Ergebnisse verdeutlichen, dass kompetenzorientierte adaptive Hochschulklausuren sicher und reliabel durchführbar sind. Da frühere Studien auch negative Effekte von CAT beim emotionalen Erleben aufgezeigt haben (z. B. PITKIN & VISPOEL, 2001), beantworteten die Studierenden einige Minuten vor der Klausurdurchführung einen Fragebogen zum emotionalen Erleben, um konventionelle Klausuren zu bewerten (Pretest), und nochmals unmittelbar nach der Klausur zur Bewertung des neuartigen Klausurkonzepts (Posttest). Der Vergleich zeigte, dass sowohl positive (Pretest:  $M = 2.77$ ,  $SD = 0.59$ ; Posttest:  $M = 2.65$ ,  $SD = 0.69$ ;  $t(80) = 2.21$ ,  $p = .03$ ), als auch negative Emotionen (Pretest:  $M = 2.31$ ,  $SD = 0.69$ ; Posttest:  $M = 2.10$ ,  $SD = 0.70$ ;  $t(81) = 4.71$ ,  $p < .01$ ) bei der neu konzipierten Klausur niedriger ausfielen, also ein gewisses Maß an Indifferenz nach der kognitiv anspruchsvollen Klausur ausgedrückt wurde. Diese Ergebnisse legen nahe, dass die Vorteile einer adaptiven Klausur nicht durch Nachteile beim emotionalen Erleben der Studierenden aufgewogen werden.

## 4 Diskussion

Der vorliegende Beitrag war mit den Zielen verknüpft, (1) adaptive kompetenzorientierte E-Klausuren als einen Ansatz für individualisierte und gleichzeitig fair und vergleichbar konzipierte Klausuren vor- und (2) die Umsetzung in drei Entwicklungsphasen am Beispiel einer Klausur zu Forschungsmethoden der Erziehungswissenschaft darzustellen. Bei adaptiven kompetenzorientierten E-Klausuren werden etablierte Methoden aus den Bereichen Educational Measurement und Psychometrie basierend auf der IRT an die spezifischen Herausforderungen von Hochschulprüfungen (z. B. der kleineren Stichprobengröße) angepasst. Die bisherigen Implementationsergebnisse (siehe auch unten) erlauben eine optimistische Einschätzung zukünftiger Nutzungsmöglichkeiten des Gesamtkonzeptes.

Der Einsatz des hier vorgestellten Konzeptes ist dabei an Voraussetzungen gebunden. Die Verlinkung von Klausuren über die Zeit und die adaptive Administration bedürfen, wie zuvor erwähnt, eines geheim gehaltenen Itempools. Das Konzept ist somit für den Einsatz bei den während der Covid-19-Pandemie eingesetzten, online über das Internet administrierten Klausuren kaum geeignet, da hier die Gefahr des Item-Diebstahls mithilfe von Screenshots hoch ist. Neben dem denkbaren Einsatz anderer Cheating-Indikatoren sollte insbesondere die Schwierigkeit wiederholt genutzter Items regelmäßig auf Item-Drift geprüft werden, um systematisches Vorbereiten auf möglicherweise publik gewordene Items aufzudecken. Ferner sollte berücksichtigt werden, dass der effiziente Einsatz von E-Klausuren eine entsprechende Infrastruktur wie bestenfalls ein E-Testzentrum notwendig macht. Interessanterweise haben sich die Infrastrukturbedingungen in einer Befragung von Prüfenden an 74 Hochschulen in Deutschland (SPODEN et al., 2020) nicht als relevanter Prädiktor der Intention zur Nutzung von E-Klausuren herausgestellt. Dies könnte damit zusammenhängen, dass die digitale Infrastruktur mit Computerpools tatsächlich besser ausgebaut ist, als zuweilen angenommen. Schließlich sollte angemerkt werden, dass die Umsetzung anspruchsvoller Teile des Konzeptes bei vielen Prüfenden Fortbildungsmöglichkeiten in der Hochschuldidaktik (und bei der Erstellung adaptiver Prüfungen in Testsoftware gegebenenfalls „handwerklicher“ Unterstützung durch die für E-Klausuren verantwortlichen Personen bedarf, die inzwischen in vielen hochschuldidaktischen Abteilungen angestellt sind). Um auf diese Herausforderung zu reagieren, wurde eine Fortbildungsveranstaltung zum Konzept erstellt, durchgeführt und evaluiert (Ergebnisse nicht publiziert). Obwohl die Anforderungen an Wissen zur IRT als anspruchsvoll eingeschätzt wurden, zeigten die Ergebnisse der Evaluation bei den Teilnehmenden deutliche Wissenszuwächse bezüglich der Konstruktion von E-Klausuren nach dem Workshop. Auch beinhalten die Ergebnisse eine positive Einschätzung der Nützlichkeit der Veranstaltung sowie eine klar geäußerte Nutzungsabsicht bezüglich des Klausurkonzeptes.

Gleichzeitig ist hervorzuheben, dass das Konzept auf alternative Prüfungsformen übertragen oder auf formative Assessments zur Lernbegleitung erweitert werden kann. Trotz der hohen Relevanz, die Modulabschlussklausuren im Prüfungswesen besitzen, haben Befragungen gezeigt, dass Studierende eine kontinuierliche Erfassung der Lernleistungen und -fortschritte als Alternative zu einer einzigen, das „Bulimielernen“ provozierenden Prüfung am Semesterende präferieren (z. B. SAMBELL, McDOWELL, & BROWN, 1997). Auch wünschen sie sich individuelles

Feedback zu ihren Ergebnissen. Das vorliegende Konzept kann hierzu mit lernförderlichen Konzepten der formativen Assessments oder des E-Learnings verknüpft werden. So könnten an adaptive Kurztests im Semesterverlauf Tutor- oder Empfehlungssysteme angebunden werden, die Hilfestellungen oder individuelle Hinweise zur Weiterarbeit bereitstellen. Die Vorteile der hier vorgestellten Prinzipien sind dabei, dass aufgrund des adaptiven Modus bei kurzen Tests bereits reliable Informationen zum Lernstand auf der Berichtsmetrik der späteren Prüfungen im Sinne einer Diskrepanz zu Zielstellungen (wie erwarteten Kompetenzen an einer bestimmten Stelle im Studium) bereitgestellt werden. Auch semesterbegleitende Teilklausuren lassen sich so zeitökonomisch durchführen und die Ergebnisse auf der etablierten Berichtsmetrik verorten. Obwohl adaptive kompetenzorientierte E-Klausuren also an bestimmte Anwendungsvoraussetzungen gebunden sind, beinhalten sie doch gleichzeitig zahlreiche Möglichkeiten, um deutliche Entwicklungssprünge bei der Konstruktion von Hochschulklausuren zu erzielen.

## 5 Literaturverzeichnis

**Biggs, J.** (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <https://doi.org/10.1007/bf00138871>

**Bloom, B.** (Hrsg.) (1956). *Taxonomy of Educational Objectives, the classification of educational goals – Handbook I: Cognitive Domain*. New York: McKay.

**DeMars, C. E.** (2004). Detection of item parameter drift over multiple test administrations. *Applied Measurement in Education*, 17, 265–300. [https://doi.org/10.1207/s15324818ame1703\\_3](https://doi.org/10.1207/s15324818ame1703_3)

**Fink, A., Born, S., Spoden, C., & Frey, A.** (2018). A continuous calibration strategy for computerized adaptive testing. *Psychological Test and Assessment Modeling*, 60, 327–346.

**Fink, A., Spoden, C., Frey, A., & Naumann, P.** (2021). Kriteriumsorientiertes adaptives Testen mit der KAT-HS-App. *Diagnostica*, 67, 110–114. <https://doi.org/10.1026/0012-1924/a000268>

**Frey, A.** (2020). Computerisiertes adaptives Testen. In H. Moosbrugger & A. Kelaiva (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (3. Aufl., S. 501–524). Berlin: Springer.

**Frey, A., & Ehmke, T.** (2007). Hypothetischer Einsatz adaptiven Testens bei der Überprüfung von Bildungsstandards. *Zeitschrift für Erziehungswissenschaft, Sonderheft 8*, 169–184. [https://doi.org/10.1007/978-3-531-90865-6\\_10](https://doi.org/10.1007/978-3-531-90865-6_10)

**Frey, A., Spoden, C., & Born, S.** (2020). Construction of psychometrically sound written university exams. *Psychological Test and Assessment Modeling*, 65(4), 472–486.

**Hachmeister, C.-D., & Grevers, J.** (2019). *Im Blickpunkt: Die Vielfalt der Studiengänge 2019. Entwicklung des Studienangebotes in Deutschland zwischen 2014 und 2019*. Gütersloh: CHE.

**Halbherr, T., Dittmann-Domenichini, N., Piendl, T., & Schlienger, C.** (2016). Authentische, kompetenzorientierte Online-Prüfungen an der ETH Zürich. *Zeitschrift für Hochschulentwicklung*, 11(2), 247–269.

**Herzberg, P. Y., & Frey, A.** (2011). Kriteriumsorientierte Diagnostik. In L. F. Hornke, M. Amelang & M. Kersting (Hrsg.), *Methoden der psychologischen Diagnostik. Enzyklopädie der Psychologie*, B/II/2 (S. 281–324). Göttingen: Hogrefe.

**Kolen, M. J., & Brennan, R. L.** (2014). *Test Equating, Scaling and Linking: Methods and Practices* (3rd ed.). New York, NY: Springer.

**Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R.** (2001). The bookmark method: Psychological perspectives. In G. J. Cizek (Hrsg.), *Setting performance standards: Concepts, methods, and perspectives* (S. 249–281). Mahwah, NJ: Erlbaum.

**Osterlind, S. J., & Everson, H. T.** (2009). *Differential item functioning*. Thousand Oaks, CA: Sage Publishing.

**Pitkin, A. K., & Vispoel, W. P.** (2001). Differences between self-adapted and computerized adaptive tests: A meta-analysis. *Journal of Educational Measurement*, 38(3), 235–247. <https://doi.org/10.1111/j.1745-3984.2001.tb01125.x>

**Sambell, K., McDowell, L., & Brown, S.** (1997). “But is it fair?": An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, 23, 349–371. [https://doi.org/10.1016/S0191-491X\(97\)86215-3](https://doi.org/10.1016/S0191-491X(97)86215-3)

**Spoden, C.** (2021). Ressourceneinsatz bei der Erstellung psychometrisch fundierter Hochschulklausuren. In C. Spoden & A. Frey (Hrsg.), *Psychometrisch fundierte E-Klausuren für die Hochschule* (S. 106–111). Lengerich: Pabst.

**Spoden, C., Frey, A., Fink, A., & Naumann, P.** (2020). Kompetenzorientierte elektronische Hochschulklausuren im Studium des Lehramts. In K. Kaspar, M. Becker-Mrotzek, S. Hofhues, J. König & D. Schmeinck (Hrsg.), *Bildung, Schule und Digitalisierung* (S. 184–189). Münster: Waxmann.

**van der Linden, W. J.** (Hrsg.). (2016). *Handbook of item response theory. Volume one: Models*. London: Chapman & Hall/CRC. <https://doi.org/10.1201/9781315374512>

**Zawacki-Richter, O.** (2015). Zur Mediennutzung im Studium – unter besonderer Berücksichtigung heterogener Studierender. *Zeitschrift für Erziehungswissenschaft*, 18, 527–549. <https://doi.org/10.1007/s11618-015-0618-6>

## Autor\*in\*en



Prof. Dr. Christian SPODEN || Hochschule Emden/Leer, Fachbereich Wirtschaft || Constantiaplatz 4, D-26723 Emden

[www.hs-emden-leer.de](http://www.hs-emden-leer.de)

[christian.spoden@hs-emden-leer.de](mailto:christian.spoden@hs-emden-leer.de)



Aron FINK || Goethe-Universität Frankfurt, Arbeitsbereich Pädagogische Psychologie || Theodor-W.-Adorno-Platz 6, D-60629 Frankfurt am Main

[https://www.psychologie.uni-frankfurt.de/73548927/Aron\\_Fink](https://www.psychologie.uni-frankfurt.de/73548927/Aron_Fink)

[a.fink@psych.uni-frankfurt.de](mailto:a.fink@psych.uni-frankfurt.de)



Prof. Dr. Andreas FREY || Goethe-Universität Frankfurt, Arbeitsbereich Pädagogische Psychologie || Theodor-W.-Adorno-Platz 6, D-60629 Frankfurt am Main

[https://www.psychologie.uni-frankfurt.de/73548872/Prof\\_Dr\\_Andreas\\_Frey](https://www.psychologie.uni-frankfurt.de/73548872/Prof_Dr_Andreas_Frey)

[frey@psych.uni-frankfurt.de](mailto:frey@psych.uni-frankfurt.de)



Hanna KÖHLER || Friedrich-Schiller-Universität Jena, Institut für Psychologie || Am Steiger 3, D-07743 Jena

<https://www.klipsy.uni-jena.de/team/hanna-koehler-m-sc>

[hanna.koehler@uni-jena.de](mailto:hanna.koehler@uni-jena.de)



Patrick NAUMANN || Goethe-Universität Frankfurt, Arbeitsbereich Pädagogische Psychologie || Theodor-W.-Adorno-Platz 6, D-60629 Frankfurt am Main

[https://www.psychologie.uni-frankfurt.de/77209044/Patrick\\_Naumann](https://www.psychologie.uni-frankfurt.de/77209044/Patrick_Naumann)

[naumann@psych.uni-frankfurt.de](mailto:naumann@psych.uni-frankfurt.de)

Miriam CLINCY<sup>1</sup>, Karin MELZER, Gunther SCHAAF, Achim EICHHORN & Nathalie VERNÉ (Esslingen)

# Online-Midterms zwischen formativer und summativer Bewertung

## Zusammenfassung

In den MINT-Fächern werden Online-Tests mit automatisierter Rückmeldung häufig für Einstufungstests eingesetzt oder in speziellen Lehrformaten mit einer hauptsächlich formativen Komponente. In diesem Bericht werden zwei Umsetzungsmodelle vorgestellt, die sich aus unterschiedlichen Rahmenbedingungen verschiedener Module ergeben. Unabhängig vom Lehrkonzept werden sie semesterbegleitend jeweils am Ende von Lerneinheiten zu einer summativen Bewertung eingesetzt. Übermöglichkeiten und formative Rückmeldungen ergeben sich für die Studierenden durch den Einsatz einer „Spielwiese“.

## Schlüsselwörter

Online-Tests, MINT, Mathematik, Bewertung

---

<sup>1</sup> E-Mail: [miriam.clincy@hs-esslingen.de](mailto:miriam.clincy@hs-esslingen.de)



## Online midterm tests between summative and formative assessment

### Abstract

In STEM subjects, online tests with automated response systems are often employed for pre-assessments or with a strong formative component tied to teaching formats. This paper presents two models designed for different modules in varying frameworks. Regardless of the teaching concept, tests are implemented during the semester at the end of each course unit. Hence, they allow for summative assessment. Practice and formative feedback are also offered in a “sandbox” environment.

### Keywords

online tests, STEM, mathematics, assessment

## 1 Online-Kurztests als Prüfungsleistung

Die Bologna-Reform des Studiums sieht für den Abschluss jeden Moduls eine Prüfungsleistung vor. Gerade in Mathematik und Physik werden diese in den ersten Semestern meist als Klausuren konzipiert, denn diese Prüfungsform bietet u. a. ein hohes Maß an Vergleichbarkeit (FRÖLICH-STEFFEN, DEN OUDEN & GIESS-MANN, 2019).

Ein Nachteil liegt darin, dass die Studierenden während des Semesters wenig Rückmeldung über ihren Lernprozess erhalten, was wenig nachhaltiges Lernen unmittelbar vor der Prüfung befördert. Gleichzeitig fallen Studierenden oft bereits reproduzierende und einfache Transferaufgaben schwer, d. h. Kompetenzen, die idealerweise bereits während des Semesters erworben werden sollten, um anwendungsorientiertere Aufgaben zu bewältigen.

Um dem entgegenzuwirken, haben sich in den letzten Jahren semesterbegleitende Onlinetests etabliert. Diese können sehr differenziert automatisierte Rückmeldungen geben, werden i. d. R. für Übungsaufgaben eingesetzt *und* geben meist forma-

tive Rückmeldungen während des Lernprozesses. Ihr Einsatz ist häufig an Lehrkonzepte gebunden.

In diesem Beitrag werden verschiedene Modelle von online-basierten Tests in Mathematik vorgestellt, die diese Feedbackmöglichkeiten aufrechterhalten und um summative Bewertungsaspekte erweitern. Gemeinsam ist allen hier vorgestellten Umsetzungen, dass die Tests jeweils *am Ende* von Lerneinheiten stehen und unabhängig vom Lehrkonzept sind. In diesem Beitrag wird der Begriff „Midterm“ für diese Tests verwendet, die auch Eingang in die Modulnote finden.

## 2 Kurztests zwischen summativer und formativer Prüfungsleistung

Die Zugangswege zu den Hochschulen für angewandte Wissenschaften sind vielfältig, entsprechend heterogen sind oft die Voraussetzungen, die die Studierenden mitbringen. Die Mathematik spielt dabei in den ingenieurwissenschaftlichen Fächern eine entscheidende Rolle für den Studienerfolg (BIEHLER, HOCHMUTH & WASSONG, 2011). Untersuchungen zeigen rückläufige Mathematikkenntnisse der Studierenden (CARPENTER & KIRK, 2017), was sich ungünstig auf die Abschlussquoten auswirkt. Dem stehen hochschulische Unterstützungsangebote gegenüber, wie fachspezifische Vorkurse und Tutorien. Sie werden zunehmend durch digitale Tests mit automatisierter Rückmeldung ergänzt, die selbstgesteuertes Lernen erlauben.

Als kompetenzorientierte Prüfungsleistung finden sich diese Tests nur bedingt. GERICK, SOMMER und ZIMMERMANN (2018) verorten semesterbegleitende Aufgaben beispielsweise eher außerhalb des MINT-Bereichs.

Bei der Bewertung lassen sich formative und summative Aspekte unterscheiden. Die Abgrenzung erfolgt hauptsächlich in Bezug auf Zweck und Zeitpunkt (HARLEN & JAMES, 1997). Unter formativer Bewertung wird eher eine Rückmeldung an die Lernenden während des Lehrprozesses verstanden, während summative Bewertungen am Ende des Prozesses zur Erfassung der Gesamtleistung stehen (SCRIVEN, 1967).

Semesterbegleitende formative Aufgaben werden oft als Zulassungsvoraussetzung oder Zusatzpunkte zur Klausur gestaltet, um Studierenden einen Anreiz für kontinuierliches Arbeiten zu geben. Dies kann unterschiedlich ausgestaltet sein: beispielsweise über Aufgaben mit sehr niedriger Bestehensgrenze als „Threshold formative assessment“ (JORDAN, 2014). Anwendungen davon finden sich in Mathematik und Physik oft als Pre-Assessment in Flipped Classroom-Formaten nach Bearbeiten des Materials (vgl. SCHÄFFLE, STANZEL, JUNKER, & ZIMMERMANN, 2017). Möglich ist auch, die Studierenden einen Test oder eine exemplarische Aufgabe beliebig oft bearbeiten zu lassen und den besten zu bewerten (wie beispielsweise bei KNEBUSCH, HEINTZ-CUSCIANNA, & WANDLER, 2021).

Die in diesem Beitrag vorgestellten online-Kurztests in Mathematikvorlesungen der Studieneingangsphase haben einen stärker summativen Charakter. Sie werden bewusst zeitlich jeweils am *Ende* einer Lerneinheit mit einer hohen Bestehensgrenze eingesetzt.

Um den Rückmeldeaspekt zu erhalten, werden Übungen ohne Bewertung eingesetzt, die eine gezielte Vorbereitung auf die Tests ermöglichen. Darüber wird eine Trennung von Lern- und Leistungssituation erreicht (WEINERT, 1999). Für diese Übungen wurde von den Autor\*innen der Begriff „Spielwiese“ als Bezeichnung für eine Entwicklungs- und Testumgebung aus der Softwareentwicklung entlehnt.

### **3 Beschreibung Midterms als Prüfungsform**

An der Hochschule Esslingen sind seit 2019 in verschiedenen Studiengängen und Fakultäten weitgehend unabhängig voneinander Online-Kurztests in Mathematik entstanden. Im Folgenden werden zwei Konzepte verglichen, die von unterschiedlichen Prämissen ausgehen, um möglicherweise Best-Practice-Parameter identifizieren zu können. Dargestellt werden die erstmaligen Durchführungen in dem beschriebenen Setting in verschiedenen Studiengängen derselben Hochschule, jeweils im ersten Semester.

- Variante 1 – IT: Mathematik 1B – Lineare Algebra (Informationstechnik), WiSe 2019/20: Digitale Umsetzung von Übungsblättern, die im Selbststudium von den Studierenden bearbeitet werden<sup>2</sup>.
- Variante 2 – FZ: Mathematik 1 (Fahrzeugtechnik/Fahrzeugsysteme), SoSe 2021: Digitale Umsetzung einer 60-minütigen Midtermklausur.

Die Tests wurden jeweils in zwei oder drei Semestergruppen parallel durchgeführt bei unterschiedlichen Lehrenden mit verschiedenen Lehrformaten, die von der klassischen Präsenzlehre bis zum Online-Flipped Classroom-Format reichten.

Aus den Rahmenbedingungen der Studiengänge sowie dem Einsatzgebiet ergeben sich unterschiedliche Gestaltungsparameter, anhand derer im Folgenden die beiden Varianten klassifiziert werden. Eine Übersicht ist in Tab. 1 gegeben.

### **3.1 Klassifizierung der Gestaltungsparameter**

Zielsetzung der Tests war die Einübung von innermathematischen Rechenfertigkeiten als Voraussetzung für komplexere Aufgabenstellungen.

#### **3.1.1 Einbettung in das Curriculum**

Variante 1 nutzt eine pauschale Midterm-Regelung der Hochschule, wonach maximal 25% der Prüfungsleistung vorgezogen werden können. In Variante 2 ist in der Prüfungsordnung eine Midtermprüfung im Umfang von 60 Minuten während des Semesters vorgesehen. Diese curricular verankerte Dauer wurde auf sechs Tests à 10 Minuten aufgeteilt.

#### **3.1.2 Bestehensgrenzen und Anrechnungsmodus**

Um die Teilleistung erfolgreich abzuschließen, wurde eine Grenze von 75% gewählt. Dies soll die hohe Erwartung an die Beherrschung der Basisfertigkeiten kommunizieren (CHICKERING & GAMSON, 1987).

---

<sup>2</sup> Weiterentwicklung der in Mathematik 1A (Analysis) eingesetzten manuell korrigierten Aufgaben (IOFFE, HELFRICH-SCHKARBANENKO, CLINCY, & KOCH, 2019).

In Variante 1 müssen kumulativ 75 % der erreichbaren Punkte erworben, in Variante 2 müssen fünf von sechs Tests bestanden werden (jeweils mind. 75 % der Punkte).

Bei bestandenem Midterm erwerben die Studierenden 10 % der Klausurpunkte.

### **3.1.3 Übungsmodus**

In beiden Varianten wurden Online-Übungen angeboten. In der Variante 1 war diese „Spielwiese“ umfangreicher als der bewertete Test und musste mit mindestens 75 % bestanden werden, um den bewerteten Test freizuschalten. In Variante 2 handelt es sich um einen identischen Test mit einer Zeitbeschränkung von jeweils 10 Minuten. Durch die Randomisierung der Aufgaben ist die Wahrscheinlichkeit einer passgenauen Vorbereitung gering, dient aber als Anreiz für die Studierenden.

Die Konzeption der „Spielwiese“ dient bei der Implementation auch einer bewertungsfreien Heranführung der Studierenden an die Syntax der Eingabe.

### **3.1.4 Testmodus**

In Variante 1 war der Test nach Freischaltung in einem beliebigen Zeitrahmen von bis zu zwei Wochen einmal durchzuführen. In Variante 2 war die Testdauer pro Test 10 Minuten. Der Test konnte einmal wiederholt werden.

Tab. 1: Übersicht über die Umsetzungsvarianten der Midterms

	Variante 1 (IT)	Variante 2 (FZ)
<b>1. Einbettung in Curriculum</b>	Vorgezogene Prüfungsleistung über Midtermregelung der Hochschule	Midtermklausur (60 Minuten) in der PO vorgesehen
<b>2. Bestehen und Anrechnung</b>	<ul style="list-style-type: none"> <li>• Kumulativ 75 % der Midterm-Punkte</li> <li>• Bonus: 10 % der Klausurpunkte</li> </ul>	<ul style="list-style-type: none"> <li>• Midterm-Tests müssen einzeln bestanden werden, ein Fehlversuch</li> <li>• Bonus: 10 % der Klausurpunkte</li> </ul>
<b>3. Übungsmodus</b>	„Spielwiese“, Aufgaben teilweise identisch mit Test, Freischaltung des Tests bei 75 % der Punkte	„Spielwiese“, identische Aufgaben zum Test und 10 Minuten Zeitbegrenzung
<b>4. Testmodus</b>	<ul style="list-style-type: none"> <li>• Ein Versuch innerhalb zwei Wochen nach Freischaltung</li> <li>• keine Zeitbegrenzung auf Einzeldurchführung</li> </ul>	<ul style="list-style-type: none"> <li>• Zwei Versuche, jeweils eine Woche freigeschaltet</li> <li>• Zeitbegrenzung von 10 Minuten</li> </ul>
<b>5. Anzahl Tests</b>	<ul style="list-style-type: none"> <li>• 5: LGS, Vektoren, Analytische Geometrie, Matrizen, Komplexe Zahlen</li> </ul>	<ul style="list-style-type: none"> <li>• 6: LGS, Funktionen, Vektoren, Differenzialrechnung, Integralrechnung, Matrizen</li> </ul>

## 3.2 Konzeption und technische Umsetzung der Aufgaben

### 3.2.1 Kompetenzorientierte Konzeption

Die Aufgaben sind nahezu ausschließlich innermathematisch. Sie zielen auf Wissen, Verstehen und einfache Anwendungen ab, also die ersten drei Stufen der Bloom'schen Lernzieltaxonomie (BLOOM & KRATHWOHL, 1986). Diese Kompetenzen sind Voraussetzungen für ein erfolgreiches Bearbeiten von Klausuraufgaben, die den Kompetenzerwerb auch in den höheren Stufen Analyse, Synthese und Evaluation von Informationen prüfen.

Für eine Orientierung an Lehrzielen vgl. (SCHINDLER, 2015) gibt es an der Hochschule über die Studiengänge hinweg eine Verständigung über zu vermittelnde

Kompetenzen in Mathematik und Physik auch unabhängig von Lehrkonzepten. Neben dem langjährigen Einsatz eines Studieneingangstests (vgl. z. B. (KURZ & KÄSS, 2019) ist hier vor allem eine Berücksichtigung der Projekte im Übergang Schule-Hochschule wie dem Mindestanforderungskatalog Mathematik (COSH-BW, 2021) zu nennen.

### 3.2.2 Technische Umsetzung der Aufgaben

Alle Midterms wurden als Tests in Moodle angelegt. Die Aufgaben sind vom Aufgabentyp STACK (SANGWIN, 2015) oder programmiert mit einer MATLAB-Moodle-Schnittstelle als Eigenentwicklung (SCHAAF, EICHHORN & HELFRICH-SCHKARBANENKO, 2019). Die Aufgaben sind dabei randomisiert, d. h. bei unterschiedlichen Aufrufen werden jeweils andere, aber vom Schwierigkeitsgrad vergleichbare Zahlenwerte, Funktionen, Aufgabenkontexte etc. angezeigt, sodass auch mehrfaches Üben durch Aufrufen desselben Tests möglich ist. Zwei Beispiele sind in Abb. 1 und Abb. 2 gezeigt.

Gegeben sind die Vektoren

$$\vec{a} = \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix} \text{ und } \vec{b} = \begin{pmatrix} 10 \\ -8 \\ 5 \end{pmatrix}.$$

1. Gesucht ist ein Vektor  $\vec{c}$ , so dass  $\vec{a}$ ,  $\vec{b}$  und  $\vec{c}$  linear abhängig sind:  
 $\vec{c} = [ \dots, \dots, \dots ]$

2. Gesucht ist ein Vektor  $\vec{d}$ , so dass  $\vec{a}$ ,  $\vec{b}$  und  $\vec{d}$  linear unabhängig sind.  
 $\vec{d} = [ \dots, \dots, \dots ]$

Hinweis: Geben Sie Vektoren als  $[x, y, z]$  mit eckigen Klammern an, der Nullvektor ist nicht zulässig.

Abb. 1: Beispiel einer offenen Aufgabenstellung zum Verständnis von linearer (Un-)Abhängigkeit von Vektoren. Die Komponenten der Vektoren werden bei wiederholtem Aufrufen variiert.

Gegeben ist die Funktion  $f(x) := \exp(-x^2)$ . Untersuchen Sie, ob die Funktion gerade oder ungerade ist (oder keins von beiden).  
Bestimmen Sie dazu zunächst

$f(-x) = \exp(-x^2)$

Ihre letzte Antwort wurde folgendermaßen interpretiert:

$\exp(-x^2)$

In Ihrer Antwort wurden die folgenden Variablen gefunden: [x]  
und beantworten Sie die folgenden Fragen

$f(x) = f(-x)$ ? Wahr

Ihre letzte Antwort wurde folgendermaßen interpretiert:

**Wahr**

$f(x) = -f(-x)$ ? Falsch

Ihre letzte Antwort wurde folgendermaßen interpretiert:

**Falsch**

Damit können Sie nun entscheiden:

- ✓ Nicht beantwortet
- Die Funktion ist weder gerade noch ungerade
- Die Funktion ist ungerade
- Die Funktion ist gerade**

Abb. 2: Beispiel einer geschlossenen Aufgabenstellung zur Symmetrie von Funktionen. Bei wiederholtem Aufrufen wird jeweils eine andere Funktion angezeigt.

## 4 Beteiligung und Ergebnisse

Aufgrund der verschiedenen Studiengänge und Zeitpunkte konnte nicht auf eine systematisch ähnliche Testevaluation durch Studierende zurückgegriffen werden. Von einem Vergleich von Klausurergebnissen (u. a. vor und während der Pandemie) wird hier ebenfalls abgesehen. Stattdessen wurden die Bearbeitungsdaten aus Moodle, wie Anzahl und Bewertung der Versuche, ausgewertet und verglichen.

Von den Studierenden wurden die Tests in beiden Varianten schon im ersten Durchlauf gut angenommen. In Vorlesungsevaluationen wurden anekdotisch vor allem die zeitnahe Rückmeldung und der Anreiz zu kontinuierlichem Lernen positiv bewertet.

### 4.1 Teilnahme

Bei Variante 1 haben bis zu 172 Studierende an Spielwiesen und Tests teilgenommen, bei Variante 2 waren es 83 (Spielwiese) bzw. 87 (Test). In Abb. 3 ist über die Anzahl der Tests dargestellt, wie sich die Teilnehmendenzahlen verändern, skaliert auf den Anfangswert. Es zeigt sich, dass beide Varianten die Studierenden in ähnlichem Maße zu kontinuierlicher Bearbeitung animieren.

Bei Variante 2 steht für viele Studierende bereits nach dem 5. Test fest, dass sie die Bonuspunkte erhalten werden, insofern spricht es für die Akzeptanz, dass mit 54% noch mehr als die Hälfte der Studierenden die letzte Spielwiese zur Übung nutzen, obwohl nur noch 22% den Test angehen.

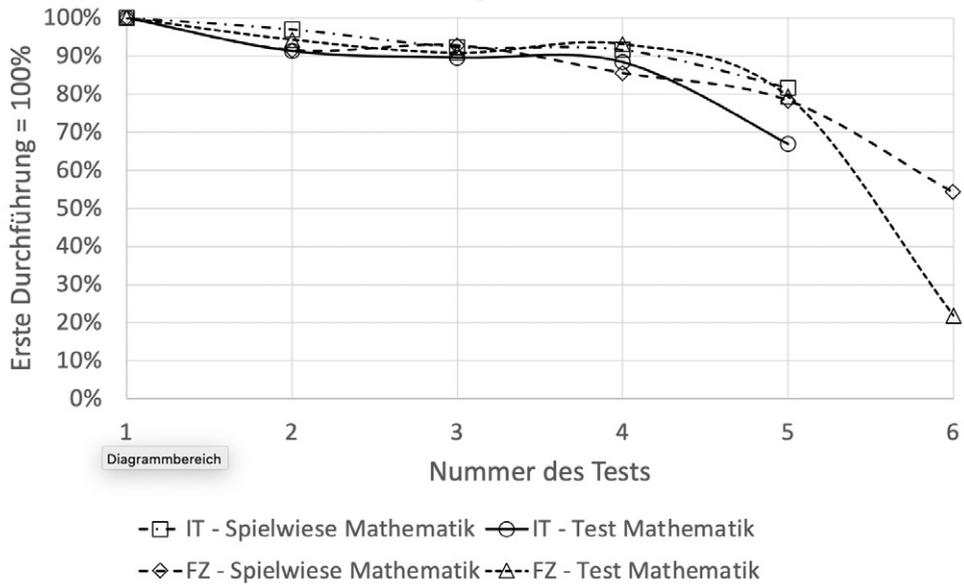


Abb. 3: Zeitlicher Verlauf der prozentualen Teilnehmendenzahlen, skaliert auf die erste Durchführung

## 4.2 Nutzung und Ergebnisse

Bei Variante 1 wurde der Test individuell freigeschaltet, sobald auf der Spielweise 75% der erforderlichen Punkte erreicht waren. Die Anzahl der Versuche pro Teilnehmende/n gemittelt über alle Spielwiesen lag bei  $2,2 \pm 0,2$  Versuchen, die maximale Anzahl der Versuche zwischen 7 und 10.

Bei der zeitbegrenzten Version 2 wurde der Test zu festen Zeitpunkten zur Verfügung gestellt, unabhängig von der Spielweise. Die Anzahl der Versuche auf der Spielweise lag hier bei  $9,9 \pm 3,0$ , einzelne Studierende nutzten die Spielweise bis zu 30-mal und mehr.

Für Variante 2 ist in Abb. 3 dargestellt, wie für die jeweiligen Spielwiesen die erreichte Punktzahl gemittelt über alle Teilnehmenden abnimmt (für die ersten 10 Versuche). Hier ist eine deutliche Konvergenz zur notwendigen Punktzahl von 15 (= 75%) zu erkennen. Der relativ langsame Anstieg könnte andeuten, dass eine befürchtete Abstimmung der Studierenden nicht flächendeckend stattfindet.

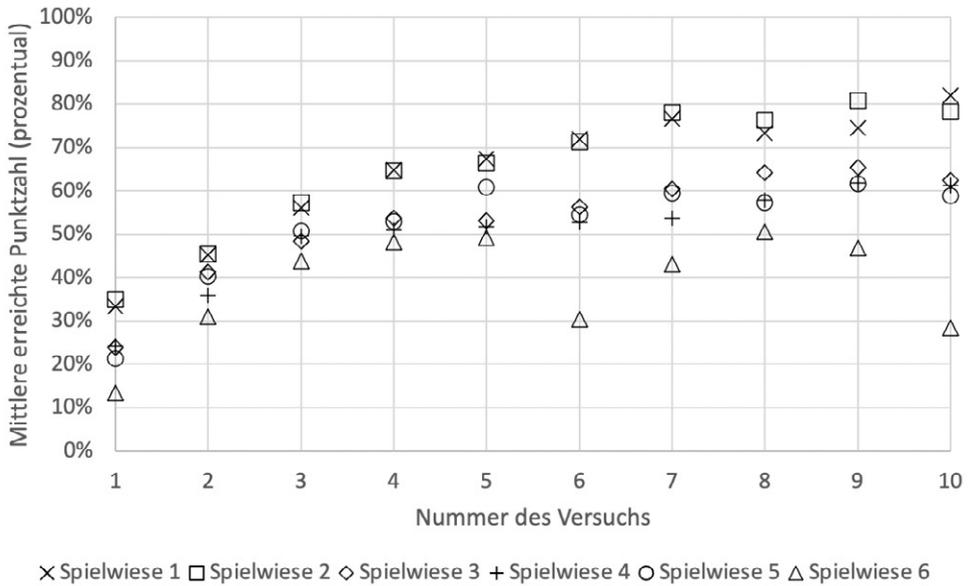


Abb. 4: Mittelwert der Bewertung in Abhängigkeit vom Versuch für die ersten 10 Versuche (Spielwiesen in Variante 2)<sup>3</sup>.

Die Bestehensquote der Midterms bei Variante 1 liegt bei 79,4%, bei Variante 2 bei 63,2%. Gegenwärtig wird untersucht, inwieweit die zeitliche Begrenzung bei Variante 2 angemessen gewählt ist. Eine befürchtete Trivialisierung des Tests durch die passgenaue Vorbereitung hat sich zumindest nicht bestätigt.

<sup>3</sup> Aufgrund der geringen Teilnehmendenzahlen wurden bei der 6. Spielwiese nur die ersten fünf Versuche berücksichtigt.

## 5 Fazit und Ausblick

In dem vorliegenden Beitrag wurden semesterbegleitende Online-Kurztests in Mathematik vorgestellt, die unabhängig vom Lehrkonzept einsetzbar sind. Unterschiedliche Umsetzungsvarianten erlauben eine Anpassung an die Rahmenbedingungen verschiedener Module.

Diese Midterms bestehen aus einer „Spielwiese“ zum Üben mit formativer Rückmeldung und einem Test mit summativer Bewertung und dem Anreiz eines Klausurbonus. Darüber wird eine Trennung von Lern- und Leistungssituation erreicht. Nach den vorliegenden Daten scheint die genaue Ausgestaltung dieser Kombination von bewertungsfreiem Üben und bewerteten Tests für die Bearbeitungsmotivation der Studierenden zweitrangig zu sein. Auswertungen der Antworten, z. B. anhand der Moodle-Statistik, sind in Arbeit.

Auch jenseits der Mathematik ist ein Prüfungskonzept von summativen Online-tests mit formativen Übungen gerade in der Studieneingangsphase vorstellbar. Die Tests stellen eine Teilprüfungsleistung dar, die auf die Förderung von Kompetenzen in den ersten Stufen der Bloom'schen Taxonomie abzielt, also der Förderung von Basis- oder auch fachübergreifenden Schlüsselkompetenzen, die Voraussetzung für die anspruchsvolleren Aufgaben in Modulprüfungen zu Semesterende sind. Denkbar ist damit auch eine stärkere Trennung der Prüfungsleistung in Wissens- und Verständnisaufgaben in semesterbegleitenden Midterms sowie anspruchsvolleren Anwendungsaufgaben in Prüfungen bei Semesterabschluss.

Ein Einsatz von kontextorientierten Aufgaben zur Einübung von Problemlöseheuristiken, die sich nicht allein über mathematische Randomisierung realisieren lassen (beispielsweise in Physik oder technischen Fächern), ist ein weiteres mögliches Anwendungsfeld im MINT-Bereich. Dies ist entsprechend aufwändiger umzusetzen, Umsetzungen hierzu sind bereits in der Erprobung.

## 6 Literaturverzeichnis

- Biehler, R., Hochmuth, R., & Wassong, T.** (2011). Transition von Schule zu Hochschule in der Mathematik: Probleme und Lösungsansätze. *Beiträge zum Mathematikunterricht*, 4. <http://hdl.handle.net/2003/32085>
- Bloom, B. S., & Krathwohl, D. R.** (1986). *Taxonomy of educational objectives. 1: Cognitive domain* (29. print). London: Longman.
- Carpenter, T. P., & Kirk, R. E.** (2017). Are psychology students getting worse at math? *Educational Studies*, 43(3), 282–295. <https://doi.org/10.1080/03055698.2016.1277132>
- Chickering, A. W., & Gamson, Z. F.** (1987). Seven Principles for Good Practice in Undergraduate Education. *AAHE Bulletin*, (March), 3–7. <https://files.eric.ed.gov/fulltext/ED282491.pdf>
- Clarke, A., & Dawson, R.** (1999). *Evaluation research: an introduction to principles, methods, and practice*. London – Thousand Oaks, Calif: SAGE.
- cosh-BW (Hrsg.)**. (2021). *Mindestanforderungskatalog Mathematik Version 3.0*. <http://cosh-mathe.de/wp-content/uploads/2021/12/makV3.0.pdf>
- Frölich-Steffen, S., den Ouden, H., & Gießmann, U. (Hrsg.)**. (2019). Klausuren. In *Kompetenzorientiert prüfen und bewerten an Universitäten* (1. Aufl., S. 101–120). Opladen – Leverkusen: Verlag Barbara Budrich. JSTOR. <https://doi.org/10.2307/j.ctvfc56gd.8>
- Gerick, J., Sommer, A., & Zimmermann, G. (Hrsg.)**. (2018). *Kompetent Prüfungen gestalten: 53 Prüfungsformate für die Hochschullehre*. Münster – New York: Waxmann.
- Harlen, W., & James, M.** (1996, April). *Creating a Positive Impact of Assessment on Learning*. Gehalten auf der Annual Meeting of the American Educational Research Association, New York. <https://files.eric.ed.gov/fulltext/ED397137.pdf>
- Ioffe, O. B., Helfrich-Schkarbanenko, A., Clincy, M., & Koch, J.** (2019). Computergenerierte Mathematiktests in der Studieneingangsphase. *Beiträge zum Mathematikunterricht*, 393. Regensburg.
- Jordan, S.** (2014). Thresholded assessment: Does it work? *ESTeEM Final Report*, 56. <https://www.open.ac.uk/about/teaching-and-learning/esteem/sites/www.open.ac.uk/about/teaching-and-learning/esteem/files/files/ecms/web-content/2014-09-Sally-Jordan-Final-Report-Thresholded-assessment.pdf>

- 
- Knebusch, A., Heintz-Cuscianna, B., & Wandler, M.** (2021). Computerbegleitetes Lernen – digital unterstütztes Selbststudium an der HFT Stuttgart. *e-teaching.org-Artikel*, 13. [https://www.e-teaching.org/etresources/pdf/erfahrungsbericht\\_2021\\_knebusch\\_heintz-cuscianna\\_wandler\\_cbl.pdf](https://www.e-teaching.org/etresources/pdf/erfahrungsbericht_2021_knebusch_heintz-cuscianna_wandler_cbl.pdf)
- Kurz, G., & Käß, H.** (2019). Physikkenntnisse von Studienanfängern des Maschinenbaus – eine Fallstudie an der Hochschule Esslingen. In V. Nordmeier & H. Grötzebauch (Hrsg.), *PhyDid B – Didaktik der Physik – Beiträge zur DPG-Frühjahrstagung Aachen 2019* (S. 189–195). Berlin: PhyDid – DPG. <http://phydid.physik.fu-berlin.de/index.php/phydid-b/article/view/1003>
- Sangwin, C.** (2015). Computer Aided Assessment of Mathematics Using STACK. In S. J. Cho (Hrsg.), *Selected Regular Lectures from the 12th International Congress on Mathematical Education* (S. 695–713). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-319-17187-6\\_39](https://doi.org/10.1007/978-3-319-17187-6_39)
- Schaaf, G., Eichhorn, A., & Helfrich-Schkarbanenko, A.** (2019, Juli 2). *Digitaler Rückenwind für Mathematikveranstaltungen – Aufgabengeneratoren*. Gehalten auf der Poster für die „MATLAB EXPO 2019“. <https://www2.hs-esslingen.de/~aeich/digitalerrueckenwind/aufgabengeneratoren/>
- Schäffle, C., Stanzel, S., Junker, E., & Zimmermann, M.** (2017, Juni). Teil 1: Aktivierende und konzeptorientierte Lehrmethoden. *Didaktik-Nachrichten*, 3–12. [https://diz-bayern.de/DiNa/06\\_2017](https://diz-bayern.de/DiNa/06_2017)
- Schindler, C.** (2015, November 18). *Herausforderung Prüfen: Eine fallbasierte Untersuchung der Prüfungspraxis von Hochschullehrenden im Rahmen eines Qualitätsentwicklungsprogramms*. [https://www.researchgate.net/publication/305800453\\_Herausforderung\\_Pruefen\\_Eine\\_fallbasierte\\_Untersuchung\\_der\\_Pruefungspraxis\\_von\\_Hochschullehrenden\\_im\\_Rahmen\\_eines\\_Qualitaetsentwicklungsprogramms](https://www.researchgate.net/publication/305800453_Herausforderung_Pruefen_Eine_fallbasierte_Untersuchung_der_Pruefungspraxis_von_Hochschullehrenden_im_Rahmen_eines_Qualitaetsentwicklungsprogramms)
- Scriven, M.** (1967). The methodology of evaluation. In R. W. Tyler (Hrsg.), *Perspectives of curriculum evaluation*. Chicago: Rand McNally education series.
- Weinert, F. E.** (1999). Die fünf Irrtümer der Schulreformer. *Psychologie heute*, 26(7), 28–34.

## Autor/innen



Prof. Dr. Miriam CLINCY || Hochschule Esslingen, Fakultät für Angewandte Naturwissenschaften, Energie- und Gebäudetechnik || Kanalstr. 33, D-73728 Esslingen

<https://www.hs-esslingen.de/personen/miriam-clincy/>

[miriam.clincy@hs-esslingen.de](mailto:miriam.clincy@hs-esslingen.de)



Prof. Dr. Karin MELZER || Hochschule Esslingen, Fakultät für Informatik und Informationstechnik || Flandernstr.101, D-73732 Esslingen

<https://www.hs-esslingen.de/personen/karin-melzer/>

[karin.melzer@hs-esslingen.de](mailto:karin.melzer@hs-esslingen.de)



Prof. Dr. Gunther SCHAAF || Hochschule Esslingen, Fakultät für Mobilität und Technik || Kanalstr. 33, D-73728 Esslingen

<https://www.hs-esslingen.de/personen/gunther-schaaf/>

[gunther.schaaf@hs-esslingen.de](mailto:gunther.schaaf@hs-esslingen.de)



Dipl.-Inf. Achim EICHHORN || Hochschule Esslingen, ZWE Studiengang und Grundstudium || Kanalstr. 33, D-73728 Esslingen

<https://www.hs-esslingen.de/personen/achim-eichhorn/>

[achim.eichhorn@hs-esslingen.de](mailto:achim.eichhorn@hs-esslingen.de)



Nathalie VERNÉ || Hochschule Esslingen, Referat Lehre und Weiterbildung || Flandernstraße 101, D-73732 Esslingen

[nathalie.verne@hs-esslingen.de](mailto:nathalie.verne@hs-esslingen.de)

Sandra NIEDERMEIER<sup>1</sup>, Michael SAILER, Lukas REMBÖCK & Matthias STADLER (Kempten/München)

# Ich packe meinen Koffer: Unterlagen und Performanz bei Open-Book-Klausuren

## Zusammenfassung

Die Nachfrage nach zukunftsgerichteten, problemlösungsorientierten Prüfungsformaten für die Hochschullehre stieg nicht zuletzt aufgrund der COVID-19-Pandemie. Open-Book-Prüfungen, bei denen Unterlagen mitgebracht werden, werden seit Langem als alternatives Prüfungsformat vorgeschlagen. Der vorliegende Beitrag beleuchtet in einer Studie mit 113 Teilnehmenden, ob die Menge und Art mitgebrachter Unterlagen einen Einfluss auf die Performanz in Multiple-Choice-Fragen zu konzeptuellem und prozeduralem Wissen aufweisen. Einen signifikanten, positiven Zusammenhang in Bezug auf die Performanz in beiden Wissensformen wies die Menge an mitgebrachten kommentierten Vorlesungsfolien auf. Diese Befunde werden in Bezug zur ICAP-Theorie gesetzt, die zugrundeliegende Lernaktivitäten beleuchtet.

## Schlüsselwörter

Open-Book, prozedurales Wissen, konzeptuelles Wissen, Prüfungsformen

---

<sup>1</sup> E-Mail: [sandra.niedermeier@hs-kempten.de](mailto:sandra.niedermeier@hs-kempten.de)



## **I pack my bag and take with me... The relationship between documentation brought by students and performance in open-book exams**

### **Abstract**

Due to the COVID-19 pandemic, there is an increasing demand for forward-thinking and problem-solving-oriented exam formats in higher education. Open-book exams, in which learning materials are brought in, have long been proposed as an alternative exam format. This study with 113 participants investigates whether the amount and type of learning materials brought in have an impact on performance in multiple-choice questions on conceptual and procedural knowledge. The amount of brought in annotated lecture slides showed a significant, positive correlation with performance in both forms of knowledge. These findings are related to ICAP theory, which highlights underlying learning activities.

### **Keywords**

Open-Book, procedural knowledge, conceptual knowledge, testing

## **1 Einleitung**

Die Covid-19-Pandemie und die mit ihr einhergehenden Reglementierungen zur Einschränkung der Virusverbreitung hatten erhebliche Auswirkungen auf die alltägliche Praxis an Hochschulen. So fanden nicht nur die Lehrveranstaltungen, sondern auch Prüfungen überwiegend online statt (CRAWFORD et al., 2020). Als eine attraktive und insbesondere eine problemlösungsorientierte Prüfungsform erscheint dabei die Open-Book-Prüfungsvariante, in der Studierenden selbstgewählte Hilfsmittel zur Verfügung stehen.

Diese Prüfungsform gilt als eine realitätsnähere Prüfung als Closed-Book-Prüfungen, in welchen keine Hilfsmittel zur Verfügung stehen, „da sowohl in der Arbeitswelt als auch im wissenschaftlichen Prozess i. d. R. Hilfsmittel erlaubt sind, um Ergebnisse zu erarbeiten“ (HORN & SCHMEES, 2020, S. 11). Vor diesem Hin-

tergrund stellt sich die Frage, welche mitzubringenden Materialien am ehesten zu empfehlen sind, um eine solche Prüfung erfolgreich zu bestreiten. So sind zum Beispiel nicht nur Mitschriften, Vorlesungsfolien oder Lehrbücher erlaubt, sondern auch bereits bearbeitete Aufgaben, welche von den Dozierenden im Vorfeld zur Verfügung gestellt wurden. Inwiefern diese Unterlagen mit dem Prüfungsergebnis zusammenhängen, ist unklar. Zunächst sollen daher im Folgenden das Prüfungsformat Open-Book und mit diesem in Zusammenhang stehende Studien genauer betrachtet werden.

## 2 Prüfung im Open-Book-Format

Als Closed-Book werden klassischerweise alle schriftlichen Prüfungen, wie u. a. Tests, Klausuren etc. bezeichnet, die ohne (oder unter nur sehr eingeschränkter) Hinzunahme von Hilfsmitteln und Unterlagen durchgeführt werden. Ihnen stehen Open-Book-Prüfungen gegenüber. Als Open-Book-Prüfungen werden Prüfungen bezeichnet, zu deren Bearbeitung es Studierenden erlaubt ist, Hilfsmittel zu verwenden. Hilfsmittel sind technische und fachbezogene Hilfen zur Bearbeitung einer Prüfung, die idealerweise vorab definiert werden (BATZ-FINKBOHNER, BESNER & GERSTNER, 2021). Fachbezogene Hilfsmittel sind z. B. Skripte, Formelsammlungen, Rechtsvorschriften, Mitschriften oder Lehrbücher. Technische Hilfsmittel sind z. B. Taschenrechner, Konstruktionswerkzeuge oder Computer bzw. bestimmte Software. Üblicherweise sind Kommunikationsmittel wie z. B. ein Smartphone zum Austausch mit anderen Studierenden nicht erlaubt. (Fern-)Prüfungen im Open-Book-Format können in vielen Varianten in beaufsichtigter sowie unbeaufsichtigter Form stattfinden (BATZ-FINKBOHNER, BESNER & GERSTNER, 2021).

Open-Book-Prüfungen (d. h. Prüfungen, bei denen die Studierenden alle verfügbaren Ressourcen nutzen können, ohne sich mit anderen Studierenden abzusprechen) werden seit einigen Jahren für Prüfungen in Online-Lernumgebungen vorgeschlagen (RAKES, 2008). Eine Open-Book-Prüfung fordert nach LEHNER (2018) eine andere Art der Prüfungsvorbereitung ein als eine Closed-Book-Prüfung. Diese ist zumindest dem Versuch ausgesetzt, Reproduktionsleistungen entsprechend einzufordern (LEHNER, 2018).

Studien haben gezeigt, dass Open-Book-Prüfungen das Lernen und die Motivation der Studierenden verbessern können (z. B. GREEN et al., 2016). GHARIB et al. (2012) lieferten Evidenz dafür, dass Studierende bei Prüfungen im Open-Book-Format, im Gegensatz zum Closed-Book-Format, eine niedrigere wahrgenommene Angst aufweisen. In Widerspruch hierzu steht zum Beispiel die etwas ältere Studie von IOANNIDOU (1997), bei welcher die Performanz, also somit das erreichte Lernergebnis, zwischen Open-Book- und Closed-Book-Prüfungen verglichen wurde. Die Tests waren so gestaltet, dass sie kritisches Denken und Denken höherer Ordnung, wozu beispielsweise das Bewusstsein der Steuerung von Lernprozessen zählt, beanspruchen sollten, was durch ein aktives und selbstgesteuertes Lernen unterstützt werden sollte. Dabei zeigte sich in den Ergebnissen der zwei Prüfungsformate kein signifikanter Unterschied in der erzielten Leistung. Eine schlechtere Performanz in Klausuren wird vermutet, da Studierende sich bei Open-Book-Prüfungen mehr auf ihre Unterlagen verlassen und sich in der Vorbereitung weniger vertieft mit den Inhalten beschäftigen (RUMMER et al., 2019).

Einen Zusammenhang zwischen den mitgebrachten Unterlagen im Open-Book-Format und dem Prüfungsergebnis könnte unter Umständen die konkrete Organisation und Elaboration der mitzubringenden Unterlagen aufweisen. Lernaktivitäten, wie das Schreiben von Zusammenfassungen, stehen im Vordergrund und wirken sich möglicherweise auf die Prüfungsleistung aus. Dieser Aspekt ist bisher noch unzureichend untersucht (CAPALDI, 2019). Generell fallen Befunde verschiedener Studien zur Performanz in Open-Book-Tests recht unterschiedlich aus. Einen strukturierten Überblick hierzu geben DURNING et al. (2016).

Im Fokus dieses Artikels steht eine Open-Book-Klausur. Eine solche wird als eine schriftliche Aufsichtsprüfung unter Zulassung inhaltlicher Hilfsmittel verstanden, die mit enger zeitlicher Begrenzung (meist 60–120 Minuten) und gleichzeitig von allen Teilnehmenden schriftlich bearbeitet wird. Hierbei gilt es, Fragen so zu konzipieren, um die reine Reproduktionsleistung oder reines „Abschreiben“ zu vermeiden. Damit werden verschiedene Wissensarten angesprochen.

## 2.1 Wissensarten und Fragekonstruktion im Prüfungskontext

Um verschiedene Wissensarten greifbar zu machen und für den folgenden Kontext vereinfacht darzustellen, bietet sich eine Differenzierung in zwei Arten an, nämlich konzeptuelles und prozedurales Wissen.

Konzeptuelles Wissen („conceptual knowledge“) bedeutet „wissen, was“ etwas ist (CARPENTER, 1986). Im Gegensatz zum konzeptuellen Wissen steht das prozedurale Wissen, welches als eine Art „Handlungswissen“ verstanden wird. Es wird gern mit dem Schlagwort „wissen, wie“ versehen und ermöglicht das Anwenden (einfacher) Prozeduren und Schemata.

Fragen zum prozeduralen Wissen umfassen in der Regel ein Szenario (oder eine Situation), welches personen-, inhalts- und kontextsensitiv ist (HAAPASALO & KADIJEVICH, 2000). Konzeptuelle Testitems erfordern ein echtes Verständnis der zugrunde liegenden Inhalte. Gelingt es erfolgreich derartige Fragen zu konzipieren, so zeigen ENGELBRECHT et al. (2005) in ihrer Studie, bei welcher 235 Erstsemester-Studierende in einem Algebra-Einführungskurs mithilfe eines Multiple-Choice-Tests geprüft wurden, dass kein Unterschied in der dargebrachten Leistung besteht. Des Weiteren hatten Studierende mehr Vertrauen in ihre Fähigkeiten, konzeptuelle Probleme zu lösen als prozedurale. Ebenso konnten sie klären, dass die Annahme, dass Missverständnisse eher bei konzeptuellen als prozeduralen mathematischen Thematiken auftreten, nicht zutrifft.

Im Kontext von Open-Book-Klausuren stellt sich nun die Frage, ob mitgebrachte Unterlagen einen Einfluss auf die Performanz in Prüfungssituationen zu konzeptuellem und prozeduralem Wissen haben. Zusätzlich stellt sich die Frage, ob bei mitgebrachten Unterlagen verschiedene Grade der Lernaktivität zugrunde liegen, welchen im Vergleich zueinander unterschiedlich stark ausgeprägte Lernprozesse zuzuschreiben sind. Als theoretische Grundlage zur Erklärung verschiedener Grade der Lernaktivität dient die ICAP-Theorie nach CHI (2009).

## 2.3 Lernaktivitäten unter der Perspektive der ICAP Theorie

Zur Unterscheidung und genaueren Definition bestimmter Lernaktivitäten von Lernenden entwickelte CHI (2009) einen theoretischen Rahmen zur genaueren Bestimmung der zugrunde liegenden Lernaktivitätsformen und -grade, um sich den kognitiven Prozessen der Lernenden anzunähern. Zur genaueren Differenzierung von Lernaktivitäten werden die Begriffe interaktiv (interactive), konstruktiv (constructive), aktiv (active) und passiv (passive) herangezogen. Sie bilden zusammen das Akronym ICAP. Zwischen diesen verschiedenen Graden besteht eine hierarchische Struktur.

Die erste Stufe in dieser Taxonomie bilden die passiven Lernaktivitäten, wie das Zuhören in einer Vorlesung, lautes oder leises Lesen eines Textes oder das Ansehen eines Videos, jedoch ohne sonstige Handlungen zu tätigen (CHI & WYLIE, 2014). Bei einer Lernaktivität aktiv zu sein, bedeutet, während des Lernens physisch aktiv zu sein. Beispiele hierfür sind unter anderem mit einem Marker hervorheben, das Paraphrasieren oder das Wiederholen, wodurch Vorwissen aktiviert wird und neue Informationen verarbeitet und in das bestehende Wissen integriert werden. So ergeben sich zum Beispiel durch das Unterstreichen von Textpassagen keine neuen Inhalte, sondern dies stellt lediglich eine Manipulation des ursprünglichen Lernmaterials dar. Konstruktive Lernaktivitäten zeichnen sich durch die Generierung von eigenen Inhalten in Form von Artefakten aus, die über das Lernmaterial hinausgehen, wie zum Beispiel das Erstellen einer Concept Map. Als interaktiv wird eine Lernaktivität bezeichnet, wenn mehrere Lernende konstruktive Lernaktivitäten ausführen und die Beiträge der einzelnen Lernenden aufeinander aufbauen.

Aus dieser Differenzierung der Grade von Lernaktivitäten stellt die ICAP-Theorie eine Erklärungsgrundlage für die Performanz in Aussicht und macht drei grundlegende Annahmen (CHI et al., 2018):

1. Das offenkundige Verhalten sowie die Lernprodukte der Lernenden stellen ein Maß für deren zugrunde liegende kognitive Prozesse dar.
2. Das offenkundige Verhalten der Lernenden und deren Lernprodukte deuten Unterschiede in Prozessen kognitiver Wissensänderung an.
3. Das beobachtbare Verhalten entspricht nicht exakt den sich daraus herleitenden Prozessen der Wissensänderung, aber es ist gut genug, um allgemeine Aussagen treffen zu können.

Die ICAP-Theorie bietet einen Rahmen an, Prüfungen zu betrachten, bei denen sich Studierende unterstützende Materialien mitbringen können. Open-Book-Klausuren bieten die Möglichkeit, Unterlagen mitzubringen, welche potenziell hilfreich für die Prüfungsleistung sind. Richtet man den Blick jedoch auf die Entstehung der Unterlagen, die mitgebracht werden, rückt der Lernprozess in den Fokus. So ist zu vermuten, dass die Basis hinter der Entstehung der Unterlagen die eben genannten Lernprozesse sind. Folien ohne Anmerkungen mitzubringen, wäre nach dieser Auffassung passiv, da diese – im besten Fall – nur gelesen wurden. Eine tiefere Verarbeitung der Inhalte der Vorlesung führt möglicherweise zu höherer Performanz in der Klausur. Eine solche tiefere Verarbeitung der Inhalte zeigt sich darin, dass die Studierenden sich im Vorfeld mit den Unterlagen, die sie mitbringen, auseinandersetzen. Beispielsweise indem sie sich eine Zusammenfassung schreiben (konstruktiv) oder Folien kommentieren (aktiv oder konstruktiv).

Die nun in diesem Zusammenhang im Folgenden vorgestellte Studie untersucht, ob mitgebrachte Unterlagen in einer Hochschulprüfung einen Einfluss auf die Performanz bei Multiple-Choice-Fragen zu konzeptuellem und prozeduralem Wissen haben.

### **3 Studie**

Ziel der vorliegenden Studie war es, den Zusammenhang zwischen mitgebrachten Unterlagen und der Performanz in einer Open-Book-Klausur zu untersuchen. Von besonderem Interesse war, welche Unterlagen günstig für das Mitbringen in eine Klausur sind. Hierbei wurde angenommen, dass ein positiver Zusammenhang zwischen den prozentual mitgebrachten Vorlesungsfolien, den prozentual zusätzlich kommentierten Vorlesungsfolien, der Seitenanzahl der mitgebrachten eigenen Aufzeichnungen sowie den prozentual mitgebrachten Trainingsaufgaben und der Performanz in den Fragen zu konzeptuellem Wissen besteht. Dasselbe wird in Bezug auf die Performanz in den Fragen zum prozeduralen Wissen angenommen.

### 3.1 Hypothesen

Die Studie beschäftigt sich damit, ob mitgebrachte Unterlagen in einer Open-Book-Prüfung mit der Performanz in Multiple-Choice Fragen zu konzeptuellem und prozeduralem Wissen zusammenhängen. Aus dieser Fragestellung resultieren folgende Hypothesen:

Hypothese 1: Es besteht ein positiver Zusammenhang zwischen mitgebrachten Vorlesungsfolien und der Performanz in Multiple-Choice-Fragen zu konzeptuellem Wissen.

Hypothese 2: Es besteht ein positiver Zusammenhang zwischen mitgebrachten Vorlesungsfolien und der Performanz in Multiple-Choice-Fragen zu prozeduralem Wissen.

Hypothese 3: Es besteht ein positiver Zusammenhang zwischen selbst kommentierten Vorlesungsfolien und der Performanz in Multiple-Choice-Fragen zu konzeptuellem Wissen.

Hypothese 4: Es besteht ein positiver Zusammenhang zwischen selbst kommentierten Vorlesungsfolien und der Performanz in Multiple-Choice-Fragen zu prozeduralem Wissen.

Hypothese 5: Es besteht ein positiver Zusammenhang zwischen der Seitenanzahl mitgebrachter eigener Aufzeichnungen und der Performanz in Multiple-Choice-Fragen zu konzeptuellem Wissen.

Hypothese 6: Es besteht ein positiver Zusammenhang zwischen der Seitenanzahl mitgebrachter eigener Aufzeichnungen und der Performanz in Multiple-Choice-Fragen zu prozeduralem Wissen.

Hypothese 7: Es besteht ein positiver Zusammenhang zwischen der Anzahl mitgebrachter Übungsaufgaben und der Performanz in Multiple-Choice-Fragen zu konzeptuellem Wissen.

Hypothese 8: Es besteht ein positiver Zusammenhang zwischen der Anzahl mitgebrachter Trainingsaufgaben und der Performanz in Multiple-Choice-Fragen zu prozeduralem Wissen.

Neben diesen Hypothesen wird auch die Annahme geprüft, ob sich die Effektstärken der Zusammenhänge zwischen Lernaktivität und Wissen nach der Qualität der Lernaktivität unterscheiden.

## **3.2 Methode**

### **3.2.1 Stichprobe und Ablauf**

Die Befragung fand im Vorlesungssaal vor dem Schreiben der Klausur zu empirischen Forschungsmethoden statt. Zunächst wurde den Studierenden der Fragebogen ausgeteilt. Für die Beantwortung hatten sie fünf Minuten Zeit. Anschließend wurde dieser wieder eingesammelt und der Klausur-Fragebogen ausgeteilt, für den wiederum 80 Minuten Zeit zur Bearbeitung angesetzt waren.

Befragt wurden ursprünglich 152 Studierende, welche an der Klausur zu empirischen Forschungsmethoden aus dem Studienfach Pädagogik/Bildungswissenschaften teilnahmen. Nach Sichtung der Daten reduzierte sich die Größe der Stichprobe auf 113 ( $w = 93$ ,  $m = 20$ ). Der Ausschluss der 39 Probandinnen und Probanden erfolgte aufgrund dessen, dass entweder keine Angaben auf dem Fragebogen getätigt wurden oder aber bei der Frage nach Anzahl der Seiten mitgebrachter eigener Aufzeichnungen eine unrealistisch (bzw. unmöglich) hohe Anzahl angegeben wurde.

### **3.2.2 Maße**

Um die mitgebrachten Materialien zu erfassen, wurden die Studierenden gebeten einzuschätzen, wie viel Prozent der Vorlesungsfolien, zusätzlich kommentierter Vorlesungsfolien und Trainingsaufgaben sowie eigene Aufzeichnungen in Seiten sie mitgebracht hatten. Diese Angaben wurden a priori den ICAP-Kategorien Passiv (unkommentierte Vorlesungsfolien), Aktiv (Trainingsaufgaben), Konstruktiv (eigene Aufzeichnungen und kommentierte Vorlesungsfolien) zugeordnet.

Die Performanzen in den Items zu konzeptuellem und prozeduralem Wissen stellen die Kriterien dieser Untersuchung dar. Diese wurde mithilfe der Multiple-Choice-Fragen einer Klausur in den empirischen Forschungsmethoden erfasst. Hierbei sind je 15 der 30 Fragen der konzeptuellen beziehungsweise der prozeduralen Wissensabfrage zuzuordnen. Das konkrete Aufgabenformat ist eine Single-Choice-Klau-

sur, was bedeutet, dass bei allen Fragen, welchen stets vier Antwortmöglichkeiten zugeordnet sind, immer nur genau eine Antwortmöglichkeit richtig sein kann. So kann jede Aufgabe entweder den Wert 1, Frage korrekt beantwortet, oder 0, Frage inkorrekt beantwortet, annehmen.

Zur besseren Unterscheidung der beiden Wissensarten folgen hier zunächst noch Beispiele:

Ein Item zur Abfrage von konzeptuellem Wissen lautet zum Beispiel:

Sie erfassen in einer Studie die Aktivitäten von Proband\*innen in einer online Lernumgebung. Um welche Form der Datenerhebung handelt es sich hierbei?

- a. Eine Beobachtung
- b. Eine Dokumentenanalyse
- c. Ein psychologischer Test
- d. Eine physiologische Messung

Abb. 1: Beispiel-Item konzeptuelle Wissensart

Die Reliabilität der gesamten Skala zu konzeptuellem Wissen wurde mithilfe des McDonald's Omega bestimmt. Der Wert beträgt  $\omega = .62$ .

Eine Frage der prozeduralen Wissensskala ist zum Beispiel:

In einer Vorstudie zu einer Untersuchung erhalten Sie folgenden Output für ein eigens entwickeltes Messinstrument zur Erfassung der Ungewissheitstoleranz aus Ihrem Statistikprogramm: Items: 3; Cronbachs Alpha: 0.42. Wie gehen Sie mit diesem Ergebnis im Hinblick auf die Hauptuntersuchung um?

- a. Items mit mittlerer Schwierigkeit sollten ausgeschlossen werden.
- b. Eine Itemanalyse ist nicht angebracht.
- c. Die Skala kann so verwendet werden.
- d. Es sollte eine Itemanalyse hinsichtlich Schwierigkeit, Trennschärfe und Verteilung durchgeführt und darauf basierend weitere Items mit ausreichender Trennschärfe erstellt werden.

Abb. 2: Beispiel-Item prozedurale Wissensart

Die Reliabilität der gesamten Skala zu prozeduralem Wissen wurde mithilfe des McDonald's Omega bestimmt. Der Wert beträgt  $\omega = .63$ .

### 3.3 Ergebnisse

Die Performanzen in den Items zu konzeptuellem und prozeduralem Wissen stellen die Kriterien dieser Untersuchung dar. Deren Ergebnisse werden nun vorgestellt. Im Folgenden werden zunächst noch die deskriptiven Befunde veranschaulicht.

#### 3.3.1 Deskriptive Befunde

Der folgenden Tabelle 1 können die Mittelwerte, Standardabweichungen sowie Minimum- und Maximum-Werte und bivariaten Korrelationen aller in dieser Arbeit untersuchten Variablen entnommen werden. Dabei ist nochmals darauf hinzuweisen, dass die Werte der mitgebrachten Vorlesungsfolien, der zusätzlich kommentierten Vorlesungsfolien sowie der mitgebrachten Trainingsaufgaben, welche alle eine verhältnismäßig hohe Standardabweichung aufweisen, prozentuale Angaben an Seitenzahlen darstellen. Bei den mitgebrachten eigenen Aufzeichnungen handelt es sich um konkrete Angaben von Seitenanzahlen. Die Daten beziehen sich auf die  $N = 113$  Personen, die an der Befragung teilnahmen und in die Analyse miteinbezogen wurden.

Bivariat besteht ein kleiner, positiver Zusammenhang zwischen der Anzahl mitgebrachter Vorlesungsfolien und der Performanz beim konzeptuellen Wissen ( $r = .21$ ,  $p = .028$ ) sowie bei der Performanz im gesamten Multiple-Choice-Test ( $r = .20$ ,  $p = .039$ ). Dies ist ebenfalls bei der Anzahl der zusätzlich kommentierten Vorlesungsfolien und der Performanz sowohl beim prozeduralen ( $r = .23$ ,  $p = .016$ ) als auch beim konzeptuellen Wissen ( $r = .27$ ,  $p = .006$ ) sowie bei der Performanz im gesamten Multiple-Choice-Test ( $r = .28$ ,  $p = .004$ ) der Fall.

Tabelle 1: Deskriptive Befunde und Interkorrelationen aller Variablen

	M	SD	Range	1	2	3	4	5	6
1. Mitgebrachte Vorlesungsfolien (in %)	53.7	41.4	0-100	-					
2. Zusätzlich kommentierte Vorlesungsfolien (in %)	30.4	33.7	0-100	.36**	-				
3. Mitgebrachte eigene Aufzeichnungen (in Seiten)	14.2	11.0	0-45	.04	.21*	-			
4. Mitgebrachte Trainingsaufgaben (in %)	31.4	49.0	0-100	.13	.17	.17	-		
5. Performanz in prozeduralen Wissensfragen	9.17	2.65	3-14	.15	.23*	.12	-.02	-	
6. Performanz in konzeptuellen Wissensfragen	7.42	2.83	1-13	.21*	.27**	.11	-.04	.59**	-
7. Performanz im gesamten MC-Test	16.6	4.89	4-27	.20*	.28**	.13	-.04	.88**	.90**

M = Mittelwert, SD = Standardabweichung, \* inidziert  $p < .050$ , \*\* inidziert  $p < .010$

### 3.3.2 Befunde zum konzeptuellen Wissen

Zur Bestimmung der relativen Bedeutung von mitgebrachten Vorlesungsfolien, zusätzlich selbst kommentierten Vorlesungsfolien, mitgebrachten eigenen Aufzeichnungen und mitgebrachten Trainingsaufgaben auf die Performanz in den Fragen zu konzeptuellem Wissen wurde eine multiple Regression errechnet. Die Inspektion eines Residuenplots zeigte, dass Homoskedastizität der Residuen gegeben ist. Die Güte des Regressionsmodells hängt also nicht von der Performanz der Studierenden ab. Ebenso liegt keine substantielle Multikollinearität der Prädiktoren vor ( $VIF_{\text{Max}} = 1.21$ ), mehrere Prädiktoren korrelieren somit nicht miteinander. Es konnte inferenzstatistisch nur abgesichert werden, dass zusätzlich selbst kommentierte Vorlesungsfolien mit der Performanz in den konzeptuellen Wissensfragen zusammenhängen

( $\beta = .23$ ;  $p = .018$ ). Die prozentual mitgebrachten Vorlesungsfolien ( $\beta = .16$ ;  $p = .056$ ), die Anzahl mitgebrachter eigener Aufzeichnungen ( $\beta = .04$ ;  $p = .348$ ) und die prozentual mitgebrachten Trainingsaufgaben ( $\beta = -.08$ ;  $p = .201$ ) erwiesen sich als keine signifikanten Prädiktoren für die Performanz in den Fragen zu konzeptuellem Wissen. Das Gesamtmodell ist hierbei signifikant ( $F(4, 97) = 2.86$ ,  $p = .007$ ), wobei die Regression 11 % der Varianz in der Performanz im konzeptuellen Wissen aufklärt ( $R^2 = .105$ ). Die entsprechenden Kennwerte hierzu sind der Tabelle 2 zu entnehmen. Von den Hypothesen 1, 3, 5 und 7 wird damit nur Hypothese 3 unterstützt, während die anderen verworfen werden müssen.

Ein explorativer Vergleich ergab, dass die Effekte der Materialien, die konstruktiven Lernaktivitäten zugeordnet wurden, nicht konsistent höher waren als die Effekte der Materialien, die niedrigeren Lernaktivitäten zugeordnet wurden.

Tabelle 2: Modellkoeffizienten Konzeptuelles Wissen

Prädiktoren	<i>b</i>	<i>SE</i>	<i>t</i>	<i>p</i>	$\beta$
Mitgebrachte Vorlesungsfolien (in %)	0.01	0.01	1.58	.056	.16
Zusätzlich kommentierte Vorlesungsfolien (in %)	0.02	0.01	2.14	.018	.23
Mitgebrachte eigener Aufzeichnungen (in Seiten)	0.01	0.03	0.39	.348	.04
Mitgebrachte Trainingsaufgaben (in %)	-0.01	0.01	-0.84	.201	-.08

### 3.3.3 Befunde zum prozeduralen Wissen

Zur Bestimmung der relativen Bedeutung von mitgebrachten Vorlesungsfolien, zusätzlich selbst kommentierten Vorlesungsfolien, mitgebrachten eigenen Aufzeichnungen und mitgebrachten Trainingsaufgaben für die Performanz in den Fragen zu prozeduralem Wissen wurde ebenfalls eine multiple Regression gerechnet. Die entsprechenden Kennwerte hierzu sind der Tabelle 3 zu entnehmen.

Tabelle 3: Modellkoeffizienten Prozedurales Wissen

Prädiktoren	<i>b</i>	<i>SE</i>	<i>t</i>	<i>p</i>	$\beta$
Mitgebrachte Vorlesungsfolien (in %)	0.01	0.01	1.09	.140	.12
Zusätzlich kommentierte Vorlesungsfolien (in %)	0.02	0.01	1.93	.029	.21
Mitgebrachte eigene Aufzeichnungen (in Seiten)	0.02	0.02	0.87	.193	.09
Mitgebrachte Trainingsaufgaben (in %)	-0.01	0.01	-1.01	.158	-.10

Wie bei den Ergebnissen zum prozeduralen Wissen wies die Inspektion eines Residuenplots auf Homoskedastizität der Residuen hin. Es lag keine substantielle Multikollinearität der Prädiktoren vor ( $VIF_{\text{Max}} = 1.21$ ), mehrere Prädiktoren korrelierten also auch hier nicht miteinander. Wie beim prozeduralen Wissen erwiesen sich nur die zusätzlich selbst kommentierten Vorlesungsfolien als signifikanter Prädiktor der Performanz in den konzeptuellen Wissensfragen ( $\beta = .21$ ;  $p = .029$ ). Die mitgebrachten Vorlesungsfolien ( $\beta = .12$ ;  $p = .140$ ), die Anzahl mitgebrachter eigener Aufzeichnungen ( $\beta = .09$ ;  $p = .193$ ) und die mitgebrachten Trainingsaufgaben ( $\beta = -.10$ ;  $p = .158$ ) erwiesen sich als keine signifikanten Prädiktoren für die Performanz in den Fragen zu konzeptuellem Wissen. Das Gesamtmodell ist hierbei signifikant ( $F(4, 97) = 2.25$ ,  $p = .035$ ), wobei die Regression 8% der Varianz in der Performanz im konzeptuellen Wissen aufklärt ( $R^2 = .085$ ). Von den Hypothesen 2, 4, 6 und 8 wird damit nur Hypothese 4 unterstützt, während die anderen verworfen werden müssen.

Ein explorativer Vergleich ergab auch hier, dass die Effekte der Materialien, die konstruktiven Lernaktivitäten zugeordnet wurden, nicht konsistent höher waren als die Effekte der Materialien, die niedrigeren Lernaktivitäten zugeordnet wurden.

## 4 Diskussion der Befunde

Ziel der vorliegenden Studie war es, mithilfe von zwei Regressionsmodellen den Einfluss mitgebrachter Unterlagen auf die geleistete Performanz in einer Open-Book-Klausur zu untersuchen. Hierbei wurde angenommen, dass ein positiver Zusammenhang zwischen den mitgebrachten Vorlesungsfolien, den zusätzlich kommentierten Vorlesungsfolien, den mitgebrachten eigenen Aufzeichnungen sowie den mitgebrachten Trainingsaufgaben und der Performanz in den Fragen zu konzeptuellem Wissen (Hypothesen 1, 3, 5 und 7) und prozeduralem Wissen (Hypothesen 2, 4, 6 und 8) besteht. Die statistischen Analysen zeigten, dass in Bezug auf die Performanz in beiden Wissensarten lediglich die zusätzlich kommentierten Vorlesungsfolien einen signifikanten, positiven Zusammenhang mit kleinem Effekt zeigten.

### 4.1.1 Theoretische Implikationen und praktische Empfehlungen

Erklärungen für die Ergebnisse lassen sich mithilfe der ICAP-Theorie herleiten. Der signifikante Befund der zusätzlich kommentierten Vorlesungsfolien auf die Performanz in beiden Wissensformen kann daraus resultieren, dass die Kommentierungen, welche auf den Folien getätigt wurden, nicht nur eine aktive Lernhandlung darstellen, sondern sogar auf der interaktiven bzw. konstruktiven Ebene anzusiedeln sind. Das bedeutet, dass womöglich Kurationsprozesse im Sinne von Vergleichen Verbindungen herstellen oder Generalisieren stattgefunden hat. Diese Aktivitäten führen nicht nur dazu, dass bestehendes Wissen mit neuem Wissen verknüpft wird, sondern Ersteres auch kohärenter, akkurater und strukturierter wird, was zu einem besseren Lernerfolg führt (CHI, 2009). In Übereinstimmung mit der Studie von ENGELBRECHT et al. (2005) gab es keinen Unterschied zwischen den zwei Wissensarten. Die Befunde bei dem Prädiktor mitgebrachte Vorlesungsfolien lassen sich wiederum so erklären, dass sie der ersten hierarchischen Ebene von Lernaktivitäten, der passiven Ebene, zuzuordnen sind. Diese wurden womöglich angesehen, jedoch fand keine weitere Auseinandersetzung mit diesen statt, wodurch kein zusätzliches Verständnis generiert und die Inhalte nicht weiter vertieft werden konnten. Überraschend sind hierbei die Ergebnisse der Seitenanzahl mitgebrachter eigener Aufzeichnungen und der mitgebrachten Trainingsaufgaben, da diese zugrundeliegende konstruktive beziehungsweise aktive Lernaktivitäten nahelegen. So wäre bei den eigenen Aufzeichnungen zu erwarten, dass diese zusätzlich erzeugte Inhalte enthalten, wie sie CHI (2009) beschreibt, und somit, wie bei den zusätzlich kommentier-

ten Vorlesungsfolien, zu einem besseren Lernverständnis führen. Da die Befunde jedoch dagegensprechen, müsste der Blick auf die konkreten Inhalte dieser eigenen Aufzeichnungen gerichtet werden, um deren Qualität zu untersuchen. Bei den Trainingsaufgaben, deren Befunde auch nicht signifikant wurden, handelt es sich um aktive Lernaktivitäten, insofern diese Aufgaben im Vorfeld auch selbst gelöst worden sind. Hier wären Lernprozesse zu erwarten, welche bereits vorhandenes Wissen aktivieren und mit neuem Wissen verknüpfen und verarbeiten (CHI, 2009). Da dies auch hier nicht der Fall ist, wäre es denkbar, dass die Trainingsaufgaben, welche über das Semester hinweg zur Verfügung gestellt wurden, nicht vertieft genug oder gar nicht bearbeitet wurden. Dies kann daraus resultieren, dass zu jeder Aufgabe eine Lösung ausgegeben wurde und die Studierenden womöglich nur diese mit in die Klausur nahmen, ohne sich weiter damit auseinanderzusetzen. Das Erstellen guter Mitschriften ist nicht selbstverständlich (WISNIEWSKI, 2019). Aus diesen Annahmen lässt sich ableiten, dass in Bezug auf die effektive Nutzung von mitgebrachten Unterlagen eine bessere Anleitung zur Gestaltung dieser notwendig wäre (CAPALDI, 2019).

#### **4.1.2 Limitationen**

Bei der Interpretation der Ergebnisse sollten die Limitationen dieser Untersuchung miteinbezogen werden. Zum einen befanden sich unter den befragten Studentinnen und Studenten solche, die die Klausur bereits einmal geschrieben haben, diese jedoch nicht bestanden hatten. Das kann bedeuten, dass sie bereits einmal sehr ähnliche Aufgaben bearbeitet haben und dies einen Einfluss auf die Performanz im Test hatte. Zudem sind eben diese Studierenden sehr wahrscheinlich aus höheren Semestern und haben daher im Gegensatz zu anderen Studentinnen und Studenten bereits mehr Erfahrung im Umgang mit mitgebrachten Unterlagen bei Open-Book-Klausuren gesammelt. Auch wurde nicht danach gefragt, inwiefern und in welchem Ausmaß genau die zusätzlich kommentierten Vorlesungsfolien beschriftet wurden, so auch bei den mitgebrachten eigenen Aufzeichnungen, bei welchen sowohl kein Wissen darum besteht, wie diese konkret aussahen, als auch darüber, wie groß oder voll beschriftet die Seiten, nach welchen gefragt wurde, genau waren. Folgestudien sollten hier mehr ins Detail gehen und auch eventuelle qualitative Studien zur Qualität der mitgebrachten Materialien beinhalten.

Zwei weitere Punkte, welche als Einschränkungen gesehen werden könnten, sind zum einen, dass die durchgeführte Befragung zur Selbsteinschätzung der prozentual mitgebrachten Unterlagen als zu vage und subjektiv gesehen werden kann. Zum anderen die Gegebenheit, dass der Fragebogen direkt vor der Klausur ausgeteilt wurde und dieser somit womöglich zu flüchtig und nicht bei voller Konzentration ausgefüllt wurde. Diese Aspekte könnten zu verzerrten Angaben beim Ausfüllen des Fragebogens geführt haben. Allerdings liefert diese Form der Befragung eine sehr niederschwellige Einschätzung, ohne den Ablauf der Klausur, und damit die Validität der Situation, einzuschränken. Abschließend ist auch die Einteilung der Fragen zu den jeweiligen Skalen der Wissensarten hinterfragbar, da, wie GÖLLER (2020, S. 65) schreibt, „diese beiden Wissenskategorien [...] nicht ganz trennscharf [sind]“. Wobei er sich auf eigene Definitionen zu deklarativem Wissen, einem ähnlichen, aber etwas weiter gefassten Begriff als konzeptuelles Wissen, und prozeduralem Wissen bezieht. Die Homogenität der Ergebnisse weist aber darauf hin, dass auch eine andere Aufteilung keine grundsätzlich anderen Befunde verursacht hätte.

Schließlich muss auch das Klausurformat selbst kritisch betrachtet werden, da die Relevanz bestimmter Hilfsmittel zwingend von den gestellten Fragen abhängen. Alle Aufgaben wurden so gestellt, dass eine bloße Wiedergabe des Inhalts der Vorlesungsfolien nicht ausreicht, um die Aufgaben zu lösen. Es überrascht daher nicht, dass die Anzahl der mitgebrachten Folien keinen substanziellen Einfluss auf den Klausurerfolg hatte. Allerdings legt der mangelnde Einfluss der eigenen Aufzeichnungen nahe, dass die Aufgaben dennoch potenziell nah an den Vorlesungsfolien erstellt wurden. Eine Replikation des gefundenen Musters an systematisch gestellten Aufgaben wäre daher sicher notwendig, um die Generalisierbarkeit der Ergebnisse abzusichern.

## 5 Fazit

Die schnelle Einführung von Online-Prüfungen im vergangenen Semester stellte für viele Lehrende und Studierende eine Herausforderung dar. Auf Basis der gewonnenen Erfahrungen und Erkenntnisse ist davon auszugehen, dass problemlösungsorientierte Prüfungsformate, wie Open-Book, zunehmen werden. Die Entwicklung der nächsten Jahre wird ferner zeigen, ob Hochschulen verstärkt Gebrauch von innovativen Prüfungsverfahren machen. Während dies für die Lehrenden höhere Ansprüche in der Fragengestaltung bedeutet, müssen sich Studierende grundsätzlich in ihrem Lernverhalten anpassen. Die vorliegende Studie kann dabei erste Erkenntnisse liefern, wie Studierende in dieser neuartigen Prüfungssituation ideal unterstützt werden können. Beispielsweise durch eine detaillierte Anleitung zur Erstellung eigener Vorlesungsaufzeichnungen, die Möglichkeiten sind zahlreich.

## 6 Literaturverzeichnis

**Batz-Finkbohner, J., Alexander Besner, A., & Gerstner, M.** (2021). *Handreichung Open-Book-Prüfung*. Fernprüfungen an bayerischen Universitäten. Bayerisches Kompetenzzentrum für Fernprüfungen. TU München. [https://www.prolehre.tum.de/fileadmin/w00btq/www/Angebote\\_Kompetenzzentrum/20211022-BayKFP-Handreichung\\_Open\\_Book\\_Pruefungen.pdf](https://www.prolehre.tum.de/fileadmin/w00btq/www/Angebote_Kompetenzzentrum/20211022-BayKFP-Handreichung_Open_Book_Pruefungen.pdf)

**Capaldi, M.** (2019). The Use of Exam Notes in an Online Mathematics Course. *The Journal of Educators Online*, 16. <https://doi.org/10.9743/JEO.2019.16.2.3>

**Carpenter, T. P.** (1986). Conceptual knowledge as a foundation for procedural knowledge. *Conceptual and procedural knowledge: The case of mathematics*, 113–132.

**Chi, M. T. H.** (2009). Active – constructive – interactive: a conceptual framework for differentiating learning activities. *Topics in cognitive science*, 1(1), 73–105. <https://doi.org/10.1111/j.1756-8765.2008.01005.x>

**Chi, M. T. H., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., Levy, R., Li, N., McEldoon, K. L., Stump, G. S., Wylie, R., Xu, D. & Yaghmourian, D. L.** (2018). Translating the ICAP Theory of Cognitive Engagement Into Practice. *Cognitive science. Vorab-Onlinepublikation*. <https://doi.org/10.1111/cogs.12626>

- Chi, M. T. H. & Wylie, R.** (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, 49(4), 219–243. <https://doi.org/10.1080/00461520.2014.965823>
- Crawford, J., Butler-Henderson, K., Rudolph, J., Malkawi, B., Glowatz, M., Burton, R., Magni, P. A., & Lam, S.** (2020). COVID-19: 20 countries' higher education intra-period digital pedagogy responses, *Journal of Applied Learning & Teaching*, 3(1), 1–20. <https://doi.org/10.37074/jalt.2020.3.1.7>
- Durning, S. J., Dong, T., Ratcliffe, T., Schuwirth, L., Artino, A. R., Boulet, J. R., & Eva, K.** (2016). Comparing open-book and closed-book examinations: a systematic review. *Academic Medicine*, 91(4), 583–599.
- Engelbrecht, J., Harding, A., & Potgieter, M.** (2005). Undergraduate students' performance and confidence in procedural and conceptual mathematics. *International Journal of Mathematical Education in Science and Technology*, 36(7), 701–712. <https://doi.org/10.1080/00207390500271107>
- Gharib, A., Phillips, W., & Mathew, N.** (2012). Cheat Sheet or Open-Book? A Comparison of the Effects of Exam Types on Performance, Retention, and Anxiety. *Online Submission*, 2(8), 469–478.
- Göller, R.** (2020). *Selbstreguliertes Lernen im Mathematikstudium. Eine qualitative Studie zur Beschreibung und Erklärung der Lern- und Problemlösestrategien von Mathematikstudierenden im ersten Studienjahr mithilfe ihrer Ziele, Beliefs und Bewertungen*. Wiesbaden: Springer.
- Green, S. G., Ferrante, C. J., & Heppard, K. A.** (2016). Using open-book exams to enhance student learning, performance, and motivation, *Journal of Effective Teaching*, 16(1), 19–35. [https://uncw.edu/jet/articles/vol16\\_1/index.htm](https://uncw.edu/jet/articles/vol16_1/index.htm) Stand vom 13. Oktober 2021.
- Haapasalo, L. & Kadjevich, D.** (2000). Two Types of Mathematical Knowledge and Their Relation. *Journal für Mathematik-Didaktik*, 21(2), 139–157. <https://doi.org/10.1007/BF03338914>
- Horn, J., & Schmees, M.** (2020). Online-Prüfungen. ELAN eV Handouts. Leitfaden. ELAN eV. <https://elan-ev.de/dateien/Online-Pr%C3%BCfungen.pdf>
- Ioannidou, M. K.** (1997). Testing and life-long learning: Open-book and closed-book examination in a university course. *Studies in Educational Evaluation*, 23(2), 131–139. [https://doi.org/10.1016/S0191-491X\(97\)00008-4](https://doi.org/10.1016/S0191-491X(97)00008-4)

**Lehner, M.** (2018). Lehren und Lernen an der Hochschule der Zukunft. In *Hochschule der Zukunft* (pp. 167–185). Wiesbaden: Springer VS.

**Rakes, G. C.** (2008). Open book testing in online learning environments, *Journal of Interactive Online Learning*, 7(1), 1–9. <http://www.ncolr.org/issues/jiol/v7/n1/open-book-testing-in-online-learning-environments.html>

**Rummer, R., Schweppe, J., & Schwede, A.** (2019). Open-Book Versus Closed-Book Tests in University Classes: A Field Experiment. *Frontiers in psychology*, 10, 463. <https://doi.org/10.3389/fpsyg.2019.00463>

**Wisniewski, K.** (2019). Mitschreiben in Vorlesungen: Ein interdisziplinärer Forschungsüberblick mit Fokus Deutsch als L2. *Bulletin VALS-ASLA*, 19, 153–170.

## Autor\*innen



Prof.in Dr.in Sandra NIEDERMEIER || Hochschule für angewandte Wissenschaften Kempten, Institut für digitale Transformation in Arbeit, Bildung und Gesellschaft || Bahnhofstraße 61 D-87435 Kempten (Allgäu)

<https://www.hs-kempten.de/personen/sandra-niedermeier>

[sandra.niedermeier@hs-kempten.de](mailto:sandra.niedermeier@hs-kempten.de)



Dr. Michael SAILER || Ludwig-Maximilians-Universität München, Department Psychologie || Leopoldstraße 13, D-80802 München

<https://www.psy.lmu.de/ffp/persons/ag-fischer/sailer-michael/index.html>

[michael.sailer@psy.lmu.de](mailto:michael.sailer@psy.lmu.de)



Lukas REMBÖCK || Ludwig-Maximilians-Universität  
München, Department Psychologie || Leopoldstraße 13,  
D-80802 München

[lukasremboeck@yahoo.de](mailto:lukasremboeck@yahoo.de)



Dr. Matthias STADLER || Ludwig-Maximilians-Universität  
München, Department Psychologie || Leopoldstraße 13,  
D-80802 München

<https://www.psy.lmu.de/ffp/persons/ag-fischer/stadler/index.html>

[matthias.stadler@uni-muenchen.de](mailto:matthias.stadler@uni-muenchen.de)



Eva OSTERESCH<sup>1</sup> (Oldenburg)

# Designing a case study approach for competence-oriented examination in large courses

## Abstract

The advantages and potentials of case study work are well known. However, the use in German large university courses and the use as a competence-oriented examination are still not common. This article aims to address this unrealized potential of case studies by presenting a scientifically sound and effective teaching concept that has been designed, implemented, and evaluated over three semesters in a bachelor module with about 400 participants. Students indicate that the teaching concept leads to high satisfaction, high motivation, high learning success, and fun. This evaluation emphasizes the use of case studies as a competence-oriented examination, even in large courses.

## Keywords

case study work, large courses, competency-oriented assessment, constructive alignment

## 1 Introduction

The advantages of case study work are well known. Case studies, when applied appropriately, are suitable to activate and motivate learners, to promote personal, social, and methodological skills, to present complex issues, and to generate long-term learning success (GERDSMAIER, 1979; GRAGG, 2013; MÜLLER 2007; PILZ, 2001). Therefore, it is surprising that they are hardly used in German university

---

<sup>1</sup> email: [eva.osteresch@uni-oldenburg.de](mailto:eva.osteresch@uni-oldenburg.de)



contexts (HOIDN, 2010; MELCHER, 2021). Especially the framework conditions of large bachelor modules (i. e., many participants, limited resources, broad teaching content, little prior knowledge) are considered unsuitable for the use of case studies. The scope of case study work is thus limited and possible benefits are not addressed (HILGERS-SEKOWSKY & HUXOLD, 2021; MELCHER, 2021).

Another unrealized potential of case study work is the use as a competence-oriented assessment. Gaining competencies primarily entails “cognitive abilities and skills to solve specific problems, as well as the associated motivational, volitional, and social readiness and skills to use the problem solutions successfully and responsibly in variable situations.” (WEINERT, 2014, p. 27). Unlike standardized teaching approaches with the structure “lecture, exercise, exam” and a focus on replicating existing knowledge without linking the teaching and testing situation (DÖBLER, 2019; SCHINDLER, 2015), case studies can promote multiple competencies and focus on the application of knowledge. One way to use case studies as an examination is the portfolio, which is already implemented in many examination regulations and offers a possibility to implement practice- and process-oriented teaching and examination. There is anecdotal evidence that the use of case studies as an examination can be useful, but implementation fails due to a lack of workable approaches. With the demands of competence-oriented assessments and constructive alignment, it is time to leverage this potential (BIGGS, 1996; DEN OUDEN et al., 2019; KANDL-BINDER, 2014).

This paper addresses both challenges at once: how can the advantages of case study work be applied in large university courses and how can case study work serve as a competence-oriented examination? A design science approach is chosen for this purpose, as it represents a valuable means for generating theoretical insights and emergent solutions to the real-world problem of case study implementation (GREGOR & HEVNER, 2013; HEVNER et al., 2004). The approach will be depicted in depth in the methodology section.

The artificial solution designed and proposed by this research is a teaching concept that facilitates case study work as a portfolio-examination in German large courses. Therefore, the teaching concept of a bachelor course in supply chain management (approx. 350–400 participants) was redesigned and evaluated over three years. The paper contributes to the theory of case study work by identifying possible applications, success factors, and positive learning effects. It contributes to competen-

cy-based examination theory by highlighting case studies as a possible examination instrument (including large courses). In terms of practice, the framework provides a practical tool for teachers to develop their own successful teaching concept with case studies.

The following section lays out the conceptual background of case study work. The subsequent section justifies and expounds on the design science approach. Chapter four presents the proposed design solution. It also highlights key findings from the evaluation, to underline the effectiveness of the approach. The concluding discussion identifies theoretical and practical implications, acknowledges limitations, and suggests future research avenues.

## 2 Academic case studies

Case study work has its origins in the 1870s at Harvard Law School and has subsequently moved to school and university contexts (ARNDT, 2013; KOSIOL, 1957; MAY, 2010). In particular, the Bologna reform and accompanying efforts to improve the quality of teaching at universities have strengthened the use of case studies (BANSCHERUS et al., 2009). The aim of case study work for learners is to identify and analyze problems, to develop and evaluate solutions independently, and then to make reasoned decisions (ARDNT, 2013). In doing so, learners expand their problem-solving and decision-making skills by discussing different solutions to the case study in groups and agreeing on a result (HERREID, 2011; KAISER, 1976). They learn to think in a differentiated manner, to weigh up alternatives, and to find and evaluate information (ARNDT, 2013; BONZ, 2009). In addition, social competence, the ability to abstract, and moral-ethical competence are promoted through case study work (REBMANN et al., 2011). Overall, case studies, when used properly, are capable of activating and motivating learners, presenting complex issues, and achieving long-term learning success (GERDSMEIER, 1979; GRAGG 2013; MÜLLER, 2007).

On a theoretical level, case studies are suitable for competence-oriented teaching and testing (WEINERT, 2014). For example, in the expert report on competence orientation in study and teaching (SCHAPER et al., 2012), case studies are emphasized for the ability to assess complex issues, the promotion of application and practical

relevance of scientific learning content processes, the promotion of interdisciplinary and professional competences, and process-oriented examinations. Especially the motivational effect of case study work is noteworthy, as it is a central factor for learning success and immanent to competency attainment (DECI & RYAN, 1993; PRENZEL et al., 1996; WEINERT, 2014). In particular, perception of autonomy, perception of competence, social inclusion, relevance of content, and instructional quality are conditions for self-determined motivation that can be fulfilled by case study work (DECI & RYAN, 1993; PRENZEL et al., 1996).

Despite the presented benefits and positive effects, case studies are hardly used in the German university context (HOIDN, 2010). Two established application scenarios of case study work are the Harvard Business School and the German vocational schools (HBS, 2021; ARDNT, 2013). The case study approach prevalent in the English-speaking university environment is the Harvard Case Method (HCM) (HBS, 2021; MELCHER, 2021). It is developed for university teaching (approx. 90 participants). Thereby 12 to 28 cases are processed per semester. The processing is obligatory and consists of an individual preparation, a small group discussion, a plenary discussion with the teacher, and a self-reflection. The HCM cases refer to a real situation in an organization at a specific time with a need for action (MELCHER, 2021). Its use at German universities is hindered by the lack of German-language cases and the inability to assess “active participation” in German examination regulations (MELCHER, 2021).

Another example is the use in the German (vocational) school context. In the classes (approx. 20 participants), the cases are usually worked on in groups under the guidance of the teachers and then discussed in plenary. According to BRETTSCHEIDER (2000), case study work should be a social process (e. g., through group work), a guided process (support from instructors), a constructive process (active action by learners, e. g., research), and an independent process (own scope of action).

Both approaches fit the basic situation to a limited extent as they were developed for smaller numbers of participants and not as an examination. At the same time, components such as using real-world settings, working through multiple cases to apply the learning content, and Brettschneider’s rules are important directions for further development.

The lack of German language cases and the intended use as an exam necessitate the design of new case studies. Therefore, established case study design criteria are considered.

A case study comprises an authentic and realistic case or basic information, (guiding) questions or tasks as well as rich information material or guiding texts (EULER & HAHN 2014; WILBERS, 2014; WOLFF, 1992). The case designer should first define the learning objectives and align the case studies accordingly. The cases should be problem-oriented, conflict-containing, and complex so that several possible solutions arise (PILZ & KRÜGER, 2013). While the HCM only deals with real companies and situations, fictional companies are used in the school context. The aim is to reduce the real-world complexity to a reasonable level (WOLFF, 1992).

One possibility to use case studies within an examination is the portfolio. In addition to the usage as a self-reflection tool or learning diary, a portfolio can also be used as an assessment tool (HÄCKER, 2007). This form of portfolio refers to a combination of several partial performances that form the examination. Nowadays, many German examination regulations include the portfolio as a possible examination method. The portfolio literature particularly emphasizes the relevance of support and feedback processes that must be included in the design of the teaching approach (STRATMANN et al., 2009). Sometimes case studies are already included in portfolios, but little research has been done on the possible combination of the two approaches (RIEBE & SCHEFFLER, 2017). The use of case studies in portfolios offers the possibility to transfer the positive effects of case study work (see above, e. g. active learners, competence orientation) into an examination situation. How this can be realized is shown below.

### **3 Design science approach**

The design science method was chosen for the development of a new case study approach because of its suitability for developing effective solutions to real-world problems (GREGOR & HEVNER, 2013; HEVNER et al., 2004). Originating in computer science, the design science approach is now established internationally and across disciplines (BUSSE et al., 2017; HOLMSTRÖM et al., 2009). The basis is always a scientific foundation (GREGOR & HEVNER, 2013). Design science

is more concerned with the question of whether something “works” than whether something is “true” (ROMME, 2003). At the same time the result “it works” always offers scientific contributions about why and under which circumstances something works. Starting point of the research is the environment and the knowledge base. Under these framework conditions, an artifact (teaching concept) is designed, implemented, and evaluated. The research design is shown in Figure 1.

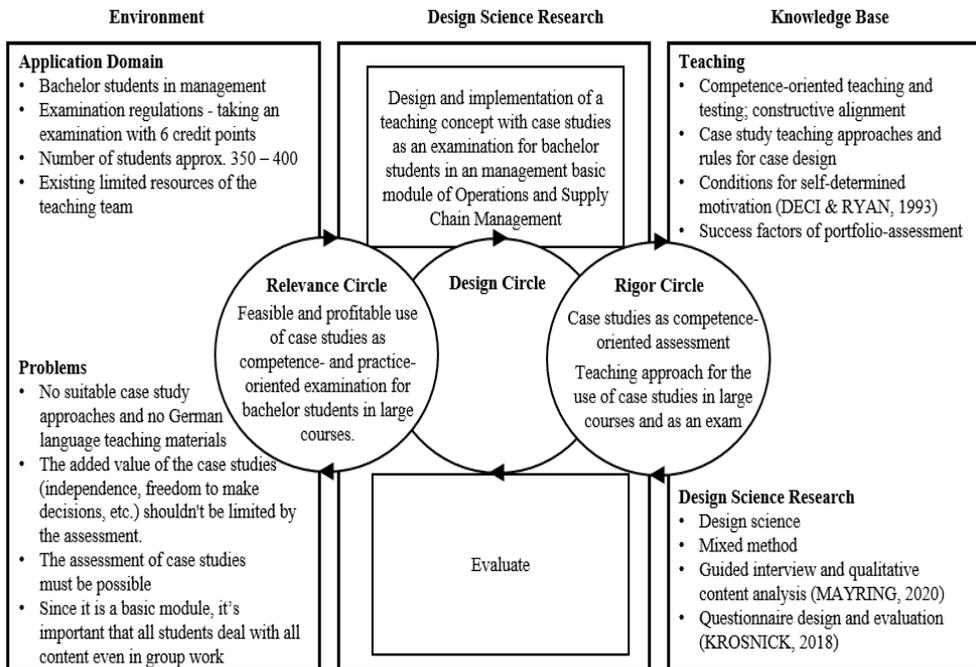


Fig. 1: Research design (own illustration, based on HEVNER et al. (2004))

The teaching concept was designed, implemented, evaluated, and refined by the author in collaboration with the teaching team in three consecutive years. In each year,

several types of evaluations were used. Table 1 details the data collections that were conducted and the basic population.

Table 1: Basic population and data base

	<b>2019</b>	<b>2020</b>	<b>2021</b>	$\Sigma$
Participants	333	410	359	1102
Official evaluation (N)	57	<i>not performed</i>	46	103
Interviews (N)	14	<i>not performed</i>	7	21
Additional survey (N)	<i>not performed</i>	103	105	208

The evaluation was conducted on an annual basis to improve the artifact in line with the design science approach. Like the teaching concept, the evaluation concept has evolved continuously. In 2021, the most comprehensive evaluation was conducted to examine the effectiveness of the final developed artifact. All individuals in the basic population had the opportunity to participate in all survey forms of the evaluation.

Part of the evaluation was the official online teaching evaluation. This could not be influenced by the researcher. The evaluation contained questions about the study situation, the course, the teachers, the students' own learning as well as an overall assessment and free text comments about positive and negative course aspects. The repetitive use of the same items over the years showed the longitudinal trend that results from the further development of the module. In 2020, there was an exception to the questionnaire due to the Corona pandemic. The official teaching evaluation was adapted to the situation and no longer contained all items. These were therefore included in a self-administered additional survey of 2020.

The additional online survey was conducted in 2020 and 2021. The questionnaire included items about the teaching concept, the teaching team, and open-ended questions about positive and negative aspects of case study work as well as suggestions for improvement. A 5point Likert scale was chosen to enable a neutral response (KROSNICK, 2018).

In 2021, the questionnaire was expanded to include the motivational conditions for self-determined motivation as it is a central factor for learning success and immanent to competency attainment (DECI & RYAN, 1993). Social integration and content interest on the part of the teacher were not surveyed, as the standardized items were developed for face-to-face teaching and were therefore only transferable to the online scenario to a limited extent. Literature-based and proven items were used (at least 4 per scale), which were presented in a mixed form (PRENZEL et al., 1996).

In 2019 and 2021, semi-structured interviews were conducted as part of final projects. Central were the questions about the positive and negative effects of the case study work, group work, motivational effects as well as possible suggestions for improvement.

The data analysis combined quantitative and qualitative methods in a mixed-method approach. The quantitative questionnaire results were analyzed using descriptive statistics via SPSS. A reliability analysis was performed for the motivation scales. The open questions and suggestions of the surveys and the interviews were evaluated using a qualitative content analysis according to MAYRING (2020). The data were examined for statements about components of the teaching concept, forming a deductive category in each case. In addition, an inductive category was formed for student sensations.

## 4 Findings

### 4.1 Case study teaching approach

The goal of the design science approach was to develop a teaching concept that meets the practical and theoretical requirements shown in Figure 1. This resulted in certain key factors of the teaching concept. First, a written elaboration with a defined appropriate scope (no presentation or oral exam) to ensure fair and uniform assessment as well as feasibility in a large course. In addition, group work was chosen to promote discussion, exchange, and social skills, and to minimize the correction effort. Transparency of learning objectives and requirements, fair assessment and strong feedback structures as prerequisites for good portfolio work and competency-based teaching were also important. The translation of these factors into a teaching concept is illustrated in Figure 2.

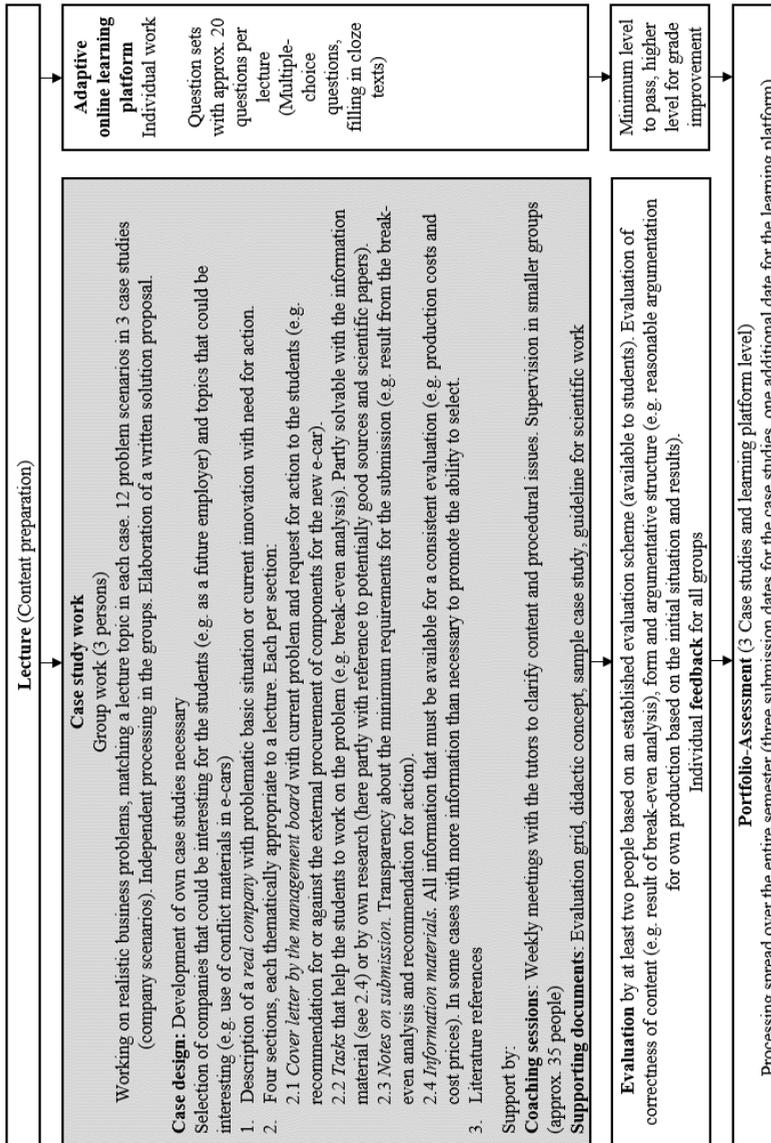


Fig. 2: Case study teaching approach

Figure 2 contains relevant components of the teaching concept and describes the didactic and methodical sequence. Central to the success was the development of own case studies tailored to the course content and auditability. The described case design can serve as a template for other teachers to convert their own teaching content into a case study.

One component teachers and students considered useful is the additional and individual learning assessment. This counteracted two disadvantages of group work formulated by students, the formation of “experts” (not all deal with all content) and group assessment.

The teaching concept followed the rules of constructive alignment and combined the teaching and testing situation. The learning objectives were made transparent. Students were put into professional situations where they had to apply their knowledge to recommend a decision for “their” board. Through group work and the construction of the tasks (e. g. research tasks, analyzing employee interviews), method and social skills were addressed in addition to professional competence.

Compared to other case study approaches, the combination of case studies and examination eliminated common steps such as a class discussion with the teacher. Moreover, the case study structure was more small-step. Students were placed in multiple decision-making situations. In each scenario, they were given a task by the management board which they had to solve (similar to a real professional situation) with the help of internal and external company materials and their own knowledge (gained through the lecture).

Whether and how the positive effects of case study work could also be achieved in this teaching setting is presented in the following.

## 4.2 Effectiveness of the case study approach

Table 2: Results of the official teaching evaluation

	2018 <sup>1</sup>	2019 <sup>1</sup>	2020 <sup>2</sup>	2021 <sup>1</sup>
	MV	MV	MV (Tra.)	MV
The course...				
...is structured in a comprehensible way	1.31	1.32	1.29 (1.22)	1.33
... is well organized	1.44	1.40	1.60 (1.45)	1.17
...stimulates critical debate	2.26	1.80	1.79 (1.59)	1.47
Related to the entire module...				
... the course has meaningfully expanded my competencies according to the module description	1.55	1.47	1.57 (1.43)	1.18
...the form of examination is appropriate	1.68	1.33	1.71 (1.53)	1.35
...the ratio of workload and credit points is appropriate	Yes (69%)	Yes (95%)	2.38 (2.04)	Yes (81%)
I am overall satisfied with the module	1.61	1.47	1.80 (1.60)	1.35

*Note:* For the years 2018 (N=87), 2019 (N=57) and 2021 (N=46), this is the official teaching evaluation of the university. In 2020 (N=103), due to online teaching, the university modified the official teaching evaluation so that the relevant items were collected through a separate survey.

<sup>1</sup> Measured with a 4-point Likert scale

<sup>2</sup> Measured with a 5-point Likert scale, mathematically transposed to a 4-Likert scale for better comparability (values in parentheses).

In 2019, the case studies were initially only used as a substitute for traditional exercises in the tutorials and not as a form of examination. However, the teaching team

perceived that students did not engage with the case studies extensively, as they did not see any added value for this for the exam. Therefore, the teaching concept and examination method were extensively revised again in order to actually achieve a constructive alignment of teaching and examination (see Figure 2). The data showed continuous improvement in almost all areas (see Table 2). Notable are the increases in the areas of critical debate, overall satisfaction, and appropriateness of the form of examination.

Table 3: Students' sensations based on the qualitative content analysis

Codes	Excerpts from the data
Case studies better than written exams/learning by heart (30)	<i>"That you don't have this bulimic learning, that's why I find such a portfolio performance much better than an exam, because it's just not this dull learning by heart. Rather, you apply the knowledge and I think that's good."</i>
Very good module (8)	<i>"Overall, this module seems to be the best module I have taken so far (content, structure, explanation and support)."</i>
High learning success (15)	<i>"In my opinion, the long-term learning success is ensured and it has more impact than an exam!"</i> <i>"Through the intensive discussion I can better retain what I have learned"</i>
High motivation (21)	<i>"More motivated than in a written exam, was much more fun than "bulimia learning" for an exam."</i>
Fun/Interest (15)	<i>"The case studies were well developed and therefore fun to work on!"</i>
Autonomy/self-reliance (25)	<i>"[I liked:] The independent work and discussion in the group."</i>
Deeper discussion/application (22)	<i>"I think the advantage of the case study is that you do it again and again and don't just learn it by heart, but you also have to really understand it in order to explain it to someone else, i.e. our imaginary board of directors."</i> <i>"That you could deal with the contents of the lectures in tasks and thus acquire more understanding."</i>

*Note:* The coding frequency is shown in parentheses after the code. Data originate from the open questions and suggestions of the official teaching evaluation (N=46), the additional questionnaire (N=105) as well as the transcripts of the seven conducted guided interviews on the topic "case study work in university teaching".

In the following, the evaluation results of the year 2021 are presented, since the development of the artifact was completed at that time. The results, therefore, provide information on the effectiveness of the artifact shown in Figure 2.

The students mentioned many positive aspects of the case study work (see Table 3). The general tenor was that case study work should be retained as an examination, as it led to high learning success, high motivation and was more fun than classic teaching approaches. The results indicate that the requirements for a competence-oriented examination could be met. Students also nominated the module as the university’s “best module” three times in a row, for the internal teaching award.

As shown in Table 4, the questionnaire results support the perceived positive effects of the teaching concept.

Table 4: Results of the additional questionnaire 2021 regarding the didactic concept

Item	MV	SD
The case study work allowed me to get a deeper understanding of the issues.	4.51	0.63
The discussions in my case study group allowed me to gain a deeper understanding of the issues.	3.94	1.12
I have internalized essential concepts through constant exposure to the topics via the learning platform.	3.90	1.66
I think the combination of online lecture, learning platform and case study work is useful.	4.08	1.72
The coaching sessions were very useful for me.	4.11	1.11
Working with individual original scientific essays was enriching.	3.16	1.86
All in all, the case study work overwhelmed me. <sup>1</sup>	1.95	0.95
I believe I gained a better understanding of the content by working through the case studies than I would have with an exam.	4.56	0.86
I estimate that the long-term learning success of a case study is higher than the learning success of an exam.	4.61	0.69

Note: N=105, measured with a 5-point Likert scale (1 = I totally disagree, 5 = I fully agree).

<sup>1</sup> The item is reverse coded.

In order to further investigate the motivational impact of the case studies, the motivation conditions content relevance, perception of competence, perception of autonomy, and instructional quality were surveyed in the 2021 questionnaire (see Table 5).

Table 5: Motivation conditions

Scale	Item	MV
Content relevance Scale mean 4.20 $\alpha^1 = 0.72$	...was shown by examples or problems how important the material is	4.16
	...I have learned that I can also use what I have learned in my later professional life	4.09
	...it was made clear which role the learning subject plays in operational processes/coherences	4.40
	...I was put in situations where I could see for myself how important the material was	4.10
Perception of competence Scale mean 4.35 $\alpha^1 = 0.62$	...I had ample opportunity to practice what I had learned	4.40
	...I was also trusted with difficult tasks	4.52
	...I was able to use my strengths	4.20
	...I was able to see what I can do already	4.25
Perception of autonomy Scale mean 4.39 $\alpha^1 = 0.69$	...I was allowed to do tasks my way	4.27
	...I was encouraged to proceed independently	4.52
	...I was able to complete demanding tasks on my own responsibility	4.36
	...I had the opportunity to deal with interesting tasks or contents in more detail	4.40
Instructional quality Scale mean 4.44 $\alpha^1 = 0.61$	...I had the opportunity to handle new areas independently	4.43
	...the material was illustrated with examples	4.40
	...tools (e.g. instructions, guidelines...) were available	4.44
	...my questions were answered	4.46
	...I was informed about the learning objectives (what I should be able to do)	4.30
...I have received an overview of the proposed course of action	4.39	

Note: N=105, measured with a 5-point Likert scale (1 = I totally disagree, 5 = I fully agree).

<sup>1</sup> Reliability analysis, Cronbach's alpha, Performed with SPSS (GEORGE & MALLERY, 2003)

Reliability was satisfactory for two scales and questionable for two scales (GEORGE & MALLERY, 2003). The reliability could not be increased by omitting individual items, and no further items were added due to the already existing length of the

questionnaire. Overall, high scores were found for all motivation conditions, reinforcing the positive benefits of case studies.

The results show that competence-oriented teaching and testing is possible even in large bachelor courses. The case studies are suitable for exposing students to professional problem situations that they have to solve independently. It was necessary to develop good evaluable case studies as well as a different way of processing the case studies (without, e. g., a class discussion). The results show that the positive effects of the case study work were not decreased by the alternative use. On the contrary, students described positive effects such as high motivation, fun and enjoyment during the work, and high learning successes. These were estimated to be higher and more long-term than they would have been in the case of a classic teaching structure.

The designed teaching concept can presumably be transferred to other subjects and modules. It can serve as a blueprint for a variety of teaching contents, provided that the teachers are willing to design their own case studies or have access to suitable or adaptable case studies. If the key components, a strong support structure, feedback processes, group work, content preparation (e. g. through a lecture), individual performance review, and clear and fair assessment structures are maintained, it can be assumed that the presented positive effects can also be achieved in other modules.

## **5 Discussion**

The design science approach has enabled the development of a theoretically sound, practical, and student-appreciated teaching concept for the use of case studies as competency-based examination in large courses (see Figure 2).

The teaching concept works and is effective. The case studies are used to put students in professional situations where they have to apply their knowledge, analyze informational material, and agree on a solution as a group to recommend a decision for “their” board. The repetitive evaluation results show an improvement in the categories satisfaction with the module, competence enhancement, structure, and organization as well as the promotion of critical thinking through the implementation and optimization of the teaching concept. The qualitative content analysis shows that in the perception of the students there is a high learning success, high motivation,

fun/interest in the processing, and high perceived autonomy/self-determination. The results indicate that a competence-oriented form of examination could be achieved through the case study work. Students describe that they prefer the presented teaching concept to a classical teaching structure with a written exam.

The presented approach has many advantages compared to a classical teaching approach (lecture, exercise, exam). The successful implementation also shows the feasibility with an appropriate teaching team (in this case professor, scientific employee, six tutors). This paper thus demonstrates that competence-oriented teaching and testing with case studies can be successful even in large courses.

## **5.1 Scientific implications**

The results provide contributions to the topic area of case study work. In particular, due to the wide range of applications, there are only a few scientifically based contributions on the impact and design of special teaching concepts. The design science approach shows that the use of case studies as a competence-oriented examination in large German courses is possible and successful. The research therefore contributes to expand the previous scope of application of case studies. The findings indicate that the use of case studies as an examination (instead of a substitute for exercises) seems to be particularly meaningful since the students have a high level of activity (especially since “active collaboration” cannot be assessed in the current German examination law). The profitable combination of case study work and portfolio work, especially with regard to a process-oriented, competence-oriented teaching and examination structure, is emphasized and a possible implementation is shown.

## **5.2 Practical implications**

The results support the efforts of competence-oriented examinations at universities and offer teachers central starting points for a successful implementation. The developed teaching concept and the design structure of the case studies (see Figure 2) can therefore serve as a blueprint for other teachers. In addition to all forms of business management modules, the teaching concept could be of interest to all teachers who are confronted with a high number of students and still want their teaching and ex-

amination structures to have a high level of practical application. Even though more effort is required from teachers and learners, the results show that it is feasible and advantageous. The positive evaluation results provide incentives for realizing the potential of case studies in practice.

### **5.3 Limitations and future research**

There are at least four potential limitations concerning the findings of this study. The first limitation concerns the scenario of the study. The teaching event is not necessarily representative of other large events. In addition, the switch to online teaching during the Corona pandemic greatly changed the setting of the study. The transfer to other courses would be interesting to explore whether the presented effects can be generalized. A second potential limitation is that the author's scientific and personal background may have influenced the coding. It should also be emphasized that in this case, the evaluations are carried out by one of the teachers. A third limitation concerns the different ways of collecting data that have evolved over the years. As a result, not all developments can be tracked over the years (e. g., self-determined motivation). Moreover, the number of students surveyed and the intensity of the survey changed over the years. Finally, the orientation of the evaluation was primarily developmental in the sense of the applied design science approach. Therefore, the observed effects are based on the students' self-assessment. After the feasibility and initial effectiveness of the teaching concept have been established, it would be useful for future research to validate it through standardized procedures in additional research settings.

Despite these limitations, this research identifies an effective teaching approach for the case study work that can be useful for other similar modules and makes contributions to case study theory, by targeting its use as a competence-oriented examination in large courses.

## 6 References

- Arndt, H.** (2013). *Methodik des Wirtschaftsunterrichts*. Opladen: Barbara Budrich.
- Banscherus, U., Himpele, K., & Staack, S.** (2009). Zehn Jahre Bologna: Studienreform zwischen Anspruch und Wirklichkeit. *Forum Wissenschaft*, 26(4), 12–15.
- Bargel, T., Multrus, F., Ramm, M., & Bargel, H.** (2009). *Bachelor-Studierende-Erfahrungen in Studium und Lehre: eine Zwischenbilanz*. Berlin: Bundesministerium für Bildung und Forschung.
- Biggs, J.** (1996). Enhancing Teaching Through Constructive Alignment. *Higher Education*, 32(3), 347–364.
- Bonz, B.** (2009). *Methoden der Berufsbildung: ein Lehrbuch* (2nd ed.). Stuttgart: Hirzel.
- Brettschneider, V.** (2000). *Entscheidungsprozesse in Gruppen: theoretische und empirische Grundlagen der Fallstudienarbeit*. Bad Heilbrunn: Klinkhardt.
- Busse, C., Schleper, M. C., Weilenmann, J., & Wagner, S. M.** (2017). Extending the Supply Chain Visibility Boundary: Utilizing Stakeholders for Identifying Supply Chain Sustainability Risks. *International Journal of Physical Distribution & Logistics Management*, 47(1), 18–40.
- Deci, E. L., & Ryan, R. M.** (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223–238.
- den Ouden, H., Frölich-Steffen, S., & Gießmann, U.** (2019). *Kompetenzorientiert prüfen und bewerten an Universitäten: Didaktische Grundannahmen, rechtliche Rahmenbedingungen und praktische Handlungsempfehlungen*. Opladen: Barbara Budrich.
- Döbler, J.** (2019). *Prüfungsregime und Prüfungskulturen. Soziologische Beobachtungen zur internen Organisation von Hochschule*. Wiesbaden: Springer Verlag.
- Euler, D., & Hahn, A.** (2014). *Wirtschaftsdidaktik* (3rd ed.). Bern: UTB GmbH.
- George, D., & Mallery, P.** (2003). *SPSS for Windows Step by Step: A Simple Guide and Reference* (4th ed.). Boston: Allyn & Bacon.
- Gerdsmeier, G.** (1979). Induktiver Wirtschaftslehre-Unterricht: Begründung, Merkmale und Verlaufsmodelle. *Bildung und Erziehung*, 32, 25–42.

- 
- Gragg, C. I.** (2013). *Because Wisdom Can't Be Told*. Cambridge, MA: Harvard University Press.
- Gregor, S., & Hevner, A. R.** (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*, 337–355.
- HBS** (2021). *Harvard Business School Online. Academic Experience – The HBS Case Method*. Harvard Business School. <https://www.hbs.edu/mba/academic-experience/>
- Häcker, T.** (2007). *Portfolio: ein Entwicklungsinstrument für selbstbestimmtes Lernen. Eine explorative Studie zur Arbeit mit Portfolios in der Sekundarstufe 1* (2nd ed.). Baltmannsweiler: Schneider Hohengehren.
- Herreid, C. F.** (2011). Case Study Teaching. *New Directions for Teaching and Learning* (128), 31–40.
- Hevner, A. R., March, S. T., Park, J., & Ram, S.** (2004). Design Science in Information Systems Research. *MIS Quarterly*, 75–105.
- Hilgers-Sekowsky, J., & Huxold, S.** (2021). Kompetenzorientierte Lehre am Beispiel eines Real-Life-Business-Case. In C. Hattula, J. Hilgers-Sekowsky & G. Schuster (Eds.), *Praxisorientierte Hochschullehre* (pp. 15–25). Wiesbaden: Springer.
- Hoidn, S.** (2010). *Lernkompetenzen an Hochschulen fördern*. Wiesbaden: Springer.
- Holmström, J., Ketokivi, M., & Hameri, A. P.** (2009). Bridging Practice and Theory: A Design Science Approach. *Decision Sciences*, 40(1), 65–87.
- Kaiser, F.-J.** (1976). *Entscheidungsstraining: die Methoden der Entscheidungsfindung. Fallstudie, Simulation, Planspiel* (2nd ed.). Bad Heilbrunn: Klinkhardt.
- Kandlbinder, P.** (2014). Constructive Alignment in University Teaching. *HERDSA News*, 36(3), 5–6.
- Kosiol, E.** (1957). *Die Behandlung praktischer Fälle im betriebswirtschaftlichen Hochschulunterricht (case method): ein Berliner Versuch*. Berlin: Duncker & Humblot.
- Krosnick, J. A.** (2018). Questionnaire Design. In *The Palgrave Handbook of Survey Research* (pp. 439–455). Basingstoke: Palgrave Macmillan.

- May, H.** (2010). *Didaktik der Ökonomischen Bildung* (8th ed.). München: Oldenbourg.
- Mayring, P.** (2020). Qualitative Inhaltsanalyse. In G. Mey & K. Mruck (Eds.), *Handbuch qualitative Forschung in der Psychologie* (pp. 495–511). Wiesbaden: Springer.
- Melcher, H.** (2021). Einsatz der Harvard Case Method in der deutschen BWL für ein verstärkt interaktives, dialog-orientiertes Lernen. In J. Noller, C. Beitz-Radio, D. Kugelmann, S. Sontheimer & S. Westerholz (Eds.), *Studierendenzentrierte Hochschullehre* (pp. 131–148). Wiesbaden: Springer.
- Metzger, C., Schulmeister, R., & Martens, T.** (2012). Motivation und Lehrorganisation als Elemente von Lernkultur. *Zeitschrift für Hochschulentwicklung*, 7(3), 36–50.
- Müller, S.** (2007). Nah dran an der Praxis: Fallstudien bieten effektives Training von Managementkompetenzen. *Das Hochschulwesen*, 55(5), 154–160.
- Pilz, M.** (2001). Der Einsatz von Fallstudien zur Förderung des vernetzten Denkens im Wirtschaftslehreunterricht – Darstellung und Evaluation eines Projektes in der Berufsfachschule. *Wirtschaft und Erziehung*, 53(6), 193–200.
- Pilz, M., & Krüger, J.** (2013). *Vernetztes Denken und Entscheidungsfindung im Ökonomieunterricht. Eine Fallstudiensammlung*. Europa-Lehrmittel.
- Prenzel, M., Kristen, A., Dengler, P., Ettle, R., & Beer, T.** (1996). Selbstbestimmt motiviertes und interessiertes Lernen in der kaufmännischen Erstausbildung. *Zeitschrift für Berufs und Wirtschaftspädagogik*, 108–127.
- Rebmann, K., Tenfelde, W., & Schlömer, T.** (2011). *Berufs- und Wirtschaftspädagogik: eine Einführung in Strukturbegriffe* (4th ed.). Wiesbaden: Gabler.
- Riebe, K., & Scheffler, R.** (2017). Lernprozesse und Prüfung: das E-Portfolio zur flexiblen Gestaltung der Lehre. In A. Mörth & U. Elsholz (Eds.), *Portfolios in der wissenschaftlichen Weiterbildung. Handreichung der wissenschaftlichen Begleitung des Bund-Länder-Wettbewerbs. Aufstieg durch Bildung: offene Hochschulen*. [https://www.pedocs.de/volltexte/2017/14892/pdf/Moerth\\_et\\_al\\_2017\\_Portfolios\\_in\\_der\\_wissenschaftlichen\\_Weiterbildung.pdf](https://www.pedocs.de/volltexte/2017/14892/pdf/Moerth_et_al_2017_Portfolios_in_der_wissenschaftlichen_Weiterbildung.pdf)
- Romme, A. G. L.** (2003). Making a Difference: Organization as Design. *Organization Science*, 14(5), 558–573.

**Schaper, N., Reis, O., Wildt, J., Horvath, E., & Bender, E.** (2012). *Fachgutachten zur Kompetenzorientierung in Studium und Lehre*. HRK Projekt nexus.

**Schindler, C.** (2015). *Herausforderung Prüfen: Eine fallbasierte Untersuchung der Prüfungspraxis von Hochschullehrenden im Rahmen eines Qualitätsentwicklungsprogramms*. [https://www.researchgate.net/publication/305800358\\_Herausforderung\\_Prufen\\_Eine\\_fallbasierte\\_Untersuchung\\_der\\_Pruferungspraxis\\_von\\_Hochschullehrenden\\_im\\_Rahmen\\_eines\\_Qualitaetsentwicklungsprogramms/references](https://www.researchgate.net/publication/305800358_Herausforderung_Prufen_Eine_fallbasierte_Untersuchung_der_Pruferungspraxis_von_Hochschullehrenden_im_Rahmen_eines_Qualitaetsentwicklungsprogramms/references)

**Stratmann, J., Preussler, A., & Kerres, M.** (2009). Lernerfolg und Kompetenz bewerten. Didaktische Potenziale von Portfolios in Lehr-/Lernkontext. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 18, 119.

**Weinert, F.** (2014). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. Weinert (Ed.), *Leistungsmessungen in Schulen* (3rd ed., pp. 17–31). Weinheim – Basel: Beltz Verlag.

**Wilbers, K.** (2014). *Wirtschaftsunterricht gestalten: eine traditionelle und handlungsorientierte Didaktik für kaufmännische Bildungsgänge* (2nd ed.). Berlin: epubli.

**Wolff, K.** (1992). Die Fallstudie als Unterrichtsmethode. *Wirtschaft und Erziehung*, 44(10), 324–332.

## Acknowledgement

I would like to thank Prof. Dr. Christian Busse for the great cooperation in the development and implementation of the teaching concept and the guidance throughout my research. I would also like to thank Ms. Maren Sandmann and Ms. Iris Ruhnau for collecting and sharing their interview data.

## Author



Eva OSTERESCH || Carl von Ossietzky Universität Oldenburg,  
Chair of Sustainability and Supply Chain Management || Ammerländer Heerstr. 114–118, D-26129 Oldenburg

<https://uol.de/en/bwl-bwp/sscm>

[eva.osteresch@uni-oldenburg.de](mailto:eva.osteresch@uni-oldenburg.de)



# Konstruktion und Durchführung eines Video-Portfolios

## Eine innovative schulpraktische Prüfung

### Zusammenfassung

In diesem Werkstattbericht wird die Konstruktion und Durchführung einer schulpraktischen Abschlussprüfung in Form eines Video-Portfolios beschrieben. Im Rahmen dieser Prüfung reichen die Studierenden die Planung einer Lektion, ein Video der Umsetzung dieser Lektion sowie schriftliche Analysen der videografierten Lektion ein. Die Studierenden haben Zugang zu den Beurteilungskriterien, zu Geräten und zu einem Online-Kurs, der Hinweise zum Videografieren und zur schriftlichen Umsetzung bietet. Das Video-Portfolio wird von zwei Expert\*innen unabhängig voneinander beurteilt. Nach dem Noteneintrag werden alle personenbezogenen Daten gelöscht und die Studierenden erhalten einen Evaluationsfragebogen.

### Schlüsselwörter

summative Prüfung, Lehrer\*innenbildung, schulpraktische Studien, professionelle Unterrichtswahrnehmung, Videoanalyse

---

1 E-Mail: [david.bisang@fhnw.ch](mailto:david.bisang@fhnw.ch)



## **Construction and implementation of a video portfolio – An innovative practice-based examination**

### **Abstract**

This workshop report describes the design and implementation of a video portfolio as a practice-based final examination. As part of this examination, students submit a lesson plan, a video of the realisation of that lesson, and a written analysis of the videotaped lesson. Students have access to assessment criteria, loan materials, and an online course that provides guidance on videography and writing. The video portfolio is independently assessed by two experts. After the grading, all personal data is deleted, and students receive an evaluation questionnaire.

### **Keywords**

summative assessment, teacher education, practice-based studies, professional vision, video analysis

## **1 Eine videobasierte Abschlussprüfung**

Wenn eine Institution Lehrpersonen ausbildet, sollte sie die Erreichung der Kompetenzziele überprüfen (TERHART, 2007). In der Lehrpersonenbildung gibt es am Ende der schulpraktischen Ausbildung von Lehramtsstudierenden deshalb häufig Lehrproben oder Prüfungslektionen, bei denen Expert\*innen vor Ort den Unterricht beobachten, beurteilen und anschließend mit den Lehramtsstudierenden besprechen. Dazu reichen Lehramtsstudierende vorher oft schriftliche Planungsdokumente ein, die in die Bewertung einfließen (z. B. BÄUERLEIN et al., 2019).

Im vorliegenden Werkstattbericht geht es um eine alternative schulpraktische Abschlussprüfung: das Video-Portfolio. Bei diesem wird auf die Arbeit mit Unterrichtsvideos anstatt auf Prüfungslektionen vor Ort gesetzt. Die Studierenden reichen die schriftliche Planung einer Lektion, die Umsetzung der Lektion auf Video und zwei schriftliche Analysen des videografierten Unterrichts ein, wobei die Analysen einer vorgegebenen Heuristik folgen.

Im vorliegenden Beitrag geht es um die Konstruktion und Durchführung des Video-Portfolios. Es werden folgende Themen behandelt: theoretische Begründung, Aufbau bzw. Bestandteile der Prüfung, die die den Studierenden zur Verfügung gestellten Hilfsmittel, datenschutzrechtliche und technische Herausforderungen und das Beurteilungsverfahren.

## **2 Professionalisierung und professionelle Unterrichtswahrnehmung**

Kompetenz umfasst nicht nur „zugrundeliegendes Wissen und vorhandene Fertigkeiten [...], sondern [auch] deren erfolgreiche Anwendung“ (FREY, 2014, S. 713). Bei einer Kompetenzüberprüfung sollte also ein gewisses Handlungsspektrum erfasst werden, nicht nur ein isolierter Bestandteil einer Kompetenz. So kann gewährleistet werden, dass nicht nur theoretisches Wissen überprüft wird, sondern auch der Transfer des Wissens auf Praxissituationen. Diese Transferleistung ist eine zentrale Voraussetzung für die professionelle Handlungskompetenz, die von Lehrpersonen erwartet wird (ebd.). Pädagogische Hochschulen sind vor allem im Bereich der schulpraktischen Ausbildung gefordert, die professionellen Handlungskompetenzen von Studierenden zu fördern, damit sie als Lehrpersonen qualitativ hochwertigen Unterricht gestalten können, bei dem die Schüler\*innen möglichst viel lernen.

Eine wichtige Grundlage der professionellen Handlungskompetenz von Lehrpersonen ist die „professionelle Unterrichtswahrnehmung“. Diese umfasst die Beobachtung und Interpretation von Unterrichtssituationen basierend auf dem Wissen über Lehr- und Lernprozesse (JAHN et al., 2014, S. 172). Um die professionelle Unterrichtswahrnehmung auszubilden, werden in der Lehrer\*innenbildung zunehmend Videos verwendet. Durch Videoaufnahmen erhalten Studierende einen breiten Zugang zu verschiedenen Unterrichtssituationen, wobei die Aufnahmen eine hohe Anschaulichkeit und Realitätsnähe gewährleisten (KRAMMER & REUSSER, 2005). Außerdem ermöglichen sie es, eine Verbindung zwischen den theoretischen Inhalten der Ausbildung und der Unterrichtstätigkeit herzustellen. Denn sie fördern die selektive Aufmerksamkeit und die wissensbasierte Analyse des Gesehenen. Durch Videos kann man lernen, die Lernprozesse der Schüler\*innen in den Blick zu nehmen, sein eigenes Verhalten besser zu verstehen und die Konsequenzen der eigenen

Entscheidungen zu begründen (GAUDIN & CHALIÈS, 2015). Unterrichtsvideos sind also ein gutes Instrument, um die professionelle Unterrichtswahrnehmung von Lehramtsstudierenden zu fördern.

In den Berufspraktischen Studien im Studiengang Sekundarstufe I der Pädagogischen Hochschule FHNW (PH FHNW) wird daher mit Unterrichtsanalysen anhand von Videoaufzeichnungen gearbeitet. Ziel ist es dabei, dass die Studierenden erfahren, wie sie mit dem fachdidaktischen und erziehungswissenschaftlichen Wissen und Können, das sie im Verlauf ihres Studiums erwerben, ihre eigenen Handlungen im Unterricht analysieren können. Die Erreichung dieses Ziels soll mit einer Abschlussprüfung, dem sogenannten Video-Portfolio, überprüft werden. Im nächsten Kapitel werden die Bestandteile und Grundlagen dieser videobasierten Prüfung beschrieben, wie sie im Studiengang Sekundarstufe I der PH FHNW durchgeführt wird.

Die meisten Dokumente, die dabei genannt werden, sind auf der folgenden Webseite öffentlich zugänglich: <https://www.fhnw.ch/plattformen/praxisportal-sek1/videoportfolio/>. Dokumente, die in diesem Artikel erwähnt werden, aber nicht auf der Website zu finden sind, können auf Anfrage zur Verfügung gestellt werden.

### **3 Die Bestandteile des Video-Portfolios**

Im Video-Portfolio werden drei Kompetenzbereiche überprüft, die in der schulpraktischen Ausbildung fokussiert werden. Deshalb erhalten die Studierenden folgende Aufgabenstellung (Berufspraktische Studien Sekundarstufe I der PH FHNW, S. 5): Sie sollen mit dem Video-Portfolio zeigen, ...

- „dass sie eine Lektion fachwissenschaftlich, fachdidaktisch und erziehungswissenschaftlich kompetent planen können,
- dass sie diese Lektion gemäß der drei Unterrichtsqualitätsdimensionen ‚kognitive Aktivierung‘, ‚Klassenführung‘ und ‚konstruktive Unterstützung‘ durchführen können,
- und dass sie aus dem durchgeführten Unterricht Schlüsselsequenzen auswählen und analysieren können, um aus den Erkenntnissen begründete Handlungs-

alternativen abzuleiten. Für die Analysen sind theoretische und empirische Wissensbestände aus den Fachdidaktiken, den Erziehungswissenschaften und den Fachwissenschaften einzubeziehen.“

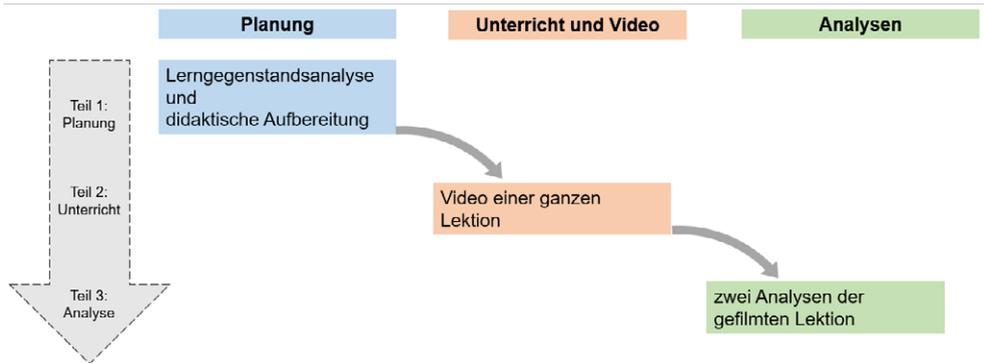


Abb. 1: Bestandteile des Video-Portfolios

Die Studierenden reichen folglich die schriftliche Planung einer Unterrichtslektion, das Video dieser Lektion und zwei schriftliche Analysen ein (vgl. Abb. 1). Damit sie diese Prüfung erfolgreich absolvieren können, erhalten die Studierenden die Beurteilungskriterien und die formalen Anforderungen.

### 3.1 Die Beurteilungskriterien

Die Beurteilungskriterien wurden basierend auf aktueller wissenschaftlicher Literatur und institutsspezifischen Konzepten (z. B. zur Planung) erarbeitet. Für das Frühjahrssemester 2021 wurden die Kriterien, die seit 2017 benutzt wurden, komplett überarbeitet. Zu jedem Beurteilungskriterium wurden mögliche Indikatoren formuliert, die den Studierenden und Beurteilenden bei der Interpretation der Kriterien helfen.

Das Beurteilungsraster umfasst die folgenden vier Bereiche:

1. Planung: 8 Beurteilungskriterien
2. Durchf hrung/Unterrichtslektion: 13 Beurteilungskriterien
3. Analyse: 8 Beurteilungskriterien
4. Formalien: 2 Beurteilungskriterien

Jedes Beurteilungskriterium wird von den Beurteilenden auf einer vierstufigen Skala beurteilt: 0 Punkte (nicht erf llt), 1 Punkt (ansatzweise erf llt), 2 Punkte (weitgehend erf llt), 3 Punkte (erf llt). Nachfolgend werden die einzelnen Bereiche kurz erl utert.

### **3.2 Anforderungen an die Planung**

Das Planen ist eine zentrale Kompetenz von Lehrpersonen. Sie erfordert die Integration des fachwissenschaftlichen, fachdidaktischen und erziehungswissenschaftlichen Wissens, das im Rahmen der Ausbildung erworben wurde. Gleichzeitig ist die Planung die Grundlage f r lernwirksamen Unterricht (T NZER & LAUTERBACH, 2010). Die acht Beurteilungskriterien der Planung orientieren sich an fachdidaktischen und erziehungswissenschaftlichen Theorien und Forschungserkenntnissen sowie an internen Planungsheuristiken.

Die Studierenden m ssen unter anderem die Auswahl des Lerngegenstands und dessen Relevanz f r die Sch ler\*innen begr nden, den Lerngegenstand sachlich vertieft analysieren, die Voraussetzungen der Sch ler\*innen aufzeigen, Lernziele formulieren und aus diesen Erkenntnissen didaktisch bedeutsame Folgerungen f r die Gestaltung der Lektion ableiten. Zus tzlich m ssen sie den geplanten Ablauf der Lektion in Form einer Tabelle darstellen, wobei die Aktivit ten der Sch ler\*innen in Bezug zu den Lernzielen im Vordergrund stehen sollen.

### **3.3 Anforderungen an die Unterrichtslektion**

Spätestens seit den Arbeiten von John HATTIE (2015) ist klar, dass die Lehrperson ein wichtiger Faktor für erfolgreiche Lernprozesse der Schüler\*innen ist. Und deshalb ist die Frage zentral, was guten Unterricht ausmacht. In der empirischen Unterrichtsforschung konnten wiederholt drei zentrale Dimensionen von Unterrichtsqualität festgestellt werden: kognitive Aktivierung, Klassenführung und konstruktive Unterstützung (z. B. PRAETORIUS et al., 2018). Die Beurteilungskriterien des Unterrichts orientieren sich an diesen drei Dimensionen und an weiteren fachdidaktischen und erziehungswissenschaftlichen Forschungserkenntnissen.

### **3.4 Anforderungen an die Analyse**

Die Studierenden müssen Sequenzen auswählen, bei denen etwas Zentrales für das Lernen der Schüler\*innen geschieht. Die Sequenzen können lernförderliche oder lernhinderliche Handlungen enthalten. Nach der entsprechenden Auswahl werden die folgenden vier Analyseschritte durchlaufen (Orientierung an SANTAGATA & GUARINO, 2011):

1. Die Schlüsselsequenz wird im Rahmen der Lektion situiert und daraus ein Thema für die Analyse festgelegt.
2. Zum festgelegten Thema werden theoretische und empirische Erkenntnisse zusammengefasst und darauf basierend Indikatoren abgeleitet und beschrieben, anhand derer die Sequenz analysiert wird.
3. Das Handeln der Lehrperson und dessen Wirkung auf die Schüler\*innen wird beschrieben und auf der Grundlage der Indikatoren interpretiert.
4. Aus der Analyse werden Handlungsalternativen abgeleitet, die mit den theoretischen bzw. empirischen Erkenntnissen zusammenhängen.

Diese vier Analyseschritte werden in verschiedenen Seminaren der Berufspraktischen Studien Sek I der PH FHNW geübt. Den Studierenden werden noch weitere Unterlagen und Hilfsmittel zur Verfügung gestellt – diese werden im nächsten Kapitel kurz erläutert.

## 4 Weitere Unterlagen und Hilfsmittel

Neben dem Beurteilungsraster erhalten die Studierenden ein Dokument, die sogenannten Ausführungsbestimmungen, in welchem sie alle notwendigen organisatorischen und formellen Informationen nachlesen können. Sie erfahren, wie der Beurteilungsprozess abläuft und worauf sie bei der Gewährleistung des Datenschutzes achten müssen. Dieses Dokument dient den Studierenden bei der Planung der Prüfung, deshalb enthält es auch einen prototypischen zeitlichen Ablauf.

Da es von den Studierenden trotz der erwähnten Unterstützung immer wieder Fragen gibt, wurde auf der Plattform „Moodle“ ein Online-Kurs mit den folgenden Inhalten erstellt:

- Termine zur Online-Beratung über ein Videokonferenz-Tool
- ein Forum, in welchem Fragen gestellt werden können
- ein freiwilliger Plagiats-Check mit der Software Turnitin

Der freiwillige Plagiats-Check wurde eingeführt, da in der Vergangenheit Plagiate aufgedeckt wurden. Die Studierenden werden so sensibilisiert, ihre Arbeit vor der definitiven Einreichung zu prüfen und kritische Stellen anzupassen.

## 5 Das Beurteilungsverfahren

Wenn die Studierenden das Video-Portfolio erstellt und eingereicht haben, wird es von zwei Expert\*innen beurteilt – von einer Fachperson aus der Erziehungswissenschaft und von einer Fachperson aus der Fachdidaktik. Da die Studierenden in beiden Bereichen ausgebildet werden, wird es als sinnvoll erachtet, dass die Beurteilung beide Perspektiven berücksichtigt. Die Expert\*innen beurteilen das Video-Portfolio unabhängig voneinander und müssen dabei das Beurteilungsraster vollständig ausfüllen.

Aus den verteilten Punkten resultiert eine ungerundete Note, die anhand einer bestimmten Formel berechnet wird. Ein Austausch zwischen den beiden Beurteilenden ist nur vorgesehen, wenn die ungerundeten Noten eine Notenstufe oder mehr voneinander abweichen oder wenn mindestens ein\*e Expert\*in eine ungenügende

Note vergibt. Die beiden Beurteilenden erhalten in beiden Fällen die Möglichkeit, aufgrund des Gesprächs ihre Noten anzupassen. Wenn sie dies im Fall der Notendifferenz nicht tun und die Noten weiterhin eine Notenstufe oder mehr voneinander abweichen, wird das Video-Portfolio von zwei weiteren Expert\*innen begutachtet.

Den Expert\*innen wird jedes Semester die Möglichkeit geboten, freiwillig an einem Workshop teilzunehmen, in welchem anhand von Beispielen diskutiert wird, welche Beurteilungskriterien wie interpretiert werden oder welche Beurteilungsfehler eine Rolle spielen können. Zusätzlich zu diesen Workshops erhalten die Expert\*innen ein Manual, das unter anderem Ausführungen zur Interpretation der Beurteilungskriterien, Anwendungsbeispiele sowie gängige Beurteilungsfehler enthält.

Wenn der Beurteilungsprozess abgeschlossen ist, erhalten die Studierenden den Beurteilungsbogen, in welchem sie sehen, welche Expert\*in wie viele Punkte pro Beurteilungskriterium vergeben hat.

## **6 Technische Herausforderungen und Datenschutz**

Qualitativ gute Videoaufnahmen zu erstellen, ist herausfordernd. Damit die Schüler\*innen gut sichtbar und die Gespräche gut hörbar sind, wurde für die Studierenden ein Erklärvideo zu diesem Thema erstellt. In diesem erhalten sie Informationen, worauf sie bei den Aufnahmen achten sollten.

Weiter werden die Studierenden auch mit geeignetem technischem Material unterstützt. So können sie auf diverse Ausleihangebote der Institution zurückgreifen und Geräte wie Funkmikrofone oder Kameras ausleihen.

Der Datenschutz ist im Rahmen dieser Prüfung sehr wichtig, da die Schüler\*innen auf den Videos erkennbar sind. Um die gültigen Datenschutzgesetze einzuhalten, müssen die Studierenden von den Schüler\*innen vor dem Filmen Einverständniserklärungen einholen. Außerdem sind die Studierenden und die PH FHNW dazu verpflichtet, alle personenbezogenen Daten nach Erhalt der Prüfungsnote zu löschen und das Löschen der Daten schriftlich zu bestätigen.

## 7 Kritische Betrachtung des Prüfungsformats

Das Video-Portfolio überprüft drei wichtige Kompetenzbereiche von Lehrpersonen: Die Planungs-, Unterrichts- und Analysekompetenz. Und es hat, da es am Ende der Ausbildung angesiedelt ist, eine selektive Funktion (MAAG MERKI & WERNER, 2014). Deshalb ist es wichtig, die Prüfung regelmäßig zu evaluieren und konzeptionell weiterzuentwickeln. Da im Frühlingssemester 2021 ein neues Beurteilungsraster eingeführt wurde, liegen derzeit erst wenige Evaluationsdaten vor, die noch keine zuverlässigen Aussagen erlauben. In diesem Kapitel wird deshalb eine kritische Betrachtung des Video-Portfolios basierend auf theoretischen und empirischen Erkenntnissen vorgenommen.

Videos ermöglichen es, die Komplexität des Unterrichts sichtbar zu machen, und sie haben den Vorteil, dass videobasierte Analysen nicht unter Handlungsdruck erstellt werden müssen (BLOMBERG et al., 2013). Diese Erkenntnis hat Vorteile für die Studierenden und die Beurteilenden: Studierende haben Zeit, ihr Professionswissen einzusetzen und ihre Handlungen vertieft zu analysieren. Und die Beurteilenden können verschiedene Sequenzen des Unterrichts wiederholt anschauen und die Beurteilung der Kriterien ebenfalls ohne Zeitdruck vornehmen. Gleichzeitig ist es im Falle einer Uneinigkeit der Beurteilenden möglich, die Inhalte nochmals anzuschauen und bestimmte Aspekte der Beurteilung zu besprechen.

Doch die Arbeit mit Videos hat auch Herausforderungen: Es stellen sich Fragen der Reaktivität (PRAETORIUS et al., 2017) und der Videoqualität. So lässt die Bild- und Tonqualität der Videos, die die Studierenden im Rahmen ihres Video-Portfolios erstellen, teilweise zu wünschen übrig. Das wiederum kann sich negativ auf die Beurteilung auswirken. Außerdem ist bei videobasierten Prüfungen zu beachten, dass die Beurteilungen hoch inferent sind, denn einige der zu beurteilenden Aspekte (z. B. kognitive Aktivierung) sind nicht direkt beobachtbar. Das macht die Beurteilungen „anfällig“ für Beurteilungsfehler (Rater Bias) (LOTZ et al., 2013). Einigen dieser potenziellen Nachteile wird versucht vorzubeugen, indem Studierende während der Ausbildung Übungsmöglichkeiten haben, Videos zu erstellen und diese mit Dozierenden zu besprechen. Und um Beurteilungsfehlern vorzubeugen, können die Expert\*innen Workshops besuchen und sie erhalten ein Manual.

Diese Ausführungen zeigen, dass das Video-Portfolio ein Instrument ist, das eine realitätsnahe Beurteilung zentraler Kompetenzen von angehenden Lehrpersonen

ermöglicht. Gleichzeitig tauschen sich Dozierende und Professurleiter\*innen aufgrund des Video-Portfolios vermehrt über die Ausbildung aus, was zu einer Weiterentwicklung des Studienkonzepts führen kann. In diesem Sinne ist das Video-Portfolio nebst anderen Formen schulpraktischer Prüfungen eine innovative Alternative.

## 8 Literaturverzeichnis

**Aeppli, J., & Lötscher, H.** (2016). EDAMA – Ein Rahmenmodell für Reflexion. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 34(1), 78–97.

**Bäuerlein, K., Senn, S., & Fraefel, U.** (2019). Die Beurteilung berufspraktischer Kompetenzen am Ende der Lehrpersonenausbildung in der Deutschschweiz. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 37, 401–420.

**Begrich, L., Fauth, B., Kunter, M., & Klieme, E.** (2017). Wie informativ ist der erste Eindruck? Das Thin-Slices-Verfahren zur videobasierten Erfassung des Unterrichts. *Zeitschrift für Erziehungswissenschaft*, 20(1), 23–47.

**Berufspraktische Studien Sekundarstufe I der PH FHNW** (2021). *Ausführungsbestimmungen: Video-Portfolio Sek I*. Windisch: PH FHNW.

**Blomberg, G., Renkl, A., Gamoran Sherin, M., Borko, H., & Seidel, T.** (2013). Five research-based heuristics for using video in pre-service teacher education. *Journal for Educational Research Online*, 5 (1), 90–114.

**Frey, A.** (2014). Kompetenzmodelle und Standards in der Lehrerbildung und im Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 712–744). Münster: Waxmann.

**Gaudin, C., & Chaliès, S.** (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review*, 16, 41–67.

**Hattie, J.** (2015). *Lernen sichtbar machen* (W. Beywl & K. Zierer, Hrsg.; 3., erweiterte Auflage). Baltmannsweiler: Schneider Hohengehren.

**Jahn, G., Stürmer, K., Seidel, T., & Prenzel, M.** (2014). Professionelle Unterrichtswahrnehmung von Lehramtsstudierenden: Eine Scaling-up Studie des Observer-Projekts. *Zeitschrift Für Entwicklungspsychologie und Pädagogische Psychologie*, 46(4), 171–180.

- Kleinknecht, M., Schneider, J., & Syring, M.** (2014). Varianten videobasierter Lehrens und Lernens in der Lehrpersonenaus- und -fortbildung. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 32(2), 210–220.
- Krammer, K., & Reusser, K.** (2005). Unterrichtsvideos als Medium der Aus- und Weiterbildung von Lehrpersonen. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 23(1), 35–50.
- Lotz, M., Gabriel, K., & Lipowsky, F.** (2013). Niedrig und hoch inferente Verfahren der Unterrichtsbeobachtung. *Zeitschrift für Pädagogik*, 59, 357–380.
- Maag Merki, K., & Werner, S.** (2014). In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 745–763). Münster: Waxmann.
- Praetorius, A.-K., McIntyre, N. A., & Klassen, R. M.** (2017). Reactivity effects in video-based classroom research: an investigation using teacher and student questionnaires as well as teacher eye-tracking. *Zeitschrift für Erziehungswissenschaft*, 20, 49–74.
- Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P.** (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM*, 50(3), 407–426.
- Santagata, R., & Guarino, J.** (2011). Using video to teach future teachers to learn from teaching. *ZDM*, 43(1), 133–145.
- Seidel, T., & Thiel, F.** (2017). Standards und Trends der videobasierten Lehr-Lernforschung. *Zeitschrift für Erziehungswissenschaft*, 20(1), 1–21.
- Tänzer, S., & Lauterbach, R.** (2010). *Sachunterricht begründet planen: Bedingungen, Entscheidungen, Modelle*. Bad Heilbrunn: Klinkhardt.
- Terhart, E.** (2007). Standards in der Lehrerbildung – eine Einführung. *Unterrichtswissenschaft*, 35(1), 2–14.

## Autor\*innen



MSc David BISANG || Pädagogische Hochschule FHNW, Institut  
Sekundarstufe I & II || Bahnhofstraße 6, CH-5210 Windisch

[david.bisang@fhnw.ch](mailto:david.bisang@fhnw.ch)



Dr. Kerstin BÄUERLEIN || Pädagogische Hochschule FHNW, Ins-  
titut Sekundarstufe I & II || Bahnhofstraße 6, CH-5210 Windisch

[kerstin.baeuerlein@fhnw.ch](mailto:kerstin.baeuerlein@fhnw.ch)



Prof. Dr. Corinne WYSS || Pädagogische Hochschule FHNW, Ins-  
titut Sekundarstufe I & II || Bahnhofstraße 6, CH-5210 Windisch

[corinne.wyss@fhnw.ch](mailto:corinne.wyss@fhnw.ch)



# Simulationen in virtuellen Lernumgebungen: Welche Vor- und Nachteile haben avatarbasierte und videokamerabasierte Formate bei der Durchführung von Rollenspielen?

## Zusammenfassung

Effektives kommunikatives Handeln wird zur Bewältigung von Aufgaben im Lehrberuf benötigt. Die Erprobung solcher Fähigkeiten kommt im Lehramtsstudium jedoch oft zu kurz. Dieser Beitrag stellt daher ein innovatives Lehr-Lern-Format vor, bei dem performanzbasiert Simulationen in Form von Rollenspielen in zwei unterschiedlichen (avatarbasiert vs. videokamerabasiert) virtuellen Lernumgebungen durchgeführt wurden. Beide Lernumgebungen wurden anhand ihrer Vor- und Nachteile miteinander verglichen. Insbesondere durch das Gefühl des Eintauchens in die avatarbasierte Lernumgebung können die Studierenden ihre kommunikativen Kompetenzen im geschützten Rahmen nahezu realitätsgetreu erproben.

## Schlüsselwörter

Simulationen, kommunikative Kompetenz, virtuelle Lernumgebung, avatarbasierte Interaktionen, Handlungsfähigkeit

---

<sup>1</sup> E-Mail: [julia.fecke@zfl.uni-giessen.de](mailto:julia.fecke@zfl.uni-giessen.de)



## **Simulations in virtual learning environments: A comparison of avatar-based and video-camera-based roleplay scenarios**

### **Abstract**

Effective communication is needed to manage tasks in the teaching profession. However, teacher training programs often fail to test such skills. This paper presents an innovative teaching-learning format in which simulations in the form of role-plays were conducted in two different (avatar-based vs. video-camera-based) virtual learning environments. The two learning environments were compared to determine their advantages and disadvantages. In particular, the feeling of immersion in the avatar-based learning environment allows students to test their communicative skills in a safe setting in a manner that is almost true to reality.

### **Keywords**

simulations, communicative competence, virtual learning environment, avatar-based interaction, vocational skills

## **1 Einführung**

Gegenwärtig gerät die berufliche Lehrkompetenz verstärkt in den Fokus (Überblick: SCHEIDIG, 2020): nicht nur politische Stimmen fordern höhere Praxisanteile in der Lehrkräfteausbildung, auch Lehramtsstudierende monieren den fehlenden Praxisbezug in der Ausbildung (MAKRINUS, 2013). Besonders für Lehramtsstudierende sind handlungsnaher (z. B. Micro-Teaching als Lehrmethode: HATTIE, 2015, S. 134) und multimediale Zugänge (KERRES, 2018) wesentlich, um sich frühzeitig mit konkreten Unterrichtsansforderungen, wie Gesprächsführung und Digitalisierung, auseinanderzusetzen. Gleichzeitig nimmt die Forderung nach innovativen digitalen Lehr-Lern-Formaten zur Förderung handlungsnaher Kompetenzen von angehenden Lehrkräften zu (KMK, 2019). Letzteres wurde durch die COVID-19-Pandemie verstärkt – der Druck, digitalisierte Lehre einzuführen sowie handlungsorientiert einzusetzen, ist hoch. Durch Fern-Studiengänge, höhere Anteile von Online-Kursen

und heterogenere Lernbedingungen wird die Relevanz digitaler Lehre auch in Post-Corona-Zeiten hoch bleiben.

Dieser Beitrag stellt daher zunächst ein innovatives Lehr-Lern-Format vor, in dem Simulationen in Form von Rollenspielen in virtuellen Lernumgebungen umgesetzt wurden, und vergleicht daran anknüpfend zwei unterschiedliche digitale Umsetzungen: avatarbasierte vs. videokamerabasierte Interaktionen. Die Lernumgebung ist hier an eine mündliche Prüfungssituation an Hochschulen zur Erfassung von kommunikativen Kompetenzen angelehnt (BRAUN et al., 2018) und wird performanzbasiert eingesetzt. Der vorliegende Beitrag liefert Erkenntnisse zu folgender Fragestellung: Welche Vor- und Nachteile haben avatarbasierte und videokamerabasierte Interaktionen als Lehr-Lern-Format bei der Durchführung von Simulationen aus Sicht der Studierenden? Die Erkenntnisse wurden aus Fokusgruppen mit Studierenden gewonnen, die im Anschluss an die Simulationen durchgeführt worden sind.

Der Beitrag gliedert sich wie folgt: In Kapitel 2 wird zunächst auf den Begriff der virtuellen Lernumgebung sowie Simulationen als Lehr-Lern-Methode eingegangen. In Kapitel 3 wird die Durchführung der Simulation in Form von Rollenspielen beschrieben. Kapitel 4 beschreibt Daten und Methoden zu den von uns behandelten Vor- und Nachteilen von avatarbasierten und videokamerabasierten Interaktionen. In Kapitel 5 werden die Ergebnisse der Fokusgruppen dargelegt und in Kapitel 6 diskutiert.

## **2 Theoretischer Rahmen: Simulationsbasiertes Lernen als Stärkung der Handlungsfähigkeit**

Während fachliche Kompetenzen vielfach erforscht werden, sind generische Kompetenzen, wie beispielsweise Kommunikationsfähigkeit, weniger stark in der hochschulischen Lehre verankert (KMK, 2017). Kommunikationsfähigkeit stellt einen wesentlichen Aspekt zentraler Handlungsfähigkeit zur Bewältigung von Aufgaben im Lehrberuf dar (BRAUN et al., 2018; GARTMEIER et al., 2012). Die Integration von Gesprächsführungskompetenzen in der Lehrkräfteausbildung ist daher dringend gefordert. Das Einüben wiederkehrender Gesprächssituationen mit Eltern,

Schüler\*innen, Vorgesetzten und Kolleg\*innen kann angehenden Lehrkräften Orientierung und Sicherheit geben (GARTMEIER et al., 2012). Hier wird mit der Simulation das Ziel verfolgt, das persönliche Handlungsrepertoire (ähnlich Core-Practice-Ansatz: FRAEFEL & SCHEIDIG, 2018) der Studierenden zu erweitern. Im Fokus der Simulation steht nicht der lineare Aufbau von Gesprächsführungskompetenzen, sondern das individuelle-experimentelle Erproben der theoretisch erlernten Praktiken sowie flexibel handlungsfähig zu werden (vgl. FRAEFEL & SCHEIDIG, 2018).

## 2.1 Performanzbasiertes Testen

Als theoretischer Hintergrund wird an das Kontinuumsmodell von BLÖMEKE, GUSTAFSSON und SHAVELSON (2015) angeknüpft. In dem Modell werden domänenspezifische Kompetenzen auf der einen Seite als kognitive Leistungsdispositionen verstanden. Demgegenüber steht das sichtbare Handeln als Fähigkeit in einer konkreten Anforderungssituation: die Performanz. Situationsspezifische Fähigkeiten (wahrnehmen, interpretieren, entscheiden) medieren zwischen der Disposition und der Performanz einer Person. Konkret bedeutet das, dass eine spezifische Situation erfasst, kognitiv verarbeitet und richtig gedeutet werden muss (BLÖMEKE et al., 2015). Die situationsspezifischen Fähigkeiten münden in einer situationsangemessenen Performanz (hier: adäquat kommunikativ handeln).

Um die kommunikative Kompetenz angehender Lehrkräfte zu überprüfen, ist es nicht ausreichend, nach der analytischen Tradition nur Wissensbestände abzufragen (z. B. mit schriftlichen Leistungstests). Hierzu bedarf es Instrumente, welche die Performanz in konkreten Handlungssituationen erfassen. Ziel solcher performanzbasierter Prüfungen ist es, die individuelle Handlungsfähigkeit in komplexen und authentischen Situationen unter standardisierten Rahmenbedingungen nachzuweisen (BRAUN et al., 2018, S. 34). Um ein solches Setting herzustellen, eignen sich Simulationen.

## 2.2 Simulationsbasiertes Lernen

Simulationsbasiertes Lernen als Lehr-Lern-Methode dient zur Heranführung an die Handlungsfähigkeit und kann Studierende dabei unterstützen, Herausforderungen ihres zukünftigen beruflichen Alltags wirklichkeitsnah unter Anleitung zu erpro-

ben (CHERNIKOVA et al., 2020). Die individuelle Handlungsfähigkeit kann in komplexen, authentischen Situationen unter standardisierten Rahmenbedingungen nachgewiesen werden.

Rollenspiele stellen eine Form von Simulationen dar und gelten als effektive Lehrmethode (CHERNIKOVA et al., 2020). Ziel von Simulationen ist die Generierung von Wissen zum besseren Verständnis verschiedener Rollen im Berufsalltag sowie die Weiterentwicklung von Handlungsfähigkeiten und Einstellungen (BRAUN et al., 2018). Durch die authentische Interaktion während der Simulation, die teilweise unerwartete Situationen hervorbringt, können Ideen und Denkmuster hinterfragt und weiterentwickelt werden, und der Nutzen von theoretischem Wissen und dessen Umsetzbarkeit wird erprobt (BRAUN et al., 2018; FRAEFEL & SCHEIDIG, 2018). Des Weiteren können Rollenspiele spätere berufliche Anforderungen repräsentieren und authentische Tätigkeitsbereiche erprobt werden (BRAUN et al., 2018). Bei dem hier vorgestellten Format liegt der Fokus auf der Erprobung von Gesprächsführungskompetenzen.

Auch wenn Rollenspiele als ein innovatives Format zur Erprobung der Handlungsfähigkeit von Studierenden angesehen werden, sind diese für die Studierenden oftmals mit Ängsten verbunden (KUNZE et al., 2016). KUNZE et al. (2016) schreiben dies dem Fehlen einer geschützten Lernumgebung zu. Gerade bei der Durchführung von Simulationen stellt eine geschützte Lernumgebung einen wesentlichen Aspekt für die Reduktion möglicher Ängste der Studierenden dar.

Aus diesem Grund wurde ein Format entwickelt, in dem Simulationen in virtuellen Lernumgebungen stattfinden.

### **2.3 Virtuelle Lernumgebungen für Simulationen**

In diesem Beitrag wird zwischen avatarbasierten und videokamerabasierten Interaktionen unterschieden. Beide Interaktionsformate finden in virtuellen Lernumgebungen, d. h. online, statt. Anzumerken ist, dass sich der Begriff „virtuell“ in den letzten Jahren stark gewandelt hat und seit der COVID-19-Pandemie 2020 häufig unscharf verwendet wird. Ursprünglich wurde der Begriff verwendet, um vom Computer simulierte Wirklichkeiten einer „künstlichen“ Welt zu beschreiben, in die sich Personen mithilfe eines entsprechenden Gerätes (z. B. Virtual-Reality-Brille)

hineinversetzen können. Mittlerweile wird der Begriff auch verwendet, wenn Veranstaltungen o. Ä. per Videokonferenz mit realen Personen stattfinden – dies trifft auch für die hier beschriebene videokamerabasierte Interaktion zu.

*Videokamerabasierte Interaktion* beschreibt die Durchführung in einem Videokonferenzformat. In diesem Format wird face-to-face per Videokamera synchron (live) interagiert. Videokamerabasierte Interaktion bedeutet für den vorliegenden Beitrag ein Gespräch zwischen zwei Personen, das eine Situation des Lehramtsalltags simuliert.

*Avatarbasierte Interaktion* beschreibt ebenfalls synchron stattfindende soziale Interaktionen in virtuellen Lernumgebungen. Die hier vorgestellte avatarbasierte Interaktion bezieht sich auf computergenerierte Umgebungen in 3D. In diesen Welten können sich mehrere Personen mittels eines Avatars innerhalb der Grenzen der Umgebung frei bewegen und miteinander kommunizieren und interagieren. Ein Avatar repräsentiert als grafische Darstellung (s. Kapitel 3.2) die eigene Person.

Allgemein kann ein virtueller Lernort eine geschützte Lernumgebung für Lehr-Lern-Prozesse sein. Das Ziel von virtuellen Lernumgebungen ist, u. a. Personen, vorzugsweise von verschiedenen Standorten aus, zusammenzubringen. Räumliche, zeitliche und konzeptuelle Einschränkungen aus der realen Welt können in der virtuellen überwunden werden. Die sogenannte Immersion – vor allem während der avatarbasierten Interaktion – der Lernumgebung stellt dabei einen essentiellen Aspekt dar. Sie beschreibt den durch die virtuelle Umgebung hervorgerufenen Effekt, wie real die Lernenden diese empfinden und in diese „eintauchen“ können (KICKMEIER-RUST & REIMANN, 2020).

Für Simulationen konnte in einer aktuellen internationalen Metaanalyse von CHERNIKOVA et al. (2020) ein großer Effekt ( $g = .85$ ) nachgewiesen werden: das Lernen in simulierten Lernumgebungen erleichtert den Erwerb von Kompetenzen (z. B. Problemlösen, kommunikative Kompetenzen). Zwar werden in der Hochschullehre vereinzelt Simulationen von Situationen des beruflichen Alltags genutzt (KICKMEIER-RUST & REIMANN, 2020; KUNZE et al., 2016), bei genauerer Betrachtung wird jedoch deutlich, dass es kaum systematische Erkenntnisse zu virtuellen Simulationen von Rollenspielen gibt (Ausnahme: KUNZE et al., 2016).

Da die Distanz und Anonymität in virtuellen Lernumgebungen größer ist als in Face-to-face-Szenarien, wird angenommen, dass dies eine gute Lerngelegenheit

darstellt, in der Lehramtsstudierende ihre Gesprächsführungskompetenzen erproben können. Daher soll vorliegender Beitrag analysieren, welche Vor- und Nachteile avatarbasierte und videokamerabasierte Interaktionen als Lehr-Lern-Format bei der Durchführung von Simulationen aus Sicht der Studierenden haben.

## 3 Durchführung der Simulationen

### 3.1 Rahmenbedingungen

Die Simulationen sind im Rahmen eines Projekts<sup>2</sup> an der JLU Gießen entstanden und wurden in drei Lehrveranstaltungen (allgemeines und berufliches Lehramt) durchgeführt.

Dabei wurde auf ein Instrument zur Simulation kommunikativer Kompetenzen für Lehramtsstudierende von BRAUN et al. (2018) zurückgegriffen. Das Instrument basiert auf HABERMAS' (1981) Kommunikationstheorie, bei der zwischen strategischer und verständigungsorientierter Kommunikation unterschieden wird. Kommunikation kann zur Erreichung eines verdeckten Gesprächsziels eingesetzt werden oder argumentativ offenbleiben (HABERMAS, 1981) – abhängig vom Kontext ist die Wahl des Kommunikationstypen wesentlich zur Zielerreichung. Das Instrument zur Testung kommunikativer Kompetenz kann als Performanztest eingesetzt werden. Hier wurde es performanzbasiert eingesetzt, d. h. zu Trainingszwecken. Die Simulation wurde nicht benotet, war aber zum Bestehen der jeweiligen Lehrveranstaltung erforderlich. Die Simulation diente als Grundlage einer Reflexion, welche bewertet wurde und in der die Rollenspiele reflektiert werden sollten.

Für die Gesprächssimulation fand ein Zweiergespräch zwischen einem/einer Studierenden und einer Schauspielperson statt. Die Schauspielperson simulierte das Gegenüber, z. B. Schüler\*in, Kolleg\*in oder Vorgesetzte\*r.

---

2 Die Gießener Offensive Berufliche Lehrerbildung (GOBeL) wird im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern aus Mitteln des Bundesministeriums für Bildung und Forschung gefördert. FKZ: 01JA2039.

### 3.2 Durchführung avatarbasierter und videokamerabasierter Interaktion

Vor der Durchführung wurden die Studierenden der jeweiligen Veranstaltung randomisiert in avatarbasierte und videokamerabasierte Interaktion eingeteilt. D. h. nicht nur die Simulation, sondern die gesamte Lehrveranstaltung hat in dem jeweiligen Format stattgefunden.

Die videokamerabasierte Interaktion fand über ein herkömmliches Videokonferenztool (Cisco Webex) statt und die avatarbasierte Interaktion in einer virtuellen 3D Lern- und Arbeitswelt von TriCAT Spaces (<https://tricat-spaces.net/>). In TriCAT Spaces wurden die Teilnehmenden über einen Avatar verkörpert und konnten diesen im virtuellen Raum navigieren und in Echtzeit mit anderen Teilnehmenden kommunizieren. Wesentlich bei diesem Interaktionsformat ist die Immersion: TriCAT Spaces bietet durch sein multimodales Interface die Beanspruchung mehrerer Sinne, wodurch bei den Nutzer\*innen das Gefühl direkter Eingebundenheit und Anwesenheit vermittelt wird (DEDE, 2009). Eine VR-Brille wurde nicht eingesetzt.

Entsprechend dem hier verfolgten Ansatz zur Förderung der Kommunikationskompetenz wurde ein entsprechendes Lehr-Lern-Format entwickelt. D. h. die Studierenden haben in den Lehrveranstaltungen Theorien zur Professionalisierung sowie die Kommunikationstheorien (auch nach Habermas) kennengelernt und sich mit e-Learning-Elementen selbst erarbeitet. Vor den Simulationen mit den Schauspielern wurden zudem mit Tutorinnen (Peer-Learning-Konzept) Simulationen erprobt. Hier haben die Studierenden den/die Gesprächspartner\*in simuliert.

Der Ablauf der Simulationen mit den Schauspielern war wie folgt: Die Studierenden erhielten nach Betreten des virtuellen Vorbereitungsraums eine schriftliche Rollenspielanleitung mit einer für angehende Lehrkräfte bedeutsamen Gesprächssituation. Dafür bekam jede Testperson insgesamt zwei Rollenspiele vorgelegt: entweder zuerst ein strategisches oder ein verständigungsorientiertes – dies wurde vorab randomisiert zugeteilt. Anschließend hatten die Studierenden ca. zehn Minuten Zeit, sich auf das Rollenspiel vorzubereiten. Nach der Vorbereitungsphase simulierte der/die Studierende mit einer Schauspielperson das vorstrukturierte Gespräch. Die Schauspielern wurden dazu geschult und waren somit über die Situationsangemessenheit der jeweiligen Rollenspiele vertraut. Nach der Simulation begab sich

der/die Studierende wieder in den Vorbereitungsraum und erhielt das zweite Rollenspiel mit gleichem Ablauf wie oben beschrieben.

Innerhalb der avatarbasierten Lernumgebung wurden die Räume für die Simulation so vorbereitet, dass die Studierenden sich leicht einfinden konnten, und die Simulation möglichst authentisch stattfinden konnte, d. h. in den Räumen standen Tische und Stühle, sodass die Möglichkeit bestand, sich mit dem Avatar hinzusetzen. Im Vorbereitungsraum waren Tutorinnen anwesend, welche die Studierenden einwiesen. Zudem wurden auf dem Konferenztisch Monitore platziert: auf diesen wurden die Rollenspielinstruktionen für die Studierenden zur Vorbereitung geteilt (s. Abb. 1).



Abb. 1: Avatarbasierte Interaktion in TriCAT Spaces: Vorbereitungsraum der Studierenden mit Monitoren und Tutorinnen im Hintergrund

Nach 10-minütiger Vorbereitung haben sich die Studierenden in die Seminarräume mit den jeweiligen Schauspielern bewegt (s. Abb. 2). Dort führten die Studierenden mit einer Schauspielperson das Rollenspiel durch. Zudem war eine Beobachtungsperson anwesend.



Abb. 2: Avatarbasierte Interaktion in TriCAT Spaces: Simulation kommunikativer Kompetenzen: Studierender (l.), Schauspieler (r.), Rater (vorne)

Bei der Durchführung in der videokamerabasierten Umgebung (Cisco Webex) war der Ablauf ähnlich: Die Studierenden hatten ebenfalls einen Vorbereitungsraum und wurden dort in separate Sitzungsräume eingeteilt, um dort die Rollenspielinstruktion durch die Tutorinnen zu erhalten. „Raum“ bedeutet hier eine geschützte videokamerabasierte Interaktion mit einem Gegenüber. Die videokamerabasierte Interaktion ähnelt einer klassischen Videokonferenz, wie sie in Corona-Zeiten allgegenwärtig sind: Hier wurde online per Videokamera kommuniziert und die Studierenden und Schauspielern konnten sich face-to-face sehen. Für die Simulation wurden die Studierenden in einen weiteren Video-Raum weitergeleitet, in dem jeweils eine Schauspielperson sowie eine Beobachtungsperson anwesend waren.

## 4 Datenerhebung und methodisches Vorgehen

### 4.1 Daten

Zur Beantwortung der Fragestellung (Welche Vor- und Nachteile haben avatarbasierte und videokamerabasierte Interaktionen als Lehr-Lern-Format bei der Durchführung von Simulationen aus Sicht der Studierenden?) wurden die Studierenden zunächst in einem experimentellen Design in avatarbasierte ( $n=29$ ) und videokamerabasierte ( $n=32$ ) Interaktion randomisiert zugeteilt. Insgesamt haben 61 Studierende an dem Experiment teilgenommen; davon studieren 29 Personen allgemeines Lehramt, 21 Personen berufliches Lehramt auf Bachelorniveau und 11 Personen berufliches Lehramt auf Masterniveau. Von den teilnehmenden Studierenden sind 49,2 % ( $n=30$ ) weiblich, 47,5 % ( $n=29$ ) männlich und 3,3 % ( $n=2$ ) haben keine Angabe gemacht. Das durchschnittliche Alter der Studierenden liegt bei 27 ( $SD=4,44$ ).

### 4.2 Instrument Fokusgruppen

Im Anschluss an die Simulationen wurden zehn Fokusgruppen (max. 8 Studierende/Gruppe: randomisiert zugeteilt) durchgeführt. Die Fokusgruppe (moderierte Gruppendiskussion) ist ein qualitatives Erhebungsinstrument, bei dem Kleingruppen durch einen Stimulus zur Diskussion über ein Thema (hier: die Erfahrungen mit den Simulationen) angeregt werden. Durch den Austausch und den gruppendynamischen Prozess sollen eine Vertiefung der Teilnehmenden mit dem Gegenstand herbeigeführt werden und vielfältige Facetten einer Thematik beleuchtet werden (SCHULZ, MACK & RENN, 2012). Ähnlich wie bei qualitativen Einzelinterviews wird der Diskussionsprozess durch einen Leitfaden strukturiert. Der Leitfaden stellt sicher, dass alle wichtigen Aspekte zur Beantwortung der Fragestellung angesprochen werden, und erhöht die Vergleichbarkeit der Ergebnisse, falls mehrere Fokusgruppen durchgeführt werden (SCHULZ et al., 2012). Die Leitfragen zielten größtenteils auf Vor- und Nachteile der Lernumgebung ab: So wurde nach den Erwartungen vor der Durchführung der Simulationen gefragt, ob die Simulationen leicht oder schwergefallen sind, nach Besonderheiten im Umgang mit der jeweiligen Lernumgebung (Avatar vs. Video) und abschließend nach positivem und negativem Gesamteindruck.

Die Fokusgruppen wurden auf Video aufgezeichnet und anschließend transkribiert. Die Auswertung erfolgte per qualitativer Inhaltsanalyse (MAYRING, 2015). Bei der Strategie der Zusammenfassung wurde das Material zunächst paraphrasiert und generalisiert. Anschließend wurden die wichtigsten Aussagen der Studierenden zur Fragestellung gebündelt, reduziert und Kategorien zugeordnet. Jede Fokusgruppe wurde einzeln ausgewertet. Kategorien, die in nahezu allen Fokusgruppen wiederkehren, werden als Gesamtergebnis nachfolgend vorgestellt.

## 5 Ergebnisse

Die Vor- und Nachteile von avatarbasierten und videokamerabasierten Interaktionen sind in Tabelle 1 zusammenfassend dargestellt. Im Anschluss werden die jeweiligen Interaktionen genauer analysiert.

Tab. 1: Übersicht Gesamtergebnis der Fokusgruppen (Vor- und Nachteile)

<b>Vorteile avatarbasierte Interaktion</b>	<b>Nachteile avatarbasierte Interaktion</b>
<ul style="list-style-type: none"> <li>- Fehlen von Gestik und Mimik: Fokus auf das gesprochene Wort</li> <li>- Hinter dem Avatar verstecken</li> <li>- Authentizität (Immersion)</li> <li>- Entspannung</li> <li>- spielerische Atmosphäre</li> <li>- Notizen nutzen</li> </ul>	<ul style="list-style-type: none"> <li>- Fehlen von Gestik und Mimik: Reaktionen können nicht eingeschätzt werden, beeinträchtigter Redefluss</li> </ul>
<b>Vorteile videokamerabasierte Interaktion</b>	<b>Nachteile videokamerabasierte Interaktion</b>
<ul style="list-style-type: none"> <li>- Schwierige Gespräche (mit Machtgefälle) fallen leicht</li> <li>- Eigenes Zuhause: Sicherheit, weniger Anspannung</li> <li>- Distanz: leichtes Ausüben von Kritik</li> <li>- Notizen nutzen</li> </ul>	<ul style="list-style-type: none"> <li>- Gestik und Mimik: erschwerte Selbst- und Fremdeinschätzung, adäquate Nutzung, Koordination zwischen verbaler und nonverbaler Kommunikation</li> <li>- Simulation als Schauspiel</li> </ul>

Bei der *avatarbasierten Interaktion* wird als Vorteil insbesondere das spielerische Lernen hervorgehoben. Dadurch, dass man sich hinter dem Avatar „verstecken“ kann, kommt es zur Entspannung und zu einer Minderung der Angst bei der Durchführung der Rollenspiele. Außerdem heben die Studierenden hervor, dass sie „ungehemmter“ waren und so den Mut fassen konnten, mehr zu sagen:

„[...] man versteckt sich auch eher hinter so'nem Avatar. In der Realität achtet man ja dann auch auf seine Körperhaltung und so, und hier war das ja nochmal wie so'ne Möglichkeit, dass der andere einen gar nicht sieht, dass das glaube ich zur Erprobung vielleicht sogar auch teilweise 'n bisschen Sicherheit gegeben hat, in dem was man sagt, weil man ja eben nicht gesehen wird [...]“

Positiv wurde hervorgehoben, dass der Wegfall der nonverbalen Kommunikation den Fokus auf die verbale Kommunikation vereinfacht. Nicht auf Gestik und Mimik achten zu müssen, erleichtert die Interaktion mit dem Gegenüber. Auch hat der Ort, an dem sich die Studierenden befanden (meist das eigene Zuhause), aufgrund des gewohnten Umfeldes für Wohlbefinden und eine entspannte Haltung gesorgt.

Während einige den Wegfall der nonverbalen Kommunikation als positiv beschreiben, sahen einige einen Nachteil darin. Das Fehlen von Gestik und Mimik erschwert die Kommunikation mit dem Gegenüber, da diese nicht direkt gelesen werden kann. Die nonverbale Kommunikation übermittelt wichtige Informationen, die zur Deutung oder gar zur Unterstreichung des Gesagten wichtig sind. Insbesondere während kurzer Gesprächspausen wird dies als Nachteil beschrieben:

„[...] mir fehlt tatsächlich wirklich Mimik, Gestik, Körpersprache dabei, wenn man ähm da wirklich genau agieren soll und genau drauf reagieren soll, ähm nur durch die Stimme, oder beziehungsweise die dann ja auch noch unterbrochen ist durch ähm ja An- und Ausschalten des Mikrofons quasi, finde ich's schwierig da ähm 'n Gespräch richtig zu führen [...]“

Bei der *videokamerabasierten Interaktion* konnten folgende Vor- und Nachteile herausgearbeitet werden: auch hier gibt das eigene Zuhause Sicherheit und mindert die Anspannung:

„Positiv ist natürlich daran, dass ich einfach in meinem Zuhause bin, an einem Ort, an dem ich mich wohlfühle. Ich hab äh meine Crocs an und halt nicht die

schicken Schuhe.[...] Ich bin ja erstmal in 'nem Raum um mich herum, der mir Sicherheit gibt, der mir Halt gibt, äh wo ich zur Ruhe kommen kann.“

Bei diesem Format fällt es den Studierenden leicht, Gespräche mit Machtgefälle (z. B. Gespräch mit Schulleitung) zu führen sowie die eigene Meinung zu vertreten, ohne sich zu stark vom Gegenüber beeinflussen zu lassen. Auch das Ausüben von Kritik gegenüber Kolleg\*innen fällt in diesem Format leicht. Hier führten die Studierenden als Referenz Präsenzformate aus eigener Erfahrung auf. Als Grund wird die Distanz, die über das Format erreicht wird, genannt:

„Also ich glaub, das macht massiv was aus, find ich. Alleine Räumlichkeit ähm. Wenn ich in das Büro des Direktors oder der Direktorin komme, allein in diesen Raum reinzutreten. Und so die Begrüßung ist schon was völlig anderes wie so in digitalen Raum reinzukommen. Und es macht auch, find ich ähm mit der körperlichen Anspannung mit der Empfindung halt komplett was anderes.“

Als Nachteil wurde genannt, dass der Wegfall von Gestik und Mimik zu Verunsicherung führt. Viele Studierende wussten nicht, wie die nonverbale Kommunikation während der videokamerabasierten Interaktion genau funktioniert, wie nonverbale Kommunikation eingesetzt werden kann und wie viel das Gegenüber tatsächlich von den eigenen Bewegungen sieht (z. B. beim Gestikulieren mit den Händen) und wie die nonverbale Kommunikation (z. B. Blickkontakt) vom Gegenüber eingeschätzt wird:

„Was ich immer schwierig finde über Kamera äh ist, wenn man sich anguckt, also wenn ich jetzt einen auf dem Bildschirm angucke, dann ist das ja kein direkter Blickkontakt, weil dafür müsste ich ja in die Kamera gucken und das finde ich immer schwierig, weil das dann irgendwie so'n passiver Blick ist.“

Das Verhältnis von Selbst- vs. Fremdeinschätzung bezogen auf den Einsatz von Gestik und Mimik ist somit erschwert. Zudem erfordert es enorme Aufmerksamkeit, Gestik und Mimik adäquat zum Beibehalten des Gesprächsflusses einzusetzen. Während es bei der avatarbasierten Interaktion von Beginn an klar ist, dass Blickkontakt nicht möglich ist, führt hier der Blickkontakt zu Irritation: Blickkontakt ist nur bis zu einem gewissen Grad möglich, da die Einschätzung, was tatsächlich beim Gegenüber ankommt, durch dieses Format erschwert wird. Von vereinzelt Studierenden wurde zudem erwähnt, dass das Hineinversetzen in die Rolle durch das Format erschwert wurde. Die Simulation wurde nicht als Gespräch wahrgenommen, sondern als eine „Rolle spielen“.

## 6 Diskussion

In diesem Beitrag wurden Vor- und Nachteile avatarbasierter und videokamerabasierter Interaktionen als Lehr-Lern-Format bei der Durchführung von Simulationen aus Sicht der Studierenden vorgestellt. Mit Bezug auf den Begriff der Performanz (BLÖMEKE et al., 2015) kann interpretiert werden, dass die Situationsanpassung und somit die Umsetzung der kognitiven Kompetenzen in tatsächlich gezeigte Handlungsfähigkeit (Performanz) in der avatarbasierten Umgebung meist gut gelingt. Das Wahrnehmen der Situation und das Entscheiden für eine kommunikative Handlung gelingt den Studierenden gut, da sie sich auf das gesprochene Wort konzentrieren können (s. a. KUNZE et al., 2016) und sich gleichzeitig als authentisch in der Situation erleben (Immersion). Dies kann mit Bezug zu den Befunden aus der videokamerabasierten Interaktion interpretiert werden. Während es bei der avatarbasierten Interaktion von Beginn an augenscheinlich ist, dass die nonverbale Kommunikation entfällt und die Studierenden sich direkt auf das Gesagte konzentrieren können, führt dies bei der videokamerabasierten Interaktion zu starker Irritation. Bei der videokamerabasierten Interaktion ist nur ein Teilausschnitt der Person ersichtlich. Somit kann hier nicht eingeschätzt werden, wie Gestik und Mimik vom Gegenüber wahrgenommen werden. Anzumerken ist, dass zu dem Zeitpunkt der Durchführung der Simulationen im Sommersemester 2021 bereits seit mehr als einem Jahr aufgrund von COVID-19 Lehre onlinebasiert stattgefunden hat und das i.d.R. über videokamerabasierte Interaktion. Dennoch hat dieses Format zu Verunsicherung und Irritationen bei den Studierenden geführt. Der Einsatz von Gestik und Mimik ist den Studierenden nicht bewusst und auch die Kameraeinstellung sowie vorhandene Ausstattung verbesserungswürdig. Fragen zur Ausstattung von Hochschulen betreffen also mitnichten nur die physischen Räume. Als Limitation ist anzumerken, dass den Studierenden das vollständige Equipment (Headset und Laptop) gestellt werden müsste, um die Bedingungen konstant zu halten.

Ein Vorteil, welcher bei beiden Formaten genannt wird, ist die Sicherheit, in vertrauter Umgebung zu agieren. Zukünftige Forschung müsste nachweisen, ob die Studierenden später in der Lage sind, ihre Handlungsfähigkeit in Face-to-face-Situationen umsetzen können, und dies nicht auf die Simulation beschränkt ist. Um diesen Vorteil in der videokamerabasierten Interaktion gewinnbringend nutzbar zu machen, scheint es Verbesserungsbedarfe zu geben: z. B. anhand von Workshops, um die Studierenden zum Einsatz von Gestik und Mimik oder auch bezüglich der

Kameraeinstellung zu schulen. Onlinelernen, hybrides Lernen oder Blended-Learning werden zunehmend wichtige Bestandteile von Studium und lebenslangem Lernen. Zudem ist deutlich geworden, dass sich die Studierenden durch die räumlichen Gegebenheiten in der avatarbasierten Lernumgebung schneller auf das Gespräch mit ihrem Gegenüber einlassen konnten. Die Räume in TriCAT Spaces konnten eine authentische Lernumgebung darstellen und das Eintauchen (Immersion) in die Rolle mittels Avatar fiel den Studierenden leicht. Die videokamerabasierte Interaktion wird dagegen teilweise als unauthentisches Schauspiel wahrgenommen.

Als Implikation für den Einsatz von performanzbasierten Simulationen in der Hochschullehre lässt sich festhalten, dass sich avatarbasierte Interaktionen sehr gut für erste Praxiserfahrungen eignen. Beide Umgebungen sorgen zunächst für mehr Sicherheit, da sie im geschützten Rahmen von zu Hause erfolgen können. In der avatarbasierten Lernumgebung kommt hinzu, dass Gestik und Mimik hier klar wegfallen, während es in der videokamerabasierten Lernumgebung große Unsicherheiten gibt, wie Gestik und Mimik zu nutzen sind und wie diese durch das Gegenüber wahrgenommen und interpretiert werden.

## 7 Literaturverzeichnis

**Blömeke, S., Gustafsson, J. E., & Shavelson, R. J.** (2015). Beyond dichotomies. *Zeitschrift für Psychologie*, 223(1), 3–13.

**Braun, E., Athanassiou, G., Pollerhof, K., & Schwabe, U.** (2018). Wie lassen sich kommunikative Kompetenzen messen? – Konzeption einer kompetenzorientierten Prüfung kommunikativer Fähigkeiten von Studierenden. *Beiträge zur Hochschulforschung*, 40(3), 34–55.

**Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F.** (2020). Simulation-based learning in higher education: A meta-analysis. *Review of Educational Research*, 90(4), 499–541.

**Dede, C.** (2009). Immersive Interfaces for Engagement and Learning. *Science*, 323, 66–69.

**Fraefel, U., & Scheidig, F.** (2018). Mit Pragmatik zu professioneller Praxis? Der Core-Practices-Ansatz in der Lehrpersonenbildung. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 36(3), 344–364.

- Gartmeier, M., Bauer, J., Noll, A., & Prenzel, M.** (2012). Welchen Problemen begegnen Lehrkräfte beim Führen von Elterngesprächen? Und welche Schlussfolgerungen ergeben sich daraus für die Vermittlung von Gesprächsführungskompetenz? *Die Deutsche Schule*, 104(4), 374–382.
- Habermas, J.** (1981). *Theorie des kommunikativen Handelns* (Vol. 2). Frankfurt am Main: Suhrkamp.
- Hattie, J.** (2015). *Lernen sichtbar machen*. Überarbeitete deutschsprachige Ausgabe von Visible Learning, besorgt von Wolfgang Beywl und Klaus Zierer (3. Aufl.). Hohengehren: Schneider.
- Kerres, M.** (2018). *Bildung in der digitalen Welt – Wir haben die Wahl*. *denk-doch-mal.de*, Online-Magazin für Arbeit-Bildung-Gesellschaft, Ausgabe 02–18 (Berufliches) Lernen in digitalen Zeiten. [https://learninglab.uni-due.de/sites/default/files/Kerres\\_denk-doch-mal.pdf](https://learninglab.uni-due.de/sites/default/files/Kerres_denk-doch-mal.pdf)
- Kickmeier-Rust, M. D., & Reimann, P.** (2020). Offenes Lernen in und mit virtuellen Welten. *Journal für LehrerInnenbildung*, 20(1), 16–26.
- KMK** (Kultusministerkonferenz) (2017). *Qualifikationsrahmen für Deutsche Hochschulabschlüsse*. [https://www.hrk.de/fileadmin/redaktion/hrk/02-Dokumente/02-03-Studium/02-03-02-Qualifikationsrahmen/2017\\_Qualifikationsrahmen\\_HQR.pdf](https://www.hrk.de/fileadmin/redaktion/hrk/02-Dokumente/02-03-Studium/02-03-02-Qualifikationsrahmen/2017_Qualifikationsrahmen_HQR.pdf)
- KMK** (Kultusministerkonferenz) (2019). *Empfehlungen zur Digitalisierung in der Hochschullehre*. Berlin. [https://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2019/2019\\_03\\_14-Digitalisierung-Hochschullehre.pdf](https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2019/2019_03_14-Digitalisierung-Hochschullehre.pdf)
- Kunze, J., Mohr, S., & Ittel, A.** (2016). Online-Rollenspiele in der Lehrkräfteausbildung. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 34(2), 188–195.
- Makrinus, L.** (2013). *Der Wunsch nach mehr Praxis. Zur Bedeutung von Praxisphasen im Lehramtsstudium*. Wiesbaden: Springer VS.
- Mayring, P.** (2015). *Qualitative Inhaltsanalyse*. Grundlagen und Techniken (12., überarbeitete Auflage). Weinheim: Beltz.
- Scheidig, F.** (2020). Lehren lernen mit digitalen Medien. Technologiegestützte Praxisbezüge in der Lehrpersonenbildung. *Zeitschrift MedienPädagogik 17 (Jahrbuch Medienpädagogik)*, 675–708.
- Schulz, M., Mack, B., & Renn, O.** (Hrsg.) (2012). *Fokusgruppen in der empirischen Sozialwissenschaft: Von der Konzeption bis zur Auswertung*. Wiesbaden: Springer-Verlag.

## Autor\*innen



Julia FECKE || Justus-Liebig-Universität Gießen, Zentrum für  
Lehrerbildung, Projekt GOBeL || Rathenastr. 8, D-35394 Gießen

[www.uni-giessen.de/gobel](http://www.uni-giessen.de/gobel)

[julia.fecke@zfl.uni-giessen.de](mailto:julia.fecke@zfl.uni-giessen.de)



Dr. Lars MÜLLER || Justus-Liebig-Universität Gießen, Zentrum  
für Lehrerbildung, Projekt GOBeL || Rathenastr. 8, D-35394  
Gießen

[www.uni-giessen.de/gobel](http://www.uni-giessen.de/gobel)

[lars.mueller@zfl.uni-giessen.de](mailto:lars.mueller@zfl.uni-giessen.de)

# Open Educational Resources (OER) in der wissenschaftlichen Weiterbildung

## Zusammenfassung

Während sich die meisten Arbeiten über (freie) digitale Bildungsressourcen im Bereich der allgemeinen Hochschulbildung bewegen, nähert sich dieser Beitrag dem Thema aus der Perspektive der wissenschaftlichen Weiterbildung. Zunächst werden freie Bildungsmaterialien (OER) entlang der verschiedenen Creative Commons Lizenzen gegenüber digitalen Bildungsressourcen allgemein begrifflich abgegrenzt. Eine Herausforderung für Weiterbildungsanbieter bzw. für das Weiterbildungspersonal besteht zudem in der Auswahl qualitativ hochwertiger Materialien. Hierzu wird auf die verschiedenen Qualitätsdimensionen für digitale Lernmaterialien und auch auf die Schwierigkeiten bei der praktischen Implementation der Qualitätssicherung eingegangen. Weiterbildungseinrichtungen bewegen sich mit ihren kostenpflichtigen Angeboten in einem Markt, in dem der Besitz bzw. das Copyright an Lernmaterialien, deren Entwicklung oft sehr kostenintensiv ist, als ein Wettbewerbsvorteil gesehen wird. Es besteht hier die Tendenz, Lernmaterialien eher nicht als OER zu veröffentlichen. Vor dem Hintergrund dieses Spannungsfeldes wird im Rahmen einer Einzelfallstudie ein Projekt vorgestellt, in dem OER in der wissenschaftlichen Weiterbildung für das in Aufbau befindliche OER-Portal Niedersachsen erstellt werden.

## Schlüsselwörter

Open Educational Resources, OER, digitale Lernmaterialien, Qualitätssicherung, wissenschaftliche Weiterbildung

---

<sup>1</sup> E-Mail: [olaf.zawacki.richter@uni-oldenburg.de](mailto:olaf.zawacki.richter@uni-oldenburg.de)



## Open Educational Resources (OER) in university continuing education

### Abstract

While most work on open educational resources (OER) is in the field of general higher education, this paper approaches the topic from the perspective of university continuing education. First, free educational materials (OER) differ conceptually from digital educational resources in general by their various Creative Commons licenses. One challenge for continuing education providers and training staff is selecting high-quality materials. Therefore, both the quality dimensions for digital learning materials and the difficulties in the practical implementation of quality assurance are discussed. Continuing education institutions operate in a market in which the ownership or copyright of learning materials, which are often very cost-intensive to develop, is seen as a competitive advantage. There is a tendency not to publish learning materials as OER. In the context of this tension, this paper presents a project within the framework of an individual case study, in which OER in continuing education are being created for the OER portal of Lower Saxony, which is currently under development.

### Keywords

Open Educational Resources, OER, digital learning materials, quality assurance, university continuing education

## 1 Einführung

Im Zuge der allgemeinen digitalen Transformation und der Digitalisierung des Lernens und Lehrens hat sich die Open Education Bewegung in den letzten Jahren dynamisch entwickelt (KERRES, 2019). Die Idee von „Open Education“ reicht jedoch weiter zurück und ist insbesondere mit dem Erscheinen von offenen Bildungsangeboten („Open Learning“) in den 1960er-Jahren verknüpft, um sogenannte nicht-traditionelle Zielgruppen zu erreichen (ZAWACKI-RICHTER & STÖTER, 2019). Dies war auch die *raison d'être* der Open Universities, die schon immer Medien

eingesetzt haben, da das Lernen und Lehren im Fernunterricht durch sie überhaupt erst ermöglicht wird (TAIT, 2008; XIAO, 2018).

Massive Open Online Courses (MOOCs) und Open Educational Resources (OER), im deutschsprachigen Raum auch als Offene Bildungsmaterialien bezeichnet, sind aktuelle und im Rampenlicht stehende Erscheinungsformen dieser Entwicklung, wobei OER in Deutschland mehr im Zentrum des gegenwärtigen Diskurses stehen.

Die wissenschaftliche Auseinandersetzung mit OER bezieht sich hauptsächlich auf den Kontext der Hochschulbildung (z. B. OTTO, 2019, 2020; MAYRBERGER, ZAWACKI-RICHTER & MÜSKENS, 2018). Auch in der Strategie der Kultusministerkonferenz „Bildung in einer digitalen Welt“ (KMK, 2017) werden die Potenziale von OER insbesondere in Hinblick auf die Hochschulen hervorgehoben.

Im Bereich der (wissenschaftlichen) Weiterbildung spielen OER bisher kaum eine Rolle, obwohl in einer Befragung der Deutschen Gesellschaft für wissenschaftliche Weiterbildung und Fernstudium (DGWF) 78 % der befragten Mitglieder der Aussage „OER bietet neue Chancen für die wissenschaftliche Weiterbildung“ zustimmten (STEINERT & DEIMANN, 2016). Im Whitepaper zu OER in der Erwachsenen- und Weiterbildung stellten auch BLEES et al. (2015) fest, dass die Debatte für die Weiterbildung Fahrt aufnimmt, da Weiterbildung und OER gut zusammenpassen (S. 9). Sie verdeutlichen dies an den vier Leitgedanken für die „Fortentwicklung [...] des Bereichs der Weiterbildung im Rahmen des lebenslangen Lernens“ der KMK (2013, S. 171): „die Stärkung der Eigenverantwortung sowie Selbststeuerung der Lernenden, der Abbau der Chancenungleichheiten, die Kooperation der Bildungsanbieter und Nutzer, die Stärkung der Bezüge zwischen allen Bildungsbereichen“ (ebd).

Vor diesem Hintergrund sollen OER in diesem Beitrag als ein mögliches Konzept für die wissenschaftliche Weiterbildung in den Blick genommen werden, die sich hier im Spannungsfeld zwischen freier Nutzung und dem Angebot auf dem wettbewerbsorientierten Weiterbildungsmarkt bewegen. Ziel dieses Beitrages ist es, Chancen und Herausforderungen in der Entwicklung, Veröffentlichung und Anwendung von OER im Hinblick auf die Spezifika der wissenschaftlichen Weiterbildung zu analysieren, wobei ein besonderer Fokus auf die Qualitätssicherung von freien Lernmaterialien gelegt wird. Insbesondere soll herausgearbeitet werden, welchen Mehrwert OER für Weiterbildungseinrichtungen jenseits direkter monetärer Erträ-

ge bieten können, indem Inhalte durch die Veröffentlichung als OER sichtbar werden.

## 2 Hemmnisse bei der Verbreitung von OER

Im internationalen Vergleich kann konstatiert werden, dass die Nutzung von OER ganz allgemein und nicht nur in der Praxis der Weiterbildung hinter den Erwartungen zurück- und Deutschland ein Nachzügler in Bezug auf diese Entwicklung geblieben ist (ORR, NEUMANN & MUUSS-MERHOLZ, 2017).

Von 2016 bis 2018 wurden im Rahmen der Förderinitiative OER-Info 25 Bildungsbereich-übergreifende Projekte (Schule, Hochschule, Weiterbildung, Berufsbildung) durch das Bundesministerium für Bildung und Forschung (BMBF) durchgeführt. Hiermit wurde eine erste breite Förderung von OER angestoßen, in der es insbesondere um die Sensibilisierung und Bekanntmachung der Potenziale von OER ging (vgl. UNIVERSITÄT HAMBURG, 2018).

Nach einer Meta-Analyse der 25 Projekte muss OTTO (2019) feststellen, dass die Projekte zwar zu einer Bewusstseinsbildung im Hinblick auf OER beigetragen haben, eine Implementierung und Anwendung in der Praxis aber eher punktuell als systematisch erfolgt.

Bereits DEIMANN und BASTIAENS (2010) sind der Frage nach Potenzialen und Hemmnissen von OER in einer Delphi-Studie nachgegangen. Sie identifizierten in einer Rangfolge vier zentrale Hürden: 1. Kulturelle Hürden und Misstrauen gegenüber fremderstellten Materialien, 2. das Fehlen einer kritischen Masse zur Verfügung stehender Materialien gerade im deutschsprachigen Raum, 3. Unsicherheiten rund um rechtliche Aspekte der Lizenzierung von OER und 4. Technische Hürden bei der Erstellung und Veröffentlichung von OER. Das folgende Zitat aus dieser Arbeit soll hier nicht unterschlagen werden: „Die OER-Bewegung hat durchaus Potential, sich weiter zu verstetigen, dies wird jedoch eher in Kleinen passieren. Ein Massenphänomen ist hier nicht zu erwarten“ (ebd., S. 17).

Die mangelnde Akzeptanz von OER wird oft mit der intransparenten Qualität der Lernmaterialien in Verbindung gebracht (vgl. JUNG, SASAKI & LATCHEM, 2016). Bereits der Bericht der Open Educational Quality Initiative (OPAL, 2011)

identifizierte den „lack of quality or fitness of OER“ (S. 8) als eine wichtige Barriere zur Nutzung von OER. Seitdem hat sich auch aus internationaler Sicht kein Verfahren durchgesetzt, das weithin anerkannt ist und breite Anwendung findet (vgl. ZAWACKI-RICHTER & MAYRBERGER, 2017).

Vor diesem Hintergrund sollen OER unter besonderer Berücksichtigung der Qualitätssicherung zunächst in das Spektrum der Open-Education-Bewegung eingeordnet und in ihren Merkmalen und Eigenschaften beschrieben werden.

### 3 OER im Kontext von Openness

„Openness“ hat viele Facetten, und OER sind dabei nur ein Phänomen der wesentlich breiteren Open-Education-Bewegung, in der es um Zugangsmöglichkeiten zu Bildungsangeboten und Durchlässigkeit in Bildungssystemen geht. Nach einem Positionspapier der COER-Gruppe<sup>2</sup> (ZAWACKI-RICHTER et al., 2020) gehören zu den Elementen von Open Education Angebote des Open & Distance Learning – heute in der Form von digitalen Lernangeboten –, die Entwicklung von OER und sogenannten Massive Open Online Courses (MOOCs) und auch Prozesse des Prior Learning Assessment and Recognition (PLAR), in denen es um die Anerkennung und Anrechnung von Kompetenzen, Fertigkeiten und Qualifikationen geht, um die vertikale und horizontale Durchlässigkeit zu erhöhen (vgl. CONRAD, 2008; MÜSKENS, 2006). Die Open Education Bewegung wird mit einer Open Educational Practice (OEP) verknüpft, die definiert werden kann als „a broad range of practices that are informed by open education initiatives and movements and that embody the values and visions of openness“ (KOSEOGLU & BOZKURT, 2018, S. 455). Eine OEP bedingt also auch eine bestimmte Haltung der Lehrenden, die vor dem Hintergrund der digitalen Vernetzung von VELETSIANOS und KIMMONS (2012) auch als Open Scholarship beschrieben wird: „a set of phenomena and practices surrounding scholars’ uses of digital and networked technologies underpinned by certain grounding assumptions regarding openness and democratization of knowledge creation and dissemination“ (S. 168).

Im Hinblick auf den Zusammenhang von OEP und OER beschreiben BELLINGER

---

2 Center of Open Education Research (<http://www.coer.education>).

und MAYRBERGER (2019) in einem Systematic Review die Produktion und Nutzung von OER im Sinne eines engen Verständnisses von OEP. Weit gefasst bezieht sich OEP auf „die uneingeschränkte Übernahme der Grundsätze einer Open Pedagogy, einschliesslich der institutionellen Rahmenbedingungen und Governance (institutionelle Meso- und Makroebene)“ (ebd., S. 39).

In Deutschland werden OER besonders bottom-up in einer recht engen Community von OER-Expert\*innen propagiert, die sich stark mit einer normativen Forderung nach Bildungsgerechtigkeit im Sinne einer Bildung für alle identifiziert (vgl. ORR, NEUMANN & MUUSS-MERHOLZ, 2017; KERRES, 2019). OTTO (2020) plädiert jedoch dafür, die (medien-)pädagogischen Mehrwerte von digitalen Lernmaterialien stärker hervorzuheben und „nicht mit einer normativen Zielsetzung zu missionieren, sondern OER als eine (Teil-)Antwort auf eine breitere Herausforderung des Lehrens und Lernens in der digitalen Welt zu begreifen, zu vermitteln und bildungsbereichsübergreifend zu integrieren“ (S. 39).

### 3.1 Merkmale von OER

In den Empfehlungen der UNESCO (2019) zu OER werden diese wie folgt definiert:

Open Educational Resources (OER) are learning, teaching and research materials in any format and medium that reside in the public domain or are under copyright that have been released under an open license, that permit no-cost access, re-use, re-purpose, adaptation and redistribution by others. (S. 5)

Im Kern geht es bei OER also darum, Lehr- und Lernmaterialien uneingeschränkt nutzbar zu machen. Digitale Materialien bieten sich hierfür in besonderer Weise an, da sie beliebig oft, verlustfrei und räumlich praktisch unbegrenzt kopiert, geteilt und verändert werden können. WILEY (2013) beschreibt in seinem Blog diese Offenheit der Nutzung mit fünf Rs: Retain, Reuse, Revise, Remix und Redistribute<sup>3</sup>. Die OER-Infostelle übersetzt dies mit den fünf V-Freiheiten: OER können verwahrt & vervielfältigt, verwendet, verarbeitet, vermischt und schließlich auch weiter verteilt werden<sup>4</sup>. Auf dieser Grundlage wird OER das Potenzial zugeschrieben, besonders

---

3 <https://opencontent.org/blog/archives/3221> (Zugriff am 25.01.2021).

4 <https://open-educational-resources.de/5rs-auf-deutsch/> (Zugriff am 25.01.2021).

die Entwicklung von kollaborativen und partizipativen Lernarrangements zu ermöglichen (OTTO, 2020).

Dabei ist das Spektrum der möglichen OER vielfältig. JUNG, SASAKI und LATCHEM (2016) sprechen in diesem Zusammenhang von der Granularität von Lernmaterialien: „OER range from entire courses and massive open online courses to smallscale learning materials, games, simulations, quizzes and other digital resources“ (S. 10). Freie Bildungsressourcen beschränken sich also nicht alleine auf Inhalte, sondern können auch Open-Source-Software und Tools zum Lernen und Lehren umfassen (z. B. ein Open-Source-Videokonferenzsystem).

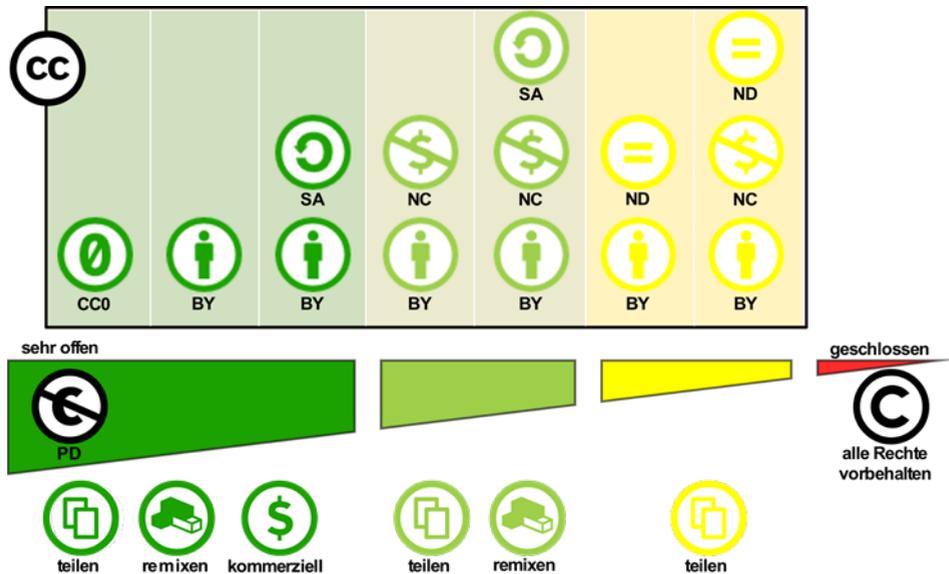


Abb. 1: Spektrum der CC-Lizenzen (von Napa unter der Lizenz CC BY 4.0 via Wikicommons)

Bei der weiteren Nutzung und Verarbeitung der Lernmaterialien ist das Urheberrecht zu beachten. Weit verbreitet bei der Erstellung und Veröffentlichung von OER ist die Vergabe von Creative-Commons (CC)-Lizenzen (siehe Abb. 1). Nach der Definition der UNESCO handelt es sich nur um „echte“ OER, wenn sie dem dunkelgrünen Spektrum der CC-Skala zuzuordnen sind: CC-0, CC-BY und CC-BY-SA, d. h. Materialien, die ohne Auflagen völlig frei sind (0), nur unter Nennung der Autor\*innen (BY = by) und/oder zusätzlich mindestens unter der gleichen Lizenz weiterzugeben sind (SA = share alike). Materialien, die nur nicht-kommerziell (NC = non-commercial) genutzt und/oder nicht verändert werden dürfen (ND = no derivatives), fallen demnach nicht unter das Label der freien Lernmaterialien.

### 3.2 Qualität von OER

Zur Qualitätsdebatte im Bildungskontext sei auf BÜLOW-SCHRAMM (2006) verwiesen. Es soll hier jedoch hervorgehoben werden, dass Bildung kein Produkt oder eine Dienstleistung ist, die an Kunden ausgeliefert wird, sondern der Mitwirkung der Lernenden bedarf. Die Lernergebnisse sind in hohem Maße auch von ihnen selbst abhängig. Es handelt sich hier also um einen „interaktiven Prozess mit offenem Ausgang“ (EHLERS, 2011, S. 61)

Die große Bedeutung der Qualitätssicherung von OER ist weitgehend unumstritten. Wie bereits ausgeführt, wird eine undurchsichtige Qualität von „fremden“ Lernmaterialien oft als ein Argument gegen die Nutzung von OER herangezogen oder als ein Hemmnis für deren Verbreitung gesehen.

Auch in der wissenschaftlichen Weiterbildung ist es bei dem großen Angebot von Lernmaterialien ganz allgemein von großer Bedeutung, deren Qualität für die Entwicklung hochwertiger Angebote beurteilen zu können. Ebenso benötigen die in den Bundesländern im Aufbau befindlichen Repositorien für OER ein Qualitätssicherungsverfahren, um die Qualität der auf den OER-Portalen veröffentlichten Materialien zu gewährleisten. Für den deutschsprachigen Raum gibt es jedoch hierfür kein etabliertes und standardisiertes Modell, das leicht anwendbar wäre. Dies trifft auch auf die internationale Literatur zu. YUAN und RECKER (2015) betonen, dass es kaum Verfahren gibt, die den wissenschaftlichen Gütekriterien der Testentwicklung genügen.

Vor diesem Hintergrund haben MAYRBERGER, ZAWACKI-RICHTER & MÜSKENS (2018) nach einer umfassenden internationalen Bestandsaufnahme von Instrumenten und Qualitätskriterien (vgl. ZAWACKI-RICHTER & MAYRBERGER, 2017) ein Qualitätsmodell für OER im Rahmen des Projektes der Hamburg Open Online University<sup>5</sup> entwickelt, das im Rahmen von mehreren Workshops mit Stakeholdern validiert wurde. Es wird hier zwischen Kriterien der pädagogisch-didaktischen Dimension (Inhalt und didaktische Konzeption) und der technischen Dimension (Zugänglichkeit und Usability) unterschieden.

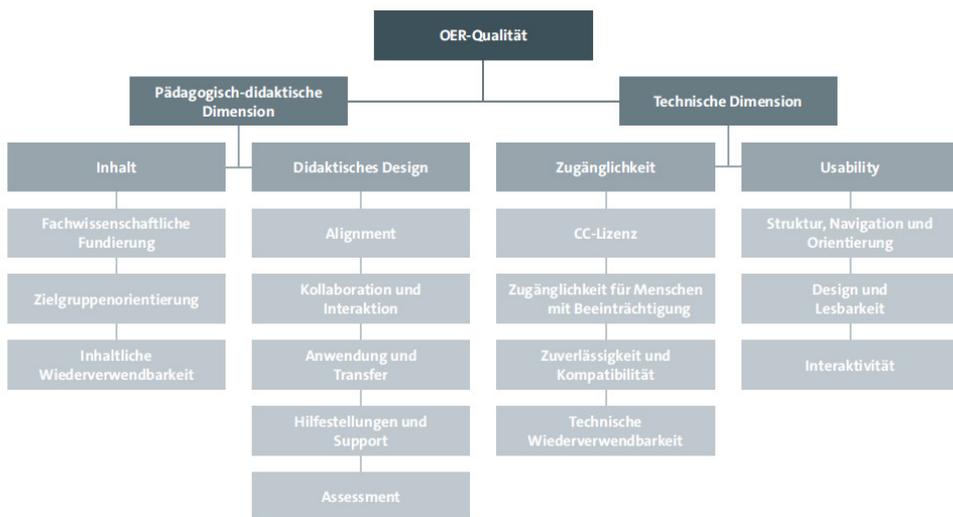


Abb. 2: Qualitätsmodell für OER (MAYRBERGER, ZAWACKI-RICHTER & MÜSKENS, 2018, S. 29, CC-BY)

Auf dieser Grundlage werden gegenwärtig im vom BMBF geförderten Projekt „Digitale Bildungsarchitekturen: Offene Lernressourcen in verteilten Lerninfrastruk-

<sup>5</sup> <http://www.hoou.de> (Zugriff am 11.02.2021).

turen“ (EduArc<sup>6</sup>) Skalen für ein Instrument (IQOer) entwickelt und erprobt. Ziel der Entwicklung ist ein objektives, reliables und valides Bewertungsinstrument, das nach Möglichkeit auf OER unterschiedlicher Granularität und Inhalte anwendbar sein soll. Hierzu wurden 15 Skalen entworfen, anhand derer in einer zweistufigen Validierungsstudie jeweils 50 OER von jeweils zwei Ratern mittels der Skalen bewertet wurden. Anhand dieser Bewertungen sollen u. a. die Interrater-Reliabilität (d. h. die Übereinstimmung der Rater hinsichtlich der Bewertung einer OER) sowie die Konstruktvalidität der Skalen bestimmt werden.

Je nach Ziel und Zeitpunkt der Qualitätserfassung kommen unterschiedliche Gruppen von Personen als Rater, d. h. für die Durchführung der Bewertung, in Frage:

- Die Autor\*innen oder Ersteller\*innen selbst können Standards, Kriterien oder Skalen während der Entwicklung der OER verwenden, um noch ausstehende Arbeitsschritte zu identifizieren.
- Inhaltsexpert\*innen können im Rahmen eines Peer-Reviewings die inhaltliche Qualität der Ressourcen summativ bewerten. Die Bewertung kann entweder vor der Veröffentlichung auf einer Plattform oder auf Grundlage einer Konzeptbeschreibung vor Beginn der Entwicklung stattfinden.
- Fachpersonal aus den Bereichen Didaktik oder Technik kann vor Veröffentlichung einer OER die Einhaltung technischer bzw. didaktischer Mindeststandards kontrollieren. Aber auch die Entwicklung kann durch eine formative Evaluation hinsichtlich dieser Standards begleitet und unterstützt werden.
- Eine Bewertung von Nutzer\*innen erfolgt üblicherweise erst nach Veröffentlichung der Ressource. Die Verwendung eher subjektiver Bewertungsskalen (z. B. als wie motivierend oder interessant wird die Ressource wahrgenommen?) kann anderen Nutzer\*innen dabei Verwendungshinweise liefern, die über objektive Bewertungsmaßstäbe hinausgehen.

Autor\*innen von OER und Anbieter von OER-Plattformen stehen jedoch vor der Herausforderung, wie eine solche Qualitätssicherung innerhalb des Entwicklungsprozesses von OER ablaufen kann. Selbst wenn reliable und valide Instrumente wie die auf dem Qualitätssicherungsmodell von MAYRBERGER et al. (2018) aufbau-

---

6 Siehe: <https://uol.de/coer/research-projects/projects/eduarc> (Zugriff am 26.01.2021).

enden Skalen zur Verfügung stehen, bleiben aus der Sicht der Akteure und Akteurinnen wesentliche Fragen unbeantwortet:

- An welchem Punkt innerhalb des Prozesses der Erstellung bzw. Nutzung eines OER soll die Qualitätserfassung stattfinden (Zeitpunkt)?
- Mit welcher Absicht erfolgt die Qualitätserfassung? (Ziel)
- Wer bewertet die Ressourcen anhand welcher Kriterien oder Skalen? (Rater)
- Auf welche Weise arbeiten die qualitätssichernde Einrichtung und Autor\*innen bzw. Entwickler\*innen der OER zusammen? (Grad der Interaktion)

Diese Fragen stellen sich bei der praktischen Entwicklung und Veröffentlichung von OER, wie in folgender Einzelfallstudie weiter ausgeführt wird.

## **4 Chancen und Herausforderungen von OER in der wissenschaftlichen Weiterbildung**

Gerade im Bereich der Weiterbildung ist die Nutzung und Verbreitung von OER besonders gering und mit hohen Widerständen verbunden (Otto, 2019, 2020). Das Thema bewegt sich hier im Spannungsfeld zwischen freier Nutzung und Kooperation und dem Angebot auf dem kompetitiven Weiterbildungsmarkt.

Die Chancen und Herausforderungen von OER in der wissenschaftlichen Weiterbildung sollen daher hier im Sinne einer explorativen Einzelfallstudie beleuchtet werden. Hierfür wird das Analyseraster von MCGREAL (2019) herangezogen, der die Implementierung von OER an 13 Hochschulen in Nord- und Südamerika, Afrika, Südost-Asien und dem Mittleren Osten untersucht hat. Hierbei werden die folgenden Fragen gestellt:

- Opportunity: Was führte zu der Entscheidung für OER?
- Innovation: Was ist neu? Was wurde entwickelt?
- Benefits: Welches sind die Ergebnisse der Intervention?

- Challenges: Welche Probleme bestanden bei der Entwicklung und Implementierung?
- Potenzial: Was verspricht man sich für die Zukunft des Projektes?

## 4.1 Institutioneller Kontext

Bei dem Beispiel, das hier in den Blick genommen werden soll, handelt es sich um das Center für lebenslanges Lernen (C3L) an der Carl von Ossietzky Universität Oldenburg. Mit über 50 Mitarbeitenden ist das C3L eine der größten Einrichtungen für wissenschaftliche Weiterbildung an deutschen Hochschulen. Das C3L nimmt als wissenschaftliches Zentrum eine führende Rolle bei der Etablierung innovativer Lehr-Lern-Formate ein. Bereits vor über 20 Jahren wurden an der Universität Oldenburg Studiengänge entwickelt (vgl. RUBIN, BERNATH & PARKER, 2004), die rein online als Fernstudium oder im Blended-Learning-Format angeboten wurden. Aktuell bietet das C3L neben verschiedenen Formaten der öffentlichen Wissenschaft sowohl zwei Bachelor- und sechs Masterstudiengänge als auch Kontaktstudienprogramme, Zertifikatsprogramme und Ausbildungen in Psychotherapie an.

## 4.2 Opportunity: Ausgangslage

Ausgangspunkt für die Entwicklung von OER am C3L war ein Strategie- und Reorganisationsprozess, der im April 2019 mit der Einführung einer neuen Zentrumsordnung und eines neu besetzten Direktoriums eingesetzt hatte. Mit Bezug auf den Hochschulentwicklungsplan und das Leitbild der Universität Oldenburg als „Offene Hochschule“ wurde u. a. festgelegt, dass die Sichtbarkeit der qualitativ hochwertigen Studienmaterialien des C3L über die Veröffentlichung als OER erhöht werden soll. Die Abteilungen „Bildungsmedien & Lerndesign“ und „Offene Hochschule“ wurden damit beauftragt, mit Eigen- und Drittmitteln vorhandene Lernmaterialien zu erschließen, als OER aufzubereiten bzw. zu erstellen. Am C3L liegen Lernmaterialien zu 105 Modulen aus berufs begleitenden Studiengängen im Blended-Learning-Format vor.<sup>7</sup> Günstig für diese Entwicklung war auch, dass im C3L bereits

---

<sup>7</sup> <https://uol.de/c3l/studiengang> (Zugriff am 11.02.2021).

Ressourcen und Kompetenzen im Bereich der Mediendidaktik und Medienproduktion vorhanden waren.

Im Rahmen der Ausschreibung „OER für den Hochschulbereich – Förderung der Erstellung, Pflege und Verbreitung von Open Educational Resources an den niedersächsischen Hochschulen“ des Niedersächsischen Ministeriums für Wissenschaft und Kultur (MWK) wird nun seit Dezember 2020 das Projekt „Open Educational Resources zum Bildungs- und Wissenschaftsmanagement“ mit rund 180 TEUR gefördert. Das zunächst auf 18 Monate festgelegte Vorhaben ist ein Tandemprojekt zwischen dem C3L der Universität Oldenburg (UOL) und der Hochschule Osnabrück (HOS, Fakultät für Wirtschafts- und Sozialwissenschaften). An beiden Standorten werden Studiengänge im Bereich des Bildungs- und Wissenschaftsmanagements angeboten, die sich an berufstätige Zielgruppen wenden.

Der Schwerpunkt liegt auf der Erschließung und Aufbereitung bereits vorhandener Materialien als OER für vier ausgewählte Module, die in unterschiedlicher Granularität in digitalen Formaten entwickelt, sich in verschiedenen Kontexten und Angebotsformaten anpassen und kombinieren lassen.

Das Projekt stellt für das C3L eine gute Möglichkeit dar, die Prozesse der Erschließung, Anpassung, Neuproduktion, Lizenzierung und Veröffentlichung zu erproben und auf andere Inhaltsbereiche zu übertragen.

### **4.3 Innovation: Digitalisierung, Repositorien und Qualitätssicherung**

Während englischsprachige Plattformen oder Repositorien für digitale Lernmaterialien wie MERLOT<sup>8</sup> bereits viele tausend Ressourcen umfassen, ist der Bestand an deutschsprachigen OER aktuell noch sehr beschränkt. OTTO (2020) kam in seiner Untersuchung zum Ergebnis, dass nicht die Einstellungen der beteiligten Akteurinnen und Akteure das zentrale Hemmnis für die Nutzung von OER in Deutschland sind, sondern deren mangelnde Verbreitung. Dies trifft insbesondere für die wissenschaftliche Weiterbildung zu. Schon die Erstellung und Veröffentlichung von OER in diesem Bildungssektor hat innovativen Charakter.

---

8 <http://www.merlot.org> (Zugriff am 11.02.2021).

Parallel zu den Projekten der Contententwicklung fördert das MWK auch den Aufbau eines niedersächsischen OER-Portals<sup>9</sup>, auf dem die Lernmaterialien veröffentlicht werden sollen. Aus der Sicht der Betreiber von OER-Plattformen ist es gegenwärtig eher das Ziel, Autor\*innen bzw. Entwickler\*innen bei der Erstellung von OER zu unterstützen, als eine Selektion eingereicherter Ressourcen vorzunehmen. Vielmehr wurde es den geförderten Projekten auferlegt, selbst für eine Qualitätssicherung zu sorgen.

Wie oben ausgeführt, stellt die Anwendung eines Evaluationsinstruments für OER in Verbindung mit einem Qualitätssicherungsprozess eine wesentliche Innovation dar. Entsprechend liegt ein zentraler Fokus des Projektvorhabens auf den unterschiedlichen didaktischen und technischen Qualitätsdimensionen (siehe Abb. 2) für digitale Lernmaterialien sowie auf der Qualitätssicherung im Hinblick auf die praktische Implementation von OER.

Geplant sind die beispielhafte Anwendung und Erprobung eines dreistufigen Review-Verfahrens für die entwickelten Materialien:

- In der ersten Stufe des Reviews erfolgt eine Evaluation der Materialien hinsichtlich der inhaltlichen und pädagogisch-didaktischen Dimension mithilfe des IQOer-Instrumentes durch Fachexpert\*innen (Peer-Review).
- In der zweiten Stufe erfolgt eine Evaluation im Hinblick auf die technischen Dimensionen des Qualitätsmodells (Zugänglichkeit für Menschen mit Beeinträchtigung, Zuverlässigkeit und Kompatibilität, technische Wiederverwertbarkeit, Usability). Dieser Teil des Reviews wird von technisch geschultem Fachpersonal der Universität Oldenburg durchgeführt.
- In der dritten Stufe bewerten potenzielle Nutzer\*innen die Materialien. Hierzu bewerten Hochschullehrende jedes der Materialien mittels der Skalen des IQOers.

Die Ergebnisse des dreistufigen Reviews bilden die Grundlage für eine anschließende Überarbeitung und Optimierung der Materialien.

Eine Voraussetzung für eine Veröffentlichung und eine spätere Rekombination und Überarbeitung der OER ist die Festlegung ihrer inhaltlichen Einheiten oder „Gra-

---

<sup>9</sup> <http://www.oernds.de> (Zugriff am 11.02.2021).

nularität“ (JUNG, SASAKI & LATCHEM, 2016), die auch in einem OER-Repositorium abbildbar sein muss. Die Elemente unterschiedlicher Granularität bilden jeweils miteinander verschachtelte Ebenen. So kann beispielsweise ein Kurs aus mehreren Lerneinheiten bestehen, die sich jeweils wiederum aus unterschiedlichen Lernobjekten zusammensetzen. Die Elemente aller Ebenen sollen eigenständig wiederverwertbar sein, um von Lehrenden je nach Bedarf in ihre eigenen Veranstaltungen eingebunden werden zu können.

Die inhaltlichen Einheiten wurden in folgenden Stufen festgelegt:

- Informationsobjekte (beispielsweise Bilder, Grafiken, Texte, Screencasts, Lehrvideos),
- Lernobjekte (z. B. Informationsobjekte verknüpft mit einer Aufgabe oder Übung),
- Lehr-/Lerneinheiten (Lernobjekte, die durch Lernziele, methodisches Vorgehen und Aufgaben miteinander verbunden sind) oder
- Kurse/Module (mehrere Lehr-/Lerneinheiten, die über einen Lernpfad/Study Guide didaktisch miteinander verbunden sind).

#### **4.4 Benefits: Ergebnisse der Intervention**

Der naheliegende Mehrwert für die beteiligten Bildungsreinrichtungen liegt zunächst darin, dass die Studienmaterialien von vier Modulen zu den Themen Strategisches Management, Qualitätsmanagement, Kosten- und Finanzmanagement sowie Methoden und Modelle digitaler Bildung in zwei Studiengängen inhaltlich überarbeitet und in vielfältige digitale Formate überführt werden (Study Guides, Lehrtexte, Lehrvideos und Screencasts, Fallstudien, Aufgaben, Open Access Literatursammlungen usw.).

Aus strategischer Sicht sind die Projektergebnisse für das C3L von höherer Bedeutung, da hier grundlegende Überlegungen zur Systematisierung von Lernmaterialien in der wissenschaftlichen Weiterbildung angestellt werden und Prozesse zur Digitalisierung von Lernmaterialien und ihrer Qualitätssicherung erprobt und evaluiert werden.

Mit der weiteren Verbreitung von OER bietet sich für das C3L ebenfalls die Möglichkeit, bestehende OER heranzuziehen und diese zur remixen und für das eigene Lehrangebot zu nutzen. Somit ließen sich Kosten für die Neuerstellung von Materialien senken, diese müssten dann nach der Überarbeitung allerdings auch wieder veröffentlicht werden.

#### **4.5 Challenges: Finanzierung, Aufwand und Akzeptanz**

Ein zentraler Punkt bei der Entscheidung, OER in der wissenschaftlichen Weiterbildung einzuführen, ist sicherlich die Finanzierung der Erstellung neuer Lernmaterialien bzw. der Erschließung und Pflege vorhandener Materialien und deren Veröffentlichung als OER. Die Kosten sind i.d.R. durch Teilnehmer\*innenentgelte zu refinanzieren, wenn nicht die Möglichkeit einer Drittmittelförderung besteht. In dem im Jahr 2020 abgeschlossenen großen Förderprogramm „Aufstieg durch Bildung – offene Hochschulen“ des BMBF, in dem zahlreiche Weiterbildungsprogramme bis hin zu berufsbegleitenden Masterstudiengängen entwickelt wurden, mussten zum Beispiel alle entwickelten Materialien veröffentlicht, nicht aber zwingend mit einer CC-Lizenz versehen werden<sup>10</sup>.

Die Kosten zur Erstellung und Pflege von Studienmaterialien sind nicht unerheblich. Nach Angaben der Geschäftsführung<sup>11</sup> des C3L belaufen sich die durchschnittlichen Ausgaben über die letzten zehn Jahre für Werkverträge, Personal und Medienproduktion im laufenden Regelbetrieb auf knapp 100.000 EUR jährlich, hinzu kommen ca. 20.000 EUR jährlich für den Versand von gedruckten Studienmaterialien. Diese Angaben beziehen sich nur auf die sieben berufsbegleitenden, weiterbildenden Bachelor- und Masterstudiengänge, wobei noch keine Kosten aus Anschubfinanzierungen über Drittmittelprojekte enthalten sind.

Die Studienmaterialien stellen für Weiterbildungseinrichtungen also ein wertvolles Gut dar. Andererseits wird kaum jemand nur aufgrund der Lernmaterialien einen Kurs buchen oder sich in einen weiterbildenden Studiengang einschreiben. Die Verfügbarkeit hochwertiger und relevanter Inhalte ist einfach eine Grundlage für die

---

10 Siehe z. B.: [https://pug-pflege-und-gesundheit.de/public\\_pages/56](https://pug-pflege-und-gesundheit.de/public_pages/56) (Zugriff am 25.01.2021).

11 Angaben per E-Mail am 26.01.2021.

Durchführung von Weiterbildungsangeboten. Es sind jedoch andere Faktoren, die für die Entscheidung zu einer Weiterbildung bei einem bestimmten Anbieter ausschlaggebend sind. Das Massachusetts Institute of Technology (MIT) hatte bereits 2002 mit der MIT OpenCourseWare Initiative<sup>12</sup> das Ziel verfolgt, alle Lernmaterialien der Universität frei online verfügbar zu machen, ohne dabei zu befürchten, dass dies auf Kosten der hohen Studiengebühren am MIT gehen würde. Es geht darum, ein Bildungsangebot an einer renommierten Institution mit einem hohen Qualitätsversprechen wahrzunehmen. Lernen ist immer ein sozialer Prozess, in dem die Beziehungen zwischen Lernenden und Lehrenden, sowie den Lernenden untereinander den größten Effekt haben (vgl. HATTIE, 2009).

Dennoch stellen Studienmaterialien eine hohe Investition dar, die auch geschützt werden muss. Eine im Wettbewerb stehende Weiterbildungseinrichtung hat sicherlich kein Interesse daran, dass ein konkurrierendes Unternehmen mit den Inhalten Erlöse erzielt, ohne dass ein Rückfluss etwa in Form von Lizenzgebühren erfolgt. Dies kann über eine entsprechende CC-Lizenzierung mit NC (non-commercial) reguliert werden – auch wenn Verfechter einer strengen Definition diese dann nicht mehr als OER bezeichnen würden.

Die Entwicklung des Qualitätssicherungsinstruments IQOer und seine Erprobung im oben dargestellten Verfahren ist sehr aufwändig und wäre ohne eine Drittmittelförderung nicht möglich gewesen. Hierfür sind dauerhaft Ressourcen zur Verfügung zu stellen, um die an der Qualitätssicherung beteiligten Akteurinnen und Akteure einbinden zu können (siehe Abschnitt 2.2).

Eine weitere Herausforderung bei der Erstellung von OER ist auch die Akzeptanz bei den Lehrenden bzw. den Autor\*innen der Studienmaterialien. Bei Veröffentlichungen in Fachzeitschriften ist man es gewohnt, sich der Kritik der Peer-Reviewer bzw. der Fachcommunity nach Veröffentlichung auszusetzen. Dies gilt aber nicht gleichermaßen für die Veröffentlichung von Lehrmaterialien. In früheren Autorenverträgen wurden dem C3L sämtliche Verwertungsrechte übertragen. Dabei wurde es den Auftragnehmenden gestattet, das erstellte Studienmaterial auch in eigenen Lehrveranstaltungen, bei Vorträgen und in ähnlichen Zusammenhängen zu verwenden. In neuen Autorenverträgen soll folgende Passage eingefügt werden:

---

12 <https://ocw.mit.edu> (Zugriff am 11.02.2021).

Das Werk soll als Open Education Ressource (OER) veröffentlicht werden. Es ist geplant, das Studienmaterial anderen Nutzer/innen mit einer CC-Lizenz in der Form BY-NC-SA online zur Verfügung zu stellen (Nennung der Autoren, keine kommerzielle Nutzung, Veränderung und Weitergabe nur nach gleichen Bedingungen). Es wird die CC-Lizenz in der jeweils aktuellen Fassung verwendet.

Obwohl die Autor\*innen so eine viel weiter gehende Kontrolle über ihre Materialien behalten, ist davon auszugehen, dass einzelne Autor\*innen weiterhin die Vertragsvariante ohne Open Access bevorzugen, um die Inhalte nicht öffentlich zu machen. Hierzu gibt es jedoch bisher keine Datenlage.

#### **4.6 Potentials: Geteilter Aufwand und erhöhte Sichtbarkeit**

Das C3L sieht sich dem Gedanken der Universität Oldenburg als „Offene Hochschule“ (vgl. HANFT & BRINKMANN, 2013) verbunden und möchte als innovativer Weiterbildungsanbieter eine führende Rolle bei der Entwicklung von offenen Bildungsressourcen einnehmen.

Die Potenziale von OER liegen dabei jenseits direkter monetärer Erträge. Die systematische Beschäftigung mit OER trägt zur Weiterentwicklung und zur Professionalisierung des mediendidaktischen und medientechnischen Prozesses der Lernmaterialerstellung und insbesondere seiner Qualitätssicherung bei. Die Universität Oldenburg führt gerade die Systemakkreditierung ein, bei der das interne Qualitätssicherungssystem im Bereich von Studium und Lehre begutachtet wird. Die in dem Projekt entwickelten Instrumente und Verfahren können hierzu einen Beitrag leisten.

Ausgelöst durch die Covid-19-Pandemie hat die Digitalisierung von Studium und Lehre einen enormen Schub erfahren (KERRES, 2020; ZAWACKI-RICHTER, 2020). Dies gilt auch für die wissenschaftliche Weiterbildung, obwohl die Angebote hier schon stärker digitalisiert waren, um zum Beispiel berufstätigen Zielgruppen in der Form des Blended Learning einen räumlich-zeitlich flexiblen Zugang zu ermöglichen (vgl. ZAWACKI-RICHTER & STÖTER, 2019). OER Projekte wie die vom MWK geförderten Initiativen können hier eine Leuchtturmfunktion einnehmen. Lehrende können so an die Entwicklung digitaler Lernmaterialien herangeführt und aktiviert werden.

Durch die Entwicklung von OER und deren Veröffentlichung auf OER-Portalen wird eine Möglichkeit der Kooperation – wie hier zwischen der Universität Oldenburg und der Hochschule Osnabrück – ermöglicht, innerhalb derer der Aufwand gemeinsam getragen wird und die digitalen Materialien dann mit der Fachcommunity der Lehrenden geteilt werden. Auf diese Weise ergeben sich auch Synergieeffekte, wenn Materialien als OER geteilt werden und nicht gleiche Inhalte parallel an unterschiedlichen Standorten produziert werden.

Schließlich kann eine Weiterbildungseinrichtung durch die Veröffentlichung ihrer Lernmaterialien als OER die Qualität der Inhalte sichtbar machen und diese durchaus auch als ein Marketinginstrument einsetzen. So könnte es doch zu monetären Erträgen über Kursbuchungen kommen. Solche Effekte werden auch im Bereich des Open Access Publishing beobachtet. Die Veröffentlichung eines Buches als Open Access führt nicht zu einem signifikanten Einbruch der Verkaufszahlen (SPEICHER, 2016). Im Gegenteil kann die verbreitete online Nutzung der Texte zu einer Steigerung der Verkäufe der gedruckten Bücher führen. Ähnlich könnte es sich mit den Kursbuchungen in der wissenschaftlichen Weiterbildung verhalten. In einer Befragung von 75 Mitgliedsinstitutionen der Deutschen Gesellschaft für wissenschaftliche Weiterbildung und Fernstudium (DGWF) denken auch nur knapp 7% der befragten Einrichtungen, dass durch OER Einnahmen verloren gehen (STEINERT & DEIMANN, 2016).

## 5 Fazit

Das Thema OER bietet zunächst die Möglichkeit, sich überhaupt mit der Entwicklung bzw. Erschließung digitaler Lernmaterialien und deren Qualitätssicherung auseinanderzusetzen. Die Relevanz muss vor dem Hintergrund der Erfahrungen des Jahres 2020 nicht weiter hervorgehoben werden. Das Fallbeispiel zeigt auch, dass immer didaktische, organisationsbezogene, technische, rechtliche und finanzielle Aspekte bei Digitalisierung des Lernens und Lehrens in der Weiterbildung zu betrachten sind.

Ob die Lernmaterialien dann als OER im Sinne der strengen UNSECO-Definition veröffentlicht werden, sei dahingestellt. Openness ist sicherlich nicht das Hauptargument für Weiterbildungseinrichtungen. Ein Bewusstseinswandel in der Breite

der wissenschaftlichen Weiterbildung im Sinne einer Kultur des Teilens oder einer „Open Educational Practice“ ist eher unwahrscheinlich. Zu diesem Schluss kommen auch BLEES et al. (2015) in ihrem Whitepaper zu OER in der Weiterbildung / Erwachsenenbildung. Weiterbildungseinrichtungen werden daran interessiert sein, zumindest die kommerzielle Nutzung ihrer Materialien zu kontrollieren. Dennoch bieten OER neue Möglichkeiten der Kooperation, der Erschließung, Nutzung und der Sichtbarkeit, die im Einzelfall abzuwägen sind. Im Sinne von Bildungspartizipation und Bildungsgerechtigkeit sind OER allemal ein belebendes Instrument.

In der Strategie der Kultusministerkonferenz „Bildung in der digitalen Welt“ wurde festgestellt (KMK, 2017): „Da es bislang keine Geschäftsmodelle zur Refinanzierung gibt, ist eine Weiterentwicklung mit Mehrkosten verbunden. Erforderlich sind daher die Schaffung eines klaren und transparenten Rechtsrahmens sowie Mittel für die Grundversorgung und für Leuchtturmprojekte“ (S. 51). Die OER-Info Initiative<sup>13</sup> hat hier für erste Impulse gesorgt und die Bekanntheit von OER erhöht. Mit den OER-Förderprogrammen auf Landesebene werden nun Infrastrukturen aufgebaut und Materialien entwickelt, denn Voraussetzung für die breite Nutzung von OER und digitalen Lernmaterialien ganz allgemein ist die Verfügbarkeit einer kritischen Masse von zugänglichen Materialien im deutschsprachigen Raum. Die Plattformen müssen einfach zu verstehen und zu bedienen sein. Gerade wenn Praktiker\*innen in der Weiterbildung bereit sind, ihre Materialien anderen zu Verfügung zu stellen, muss die Hürde zur Veröffentlichung auf diesen Portalen möglichst niedrig sein.

Die Herausforderungen und Möglichkeiten der Digitalisierung in der wissenschaftlichen Weiterbildung sind vielfältig. Die Beschäftigung mit OER ist daher lohnend, auch wenn weniger die direkten Benefits als die langfristigen Potenziale im Vordergrund stehen.

### **Acknowledgement**

Der Beitrag wurde gefördert mit Mitteln aus den Digitalisierungsmaßnahmen im Geschäftsbereich des Niedersächsischen Ministeriums für Wissenschaft und Kultur (MWK) Open Educational Resources (OER)-Portal-Niedersachsen (Förderkennzeichen: 56000016-15/UOL).

---

13 <https://open-educational-resources.de/> (Zugriff am 11.02.2021)

## 6 Literaturverzeichnis

**Bellinger, F., & Mayrberger, K.** (2019). Systematic Literature Review zu Open Educational Practices (OEP) in der Hochschule im europäischen Forschungskontext. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 34 (Research and OER), 19–46. <https://doi.org/10.21240/mpaed/34/2019.02.18.X>

**Blees, I., Deimann, M., Seipel, H., Hirschmann, D., & Muuß-Merholz, J.** (2015). *Whitepaper Open Educational Resources (OER) in Weiterbildung/Erwachsenenbildung*. Bertelsmann Stiftung, Internet & Gesellschaft Co:laboratory, MinD-Stiftung, Open Knowledge Foundation Deutschland und Technologiestiftung Berlin in Kooperation mit open-educational-resources.de – Transferstelle für OER. <https://open-educational-resources.de/wp-content/uploads/Whitepaper-OER-Weiterbildung-2015.pdf>

**Bülow-Schramm, M.** (2006). *Qualitätsmanagement in Bildungseinrichtungen*. Münster: Waxmann.

**Conrad, D.** (2008). Revisiting the Recognition of Prior Learning (RPL): A Reflective Inquiry into RPL Practice in Canada. *Canadian Journal of University Continuing Education*, 34(2), 89–110. Education Source.

**Deimann, M., & Bastiaens, T.** (2010). Potenziale und Hemmnisse freier digitaler Bildungsressourcen – eine Delphi-Studie. *Zeitschrift Für E-Learning, Lernkultur und Bildungstechnologie*, 5(3), 7–18.

**Ehlers, U.-D.** (2011). *Qualität im E-Learning aus Lernaltersicht* (2. Aufl.). Wiesbaden: Verlag für Sozialwissenschaften.

**Hanft, A., & Brinkmann, K.** (Eds.). (2013). *Offene Hochschulen: Die Neuausrichtung der Hochschulen auf Lebenslanges Lernen*. Münster: Waxmann.

**Hattie, J.** (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

**Jung, I., Sasaki, T., & Latchem, C.** (2016). A framework for assessing fitness for purpose in open educational resources. *International Journal of Educational Technology in Higher Education*, 13(1), 3. <https://doi.org/10.1186/s41239-016-0002-5>

**Kerres, M.** (2019). Offene Bildungsressourcen und Open Education: Openness als Bewegung oder als Gefüge von Initiativen? *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 34(Research and OER), 1–18. <https://doi.org/10.21240/mpaed/34/2019.02.17.X>

- Kerres, M.** (2020). Against all odds: Education in Germany coping with Covid-19. *Postdigital Science and Education*, 1–5. <https://doi.org/10.1007/s42438-020-00130-7>
- KMK.** (2013). *Das Bildungswesen in der Bundesrepublik Deutschland 2011/2012. Darstellung der Kompetenzen, Strukturen und bildungspolitischen Entwicklungen für den Informationsaustausch in Europa*. Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. [http://www.kmk.org/fileadmin/doc/Dokumentation/Bildungswesen\\_pdfs/erwachsenenbildung.pdf](http://www.kmk.org/fileadmin/doc/Dokumentation/Bildungswesen_pdfs/erwachsenenbildung.pdf)
- KMK.** (2017). *Strategie der Kultusministerkonferenz „Bildung in der digitalen Welt“*. Kultusministerkonferenz. [https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2018/Digitalstrategie\\_2017\\_mit\\_Weiterbildung.pdf](https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2018/Digitalstrategie_2017_mit_Weiterbildung.pdf)
- Koseoglu, S. & Bozkurt, A.** (2018). An exploratory literature review on open educational practices. *Distance Education*, 39(4), 441–461. <https://doi.org/10.1080/01587919.2018.1520042>
- Mayrberger, K., Zawacki-Richter, O., & Müskens, W.** (2018). *Qualitätsentwicklung von OER – Vorschlag zur Erstellung eines Qualitätssicherungsinstrumentes für OER am Beispiel der Hamburg Open Online University*. Universität Hamburg. <https://doi.org/10.25592/978.3.924330.67.5>
- McGreal, R.** (2019). A Survey of OER Implementations in 13 Higher Education Institutions. *The International Review of Research in Open and Distributed Learning*, 20(5), 141–145. <https://doi.org/10.19173/irrodl.v20i5.4577>
- Müskens, W.** (2006). Pauschale und individuelle Anrechnung beruflicher Kompetenzen auf Hochschulstudiengänge – das Oldenburger Modell. *Hochschule Und Weiterbildung*, 2006(1), 23–30.
- OPAL.** (2011). *Beyond OER: Shifting Focus from Resources to Practices*. Open Educational Quality Initiative (OPAL). [https://duepublico2.uni-due.de/servlets/MCRFileNodeServlet/duepublico\\_derivate\\_00025907/OPALReport2011\\_Beyond\\_OER.pdf](https://duepublico2.uni-due.de/servlets/MCRFileNodeServlet/duepublico_derivate_00025907/OPALReport2011_Beyond_OER.pdf)
- Orr, D., Neumann, J., & Muuß-Merholz, J.** (2017). *German OER Practices and Policy – From Bottom-up to Top-down Initiatives*. UNESCO. [https://open-educational-resources.de/wp-content/uploads/UNESCO\\_Report\\_-German\\_OER.pdf](https://open-educational-resources.de/wp-content/uploads/UNESCO_Report_-German_OER.pdf)
- Otto, D.** (2019). Adoption and Diffusion of Open Educational Resources (OER) in Education: A Meta-Analysis of 25 OER-Projects. *International Review of Research in Open & Distributed Learning*, 20(5), 122–140.

- Otto, D.** (2020). Grosse Erwartungen: Die Rolle von Einstellungen bei der Nutzung und Verbreitung von Open Educational Resources. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 2020(Occasional Papers), 21–43. <https://doi.org/10.21240/mpaed/00/2020.02.26.X>
- Rubin, E., Bernath, U., & Parker, M.** (2004). The Master of Distance Education program: A collaboration between the University of Maryland University College and Oldenburg University. *Journal of Asynchronous Learning Networks*, 8(3), 52–70.
- Speicher, L.** (2016). Does publishing a book as open access affect print sales? *TXT*, 2016(1), 124–127. <https://hdl.handle.net/1887/42723>
- Steinert, F., & Deimann, M.** (2016). *Open Educational Resources – Ein Konzept für die wissenschaftliche Weiterbildung?* Unpublished. <http://rgdoi.net/10.13140/RG.2.2.10793.70243>
- Tait, A.** (2008). What are open universities for? *Open Learning: The Journal of Open, Distance and e-Learning*, 23(2), 85–93. <https://doi.org/10.1080/02680510802051871>
- UNESCO** (2019). *Recommendation on Open Educational Resources (OER)*. <https://unesdoc.unesco.org/ark:/48223/pf0000373755/PDF/373755eng.pdf.multi.page=3>
- Universität Hamburg.** (2018). *Projekte der BMBF-Förderung OERInfo 2017/2018: Sonderband zum Fachmagazin Synergie* (284 Seiten) [PDF]. Universität Hamburg. <https://doi.org/10.25592/978.3.924330.64.4>
- Veletsianos, G., & Kimmons, R.** (2012). Assumptions and challenges of open scholarship. *The International Review of Research in Open and Distributed Learning*, 13(4), 166. <https://doi.org/10.19173/irrod.v13i4.1313>
- Wiley, D.** (2013). *What is Open Pedagogy?* <https://opencontent.org/blog/archives/2975>
- Xiao, J.** (2018). On the margins or at the center? Distance education in higher education. *Distance Education*, 39(2), 259–274. <https://doi.org/10.1080/01587919.2018.1429213>
- Yuan, M., & Recker, M.** (2015). Not All Rubrics Are Equal: A Review of Rubrics for Evaluating the Quality of Open Educational Resources. *International Review of Research in Open & Distributed Learning*, 16(5), 16–38.

**Zawacki-Richter, O.** (2020). Halb zog sie ihn, halb sank er hin ... Covid-19 als Chance für die Digitalisierung von Studium und Lehre? *Das Hochschulwesen*, 68(4+5), 1–9.

**Zawacki-Richter, O., Conrad, D., Bozkurt, A., Aydin, C. H., Jung, I., Kondakci, Y., Stöter, J., Veletsianos, G., Bedenlier, S., Bond, A. J., Blaschke, L., Bronens, A., Dolch, C., Kalz, M., Loglo, F., Mayrberger, K., Müskens, W., Naidu, S., Prinsloo, P., & Xiao, J.** (2020). Elements of Open Education: An Invitation to Future Research. *International Review of Research in Open and Distributed Learning*, 21(3), 319–334.

**Zawacki-Richter, O., & Mayrberger, K.** (2017). *Qualität von OER: Internationale Bestandsaufnahme von Instrumenten zur Qualitätssicherung von Open Educational Resources (OER) – Schritte zu einem deutschen Modell am Beispiel der Hamburg Open Online University*. Hamburg Open Online University. <https://www.synergie.uni-hamburg.de/media/sonderbaende/qualitaet-von-oer-2017.pdf>

**Zawacki-Richter, O., & Stöter, J.** (2019). Angebotsformen des Fernstudiums mit digitalen Medien in der wissenschaftlichen Weiterbildung. In W. Jütte & M. Rohs (Hrsg.), *Handbuch Wissenschaftliche Weiterbildung* (S. 1–16). Wiesbaden: Springer Fachmedien. [https://doi.org/10.1007/978-3-658-17674-7\\_15-1](https://doi.org/10.1007/978-3-658-17674-7_15-1)

## Autor\*innen



Prof. Dr. Olaf ZAWACKI-RICHTER || Carl von Ossietzky Universität Oldenburg, Institut für Pädagogik, Center for Open Education Research (COER) ||

Ammerländer Heerstraße 136, D-26111 Oldenburg

<https://uol.de/coer/coer-members/dr-olaf-zawacki-richter>

[olaf.zawacki.richter@uni-oldenburg.de](mailto:olaf.zawacki.richter@uni-oldenburg.de)



PD. Dr. Wolfgang MÜSKENS || Carl von Ossietzky Universität Oldenburg, Institut für Pädagogik, Center for Open Education Research (COER) ||

Ammerländer Heerstraße 136, D-26111 Oldenburg

<https://uol.de/coer/coer-members/dr-wolfgang-mueskens>

[wolfgang.mueskens@uni-oldenburg.de](mailto:wolfgang.mueskens@uni-oldenburg.de)



Nadine DEMBSKI || Carl von Ossietzky Universität Oldenburg, Center für lebenslanges Lernen (C3L) || Ammerländer Heerstraße 136, D-26111 Oldenburg

<https://uol.de/c3l/>

[nadine.dembski@uni-oldenburg.de](mailto:nadine.dembski@uni-oldenburg.de)



Sonja LÜBBEN || Carl von Ossietzky Universität Oldenburg, Center für lebenslanges Lernen (C3L) || Ammerländer Heerstraße 136, D-26111 Oldenburg

<https://uol.de/c3l/>

[sonja.luebben@uni-oldenburg.de](mailto:sonja.luebben@uni-oldenburg.de)





9 783755 740636

<fnma>

[www.zfhe.at](http://www.zfhe.at)